European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

College admissions with ties and common quotas: Integer programming approach☆

Kolos Csaba Ágoston[a], Péter Biró[a,b,1,*], Endre Kováts[c], Zsuzsanna Jankó[a,b]

[a] Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest, H-1093, Fővám tér 13-15., Budapest, Hungary
[b] Institute of Economics, Research Centre for Economic and Regional Studies, Hungarian Academy of Sciences, H-1097, Tóth Kálmán u. 4., Budapest, Hungary
[c] Budapest University of Technology and Economics, H-1111, Műegyetem rakpart 3., Budapest, Hungary

ARTICLE INFO

Article history:
Received 28 January 2020
Accepted 19 August 2021
Available online xxx

Keywords:
Assignment
Stable matching
College admission
Distributional constraints
Integer programming

ABSTRACT

Admission to universities is organised in a centralised scheme in Hungary. In this paper we investigate two major specialities of this application: ties and common quotas. A tie occur when some students have the same score at a programme. If not enough seats are available for the last tied group of applicants at a programme then there are three reasonable policies used in practice: 1) all must be rejected, as in Hungary 2) all can be accepted, as in Chile 3) a lottery decides which students are accepted from this group, as in Ireland. Even though student-optimal stable matchings can be computed efficiently for each of the above three cases, we developed (mixed) integer programming (IP) formulations for solving these problems, and compared the solutions obtained by the three policies for a real instance of the Hungarian application from 2008. In the case of Hungary common quotas arise from the faculty quotas imposed on their programmes and from the national quotas set for state-financed students in each subject. The overlapping structure of common quotas makes the computational problem of finding a stable solution NP-hard, even for strict rankings. In the case of ties and common quotas we propose two reasonable stable solution concepts for the Hungarian and Chilean policies. We developed (mixed) IP formulations for solving these stable matching problems and tested their performance on the large scale real instance from 2008 and also for one from 2009 under two different assumptions. We demonstrate that the most general case is also solvable in practice by IP technique.

1. Introduction

Gale and Shapley gave a standard model for college admissions (Gale & Shapley, 1962), and suggested stable matching for its solution. Intuitively, a matching is stable if an application to a college is rejected because the college is already full with higher ranked students. Gale and Shapley showed that a stable matching can always be found by the *deferred-acceptance algorithm*, which runs in linear time in the number of applications, see e.g. Manlove (2013). Moreover, the student-oriented variant results in a student-optimal stable matching, meaning that no student could get a better assignment in any other stable matching. The theory of stable matchings has intensively been studied since 1962 by mathematicians/computer scientists (see e.g. Manlove, 2013) and economists/game theorists (see e.g. Roth & Sotomayor, 1990). The Gale-Shapley algorithm has also been used in practice all around the world (Biró, 2017), first in 1952 in the US resident allocation programme, called NRMP (Roth, 1984), then also in school choice, e.g. in Boston (Abdulkadiroğlu, Pathak, & Roth, 2005a) and New York (Abdulkadiroğlu, Pathak, Roth, & Sönmez, 2005b). In Hungary, the national admission scheme for secondary schools follows the original Gale-Shapley model and algorithm (Biró, 2014a), and the higher education admission scheme also uses a heuristic solution based on the Gale-Shapley algorithm (Biró, 2014b).

The Hungarian higher education admission scheme have at least four important special features: presence of ties, lower and common quotas, and paired applications. The students submit preference list on the university programmes they apply to, and

☆ Earlier results of this paper have been presented in two conference papers [5,6].

* Corresponding author.
E-mail addresses: kolos.agoston@uni-corvinus.hu (K.C. Ágoston), peter.biro@krtk.mta.hu (P. Biró), endre.kovats.92@gmail.com (E. Kováts), zsuzsanna.janko@uni-corvinus.hu (Z. Jankó).

[1] Supported by the Hungarian Academy of Sciences under its Momentum Programme (*Engineering Economics in Matching Markets*, no. LP2021-2), and by the Hungarian Scientific Research Fund – OTKA (no. K129086).

https://doi.org/10.1016/j.ejor.2021.08.033

are ranked according to their scores, which are of integer values currently in the range of [0,500] higher score meaning better performance. The solution by the coordinating agency is announced in terms of cutoff scores to be understood as follows: every student is admitted to the best programme of her preference where she achieved the cutoff score. A tie can occur when two or more students have the very same score at a programme they apply for. Ties are never broken in Hungary, either all or none of the students in the tie are admitted, depending on the cutoff score. Lower quotas are minimum requirements for the number of admitted students for each programme, which are set by the universities to make the education economical. Applications by the students also include the contract term of the study, i.e., whether their study is funded by the state or privately. For every programme there is a common upper bound for the number of admitted students under any contract term, and there are also nationwide common quotas for the number of students getting state funds in each subject (e.g., Chemistry). Finally, the students applying for teachers' programmes should apply for pairs of programmes (such as Math-Physics). The latter feature was re-introduced in the application in 2010, but it was not present in 2008 and 2009, the years our analysis bears on in this paper. Further details of the application can be found in Biró (2014b).

Each of the three special features (lower and common quotas and paired applications) makes the problem NP-hard (Biró, Fleiner, Irving, & Manlove, 2010), only the case of ties is resolvable efficiently (Biró & Kiselgof, 2015). In a recent paper (Ágoston, Biró, & McBride, 2016) Ágoston et al. studied the usage of integer programming techniques for finding stable solutions with regard to each of these four special features separately, and they solved the case of lower quotas for the real instance of 2008. We refer to this instance as 2008-Educatio instance, which was provided for research purpose by the coordinating governmental agency called Educatio kht, and contained all the relevant upper and lower quotas in an anonym dataset. In this follow-up work we develop and test new IP formulations for the case of ties and common quotas separately and then also for the case when both features are present. So, the ultimate goal of this work was to suggest a solution concept for the college admission problem where ties and common quotas are also present, and to provide integer programming formulations that are suitable to compute this solution for large scale applications, such as the Hungarian university admission scheme with over 100,000 students.

The presence of ties and equal treatment policy (i.e., not breaking the ties) is also a feature in the Chilean university admissions (Rios, Larroucau, Parra, & Cominetti, 2014). However, the policy used there is more permissive than the Hungarian one, since if two students with the same score are competing for the last seat at a programme then they are both accepted in Chile, but both rejected in Hungary, whilst a random tie-breaking is used in Ireland (Chen, 2012) to decide which student will be admitted. These solution concepts have been studied theoretically in Biró & Kiselgof (2015) under the name of H-stability and L-stability. The intuitive result proved in that paper is that when student-optimal stable solutions are compared for the same instance then the cutoff scores are at least as high in Ireland as in Chile, and at least as high in Hungary as in Ireland. So the students are always getting the worst assignments in Hungary, a better assignment in Ireland, and the best one in Chile. In this paper we quantify these differences on the Hungarian university admission instances from 2008 and 2009, presented in Section 5.

Common quotas are also present in many other applications. A recent paper (Baswana, Chakrabarti, Chandran, Kanoria, & Patange, 2019) describes the admission to Engineering Colleges in India, where common quotas are used for different, possibly overlapping types, just as in the Hungarian case. This means that a stable solution may not exist and the problem is NP-hard (Biró et al., 2010), thus the authors have proposed a heuristic algorithm. Interestingly the authors were aware of the possibility of using IP solutions for this problem, as described in Ágoston et al. (2016) for the Hungarian case[2], but they decided not to use that approach because of the possibly long run time. In this paper we demonstrate the case of common quotas is tractable for large instances, even if the quotas are overlapping and the problem is further complicated by the presence of ties, as in the Hungarian case. The Indian applications have also been studied by Sönmez & Yenmez (2019a,b), where the case of nested set systems have been proved to be solvable by a generalised deferred-acceptance algorithm, which corresponds to the finding of Biró et al. (2010) on the Hungarian college admissions. Furthermore, the same kind of requirements are implemented in college admission schemes with affirmative action, such as the Brazilian college admission system (Aygün & Bo, 2013).

Similar distributional requirements are present for the Israeli Mechinot gap-year programs (Gonczarowski, Kovalio, Nisan, & Romm, 2019), where the authors developed and implemented a new Gale-Shapley type heuristic solution for the application. Ágoston, Biró, & Szántó (2018b) used integer programming techniques for allocating students to companies at CEMS universities under complex distributional constraints with respect to the types of students. Distributional constraints are present in school choice programmes as well, where the decision makers want to control the socio-ethnical distribution of the students (Abdulkadiroğlu, 2005; Abdulkadiroğlu & Ehlers, 2007; Bo, 2016; Echenique & Yenmez, 2015; Ehlers, Hafalir, Yenmez, & Yildirim, 2014). Another well-documented case is the Japanese resident allocation, where the government wants to ensure that the doctors are evenly distributed across the country. They imposed lower quotas on the number of doctors allocated in each region (Goto, Kojima, Kurata, Tamura, & Yokoo, 2017; Kamada & Kojima, 2014; 2017a; 2017b).

Assignments problems are extensively studied in the OR literature (see e.g. Pentico, 2007). There are many examples of practical matching problems, such as papers assignment to reviewers (Garg, Kavitha, Kumar, Mehlhorn, & Mestre, 2010), course allocation (Diebold & Bichler, 2017), marriage assignment (Cao, Fragniére, Gautier, Sapin, & Widmer, 2010) and kidney exchanges (Biró, van de Klundert, & Manlove, 2019). However, the usage of integer programming techniques is relatively new for two-sided matching markets under preferences. This may well be caused by the good performance of the Gale-Shapley type heuristics in practice (see e.g. Roth & Peranson, 1999). With their short run times, they apparently have been preferred over integer programming approaches to solve the sometimes large instances. Besides a previous paper (Ágoston et al., 2016) motivated by the Hungarian university admissions, there were only a couple of studies in this direction for finding maximum size weakly stable matchings in resident allocation problem with ties (Delorme et al., 2019; Kwanashie & Manlove, 2014), for finding stable matching in the presence of couples (Biró, McBride, & Manlove, 2014), and under distributional constraints (Ágoston et al., 2018b).

The paper most closely related to our work is the recent study of Delorme et al. (2019), where IP techniques have been developed and tested to solve a two-sided stable matching problem in a real application, pairing children with adoptive families. Our IP formulations for solving the classical college admissions problem (in their terminology, the Hospitals/Residents problem) are very similar. We also find that introduction of additional variables and

[2] The authors wrote "In principle, one could appeal to the integer programming method devised by Biró and McBride (2014) for this problem, that finds a stable outcome when it exists. However, such an approach was untenable in practice due to complexity, relative opaqueness, and the likelihood of an unreasonably large run time on our large problem."

binary cutoff scores variables drastically improves the efficiency of the IP solution. Thereafter they focus on the NP-hard problem of finding a maximum size weakly stable matching in case of ties, whilst we investigate the Hungarian equal treatment policy when considering the ties and also the feature of common quotas, both present in the Hungarian application. The presence of common quotas makes the problem NP-hard. We test our IP formulations on the 2008 and 2009 instances of Hungarian university admission.

Real instances analysed

Besides the 2008-Educatio instance, we have also had access to the Hungarian university admission data from another source, the KRTK Databank for the years 2001–2017. However the latter instances do not contain capacity constraints, and the identifiers for the programmes also differ from the 2008-Educatio instance. Nevertheless, using the 2008-Education instance, we extended our computational analyses for the 2009-KRTK instance under two reasonable assumptions after linking the 2008-Education and 2008-KRTK instances and then the 2008-KRTK and 2009-KRTK instances.[3] The linkage of the two 2008 instances involved matching of the students and programmes of the two instances. This was not straightforward due to some limitations of the instances (e.g. the KRTK instance contained only the first six applications of each student) and the possibility that the two instances reflect different snapshots of the applications. Nevertheless, approximately 98,5% of the programmes have been identified. Then we also needed to match the programmes in the 2008 and 2009 KRTK instances, which was also non-trivial due to the changes in the list of programmes offered (sometimes only the name of the university or the faculty has changed, but the programmes remained essentially the same, which required manual checks). When the linkage between the programmes of the 2008-Educatio and 2009-KRTK instances were ready, we added the capacity constraints for the 2009-KRTK instance in two reasonable ways: a) we used the same constraints as in the 2008-Educatio instance, b) we used the number of admitted students in 2009 for all programmes and common quotas identified from the 2008-Educatio instance. We refer to these two cases as 2009-KRTK-previous and 2009-KRTK-admitted, respectively.

Regarding the main statistics of the instances, in the 2008-Educatio instance we have 81,427 applicants, 353,618 applications, 3298 programmes, 2275 faculty quotas, and 206 national common quotas. Whilst in the 2009-KRTK instance we have 105,739 applicants, 310,346 applications, 2992 programmes, 1828 faculty quotas, and 197 national common quotas.

Our contribution

Our research is a follow-up of the work of Ágoston et al. (2016), where the same application, the Hungarian university admission was studied with its four special features: ties, common quotas, lower quotas and paired applications. In Ágoston et al. (2016) the special features were considered one-by-one and their main result was a practically tractable IP solution for the NP-hard case of lower quotas, demonstrated on the 2008-Education instance. In this paper we continue the investigation and first we look more deeply into the classical college admission problem, where we compare several (mixed) IP formulations. The cutoff score formulation (already described in Ágoston et al. (2016)) turns out to be viable to solve for the 2008-Educatio instance even without any preprocessing. For the still efficiently solvable case of ties we find that the new binary cutoff formulation (that is similar to the IP suggested in Delorme et al. (2019)) performs the best among the IP-s studied. We then compare the solutions of the Hungarian, Irish and Chilean policies for the case of ties. Confirming the theories described in Biró & Kiselgof (2015), we find that indeed the student-optimal cutoff scores are always the highest in Hungary, followed by the Irish cutoffs and the lowest in Chile, if considered for the same instance. Finally, we define stability through cutoff scores for the case of ties and common quotas with respect to the Hungarian and Chilean policies, and we propose IP formulations with binary cutoff score variables. We find that these IP formulations work well for the 2008-Educatio instance. We compare the solutions with respect to the Hungarian, Irish and Chilean policies. We also extended the computational analyses for the 2009-KRTK instance after linking the 2008-Education and 2008-KRTK instances as well as the 2008 and 2009 KRTK instances under two assumptions: a) by using the 2008 quotas for 2009 everywhere, and b) setting the quotas equal to the number of students admitted. In order to speed up the computations we introduced a preprocessing phase that fixes some variables in the IP model corresponding to students' applications that are either surely accepted or surely rejected in the stable solutions.

Layout of the paper

In Section 2, we start by investigating the basic Gale-Shapley case and testing different IP formulations for a simplified instance of the 2008-Education instance. We find that the cutoff formulations perform better than the standard ones regarding their run time. In Section 3, we consider the special feature of ties under the Hungarian policy, where the quotas are strictly obeyed, so the last group of students with the same score (that cannot fit in the quota) is rejected. Here we observe that the cutoff formulation with binary variables outperforms the cutoff formulation with continuous variables. Then, in Section 4, we describe IP formulations also for the Chilean policy, where the last group of students is still accepted (without whom there remains an empty seat, but with whom the quota may be violated). We compare the results obtained for the Hungarian, Irish and Chilean policies. Then we turn our attention to common quotas, which are present in the Hungarian application in a structure that make the problem NP-hard to solve (Biró et al., 2010). We test different IP-s for solving the problem under strict preferences in Section 5. Finally, in Section 6, we tackle the real case when both ties and common quotas are present. We develop IP-s again for both the Hungarian and Chilean policies and we compare the results for both the 2008-Educatio and the 2009-KRTK instances. We conclude in Section 7.

2. The Gale-Shapley model

In this section we provide various IP formulations for the classical Gale-Shapley college admission model and then we test these formulations on the 2008-Educatio instance.

2.1. Definitions and preliminaries

In the classical college admissions problem by Gale & Shapley (1962) the students are assigned to colleges.[4] In the following we will refer to the two sets as *applicants* $A = \{a_1, \dots, a_n\}$ and *colleges* $C = \{c_1, \dots c_m\}$. Throughout the manuscript, we use the convention of $i = 1, \dots, n$ and $j = 1, \dots, m$. Let u_j denote the upper quota of

[3] This data matching challenge was conducted as part of a student project, the details are available in a BSc thesis upon request.

[4] 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.

college c_j. Regarding the preferences, we assume that the applicants provide strict rankings over the colleges, where r_{ij} denotes the rank of the application (a_i, c_j) in applicant a_i's preference list. We suppose that the students are ranked according to their scores at the colleges, so college c_j ranks applicant a_i according to her score s_{ij} at c_j, where the greater the score the more preferred is the student by the college. Let $E \subseteq A \times C$ denote the set of applications. A *matching* is a set of applications, where each student is admitted to at most one college and each college has at most as many assignees as its quota. For a matching M let $M(a_i)$ denote the college where a_i is admitted to (or $\emptyset$ if a_i is not allocated to any college) and let $M(c_j)$ denote the set of applicants admitted to c_j in M. Thus the feasibility of a matching $M \subset E$ means that for every applicant a_i, $|M(a_i)| \leq 1$ and for every college c_j, $|M(c_j)| \leq u_j$. A matching $M \subset E$ is *stable* if for any application (a_i, c_j) outside M either a_i prefers $M(a_i)$ to c_j or c_j filled its seats with u_j applicants who all have higher scores than a_i has. The deferred-acceptance algorithm of Gale and Shapley provides a student-optimal stable matching in linear time (Gale & Shapley, 1962).

The notion of *cutoff scores* is important for both the classical Gale-Shapley model and its generalisations with ties and common quotas. Let t_j denote the cutoff score of college c_j and let $\mathbf{t}$ denote a set of cutoff scores for all colleges. A student a_i is *admissible* to a college c_j with cutoff score t_j if $s_{ij} \geq t_j$. We say that matching M is implied by cutoff scores $\mathbf{t}$ if every student is admitted to the most preferred college in her list, where she is admissible (i.e., achieved the cutoff score). We say that a set of cutoff scores $\mathbf{t}$ *corresponds* to a matching M if $\mathbf{t}$ implies M. For a matching M an applicant a_i has *justified envy* towards another applicant a_k at college c_j if $M(a_k) = c_j$, a_i prefers c_j to $M(a_i)$ and a_i is ranked higher than a_k at c_j (i.e. $s_{ij} > s_{kj}$). A matching with no justified envy is called *envy-free* (see Wu & Roth, 2018 and Yokoi, 2020).

It is not hard to see that a matching is envy-free if and only if it is implied by some cutoff scores (Ágoston & Biró, 2017). Note that an envy-free matching can be wasteful in the sense that it can leave many desired seats empty (in fact the empty matching is also envy-free). More precisely, when a student a_i prefers c_j to $M(a_i)$ and c_j is not saturated (i.e. $|M(c_j)| < u_j$) then we say that M is *wasteful*. By definition it follows that a matching is stable if and only if it is envy-free and non-wasteful (see also Azevedo & Leshno, 2016). To achieve non-wastefulness we can require the cutoff of any unsaturated college to be minimum (zero in our case). Alternatively, we may require that no cutoff score be decreased without violating the quota of that college, while keeping the other cutoff scores. Furthermore, we may also satisfy the latter condition by ensuring that we select the student-optimal envy-free matching, which is the same as the student-optimal stable matching (Wu & Roth, 2018). To obtain this solution we only need to use an appropriate objective function. We will use the above described connections when developing our IPs.

2.2. IP formulations

Here we will describe three different formulations.

The Baïou-Balinski formulation

First we describe the basic IP formulation for the Gale-Shapley model, proposed in Baïou & Balinski (2000). All of our formulations are based on the binary variables corresponding to applications, where $x_{ij} = 1$ denotes that the application (a_i, c_j) is accepted in the solution (and $x_{ij} = 0$ denotes that it is not). The feasibility of a matching can be ensured with the following two sets of constraints, which are part of all our IPs.

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

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

Note that (1) implies that no applicant can be assigned to more than one college, whereas (2) implies that the upper quotas of the colleges are respected.

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

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

Note that for each $(a_i, c_j) \in E$, if a_i is matched to c_j or to a more preferred college then the first term ensures 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.

$$\min \sum_{(a_i,c_j)\in E} r_{ij} \cdot x_{ij} \tag{4}$$

We abbreviate this formulation based on constraints (1), (2) and (3), and objective function (4) as SO-BB (referring to *student optimal Baïou-Balinski* model). This IP results in the student-optimal stable matching.

The cutoff score formulation

For each college c_j we define a nonnegative real variable t_j denoting its cutoff score.

$$t_j \leq \left(1 - x_{ij}\right) \cdot (\bar{s} + 1) + s_{ij} \text{ for each } (a_i, c_j) \in E \tag{5}$$

and

$$s_{ij} + \epsilon \leq t_j + \left(\sum_{k:r_{ik}\leq r_{ij}} x_{ik}\right) \cdot (\bar{s} + 1) \text{ for each } (a_i, c_j) \in E \tag{6}$$

where $\bar{s}$ is an upper bound for the scores (currently 500 in Hungary) and ϵ is a small positive number.[5] Here (5) implies that if a student a_i is admitted to college c_j then her score (s_{ij}) has reached the cutoff score. The second Eq. (6) ensures the envy-freeness, namely that if a_i is not admitted to c_j or to any better according to her preference then it must be the case that she has not reached the cutoff at c_j. Thus these two sets of conditions create the connection between the cutoff scores and the matching, ensuring that the matching implied by the cutoff scores is envy-free.

To require stability of the matching we need to rule out the possibility of blocking with an empty seat (i.e. wastefulness). This can be achieved by forcing the cutoff score of unsaturated colleges to be minimum (i.e. zero in our case) by the following constraints, where f_j is a binary variable indicating whether c_j rejects any student in the solution.

$$f_j \cdot u_j \leq \sum_{(a_i,c_j)\in E: c_j\in C} x_{ij} \qquad \forall c_j \in C \tag{7}$$

and

$$t_j \leq f_j(\bar{s} + 1) \qquad \forall c_j \in C \tag{8}$$

Our second IP is then constructed from feasibility constraints (1), (2), cutoff score constraints (5), (6), non-wastefulness constraints (7), (8), and the objective function (4) enforcing student-optimality. We abbreviate this IP as SO-NW-CUT, referring to *student-optimal non-wasteful cutoff scores*.

[5] Note in Ágoston et al. (2016) 1 was used instead of ϵ, but we found that the latter choice makes the constraints tighter and the solution more efficient.

As an alternative, we can drop the non-wastefulness constraints and enforce stability directly by obtaining the student-optimal envy-free matching by using either of the following objective functions.

$$\min \sum_{c_j \in C} t_j \tag{9}$$

or

$$\max \sum_{(a_i,c_j)\in E} (K - r_{ij}) \cdot x_{ij} \tag{10}$$

with a large enough constant K. When combined with the feasibility constraints (1), (2), and cutoff score constraints (5), (6), we abbreviate the IP using objective function (9) as MIN-CUT, referring to *minimum cutoff scores*. Likewise, when combined with the feasibility constraints (1), (2), and cutoff score constraints (5), (6), we abbreviate the IP using objective function (10) as MSMR-CUT, referring to *maximum size minimum rank cutoff scores*. Note that as explained earlier both MIN-CUT and MSMR-CUT will output the student-optimal stable matching.

The binary cutoff score formulation

We can make the cutoff formulations discrete by replacing the continuous cutoff variables by binary variables, as follows. For a college c_j, let S_j denote the set of scores the students have there, i.e. $S_j = \{s_{ij} : (a_i, c_j) \in E\}$. Suppose also that the elements of S_j are sorted in an increasing order, so $S_j = \{s_j^1, s_j^2, \dots, s_j^m\}$, where $s_j^k < s_j^{k+1}$. For each college c_j, let us now introduce $|S_j|$ binary cutoff variables: $t_j^1, t_j^2, \dots, t_j^m$ with the following constraints.

$$x_{ij} \le t_j^k \text{ for each } (a_i, c_j) \in E, s_{ij} = s_j^k \tag{11}$$

and

$$t_j^k \le t_j^{k+1} \text{ for each } k = 1.(|S_j| - 1) \tag{12}$$

Here, $t_j^k = 0$ means that the cutoff score at c_j is greater than s_j^k. Furthermore, (12) ensures the monotonicity of the binary cutoff variables and (11) requires that an application can only be accepted if the cutoff score is reached, corresponding to the continuous constraint (5). Regarding the second continuous constraint (6) we add the following simpler equations.

$$1 \le \sum_{h: r_{ih} \le r_{ij}} x_{ih} + (1 - t_j^k) \text{ for each } (a_i, c_j) \in E, s_{ij} = s_j^k \tag{13}$$

Therefore, constraints (11), (12) and (13) replace (5) and (6), and together with the feasibility constraints (1), (2) they make the link between the binary cutoff scores and the envy-free matchings.

To achieve stability, we can use the same techniques as in the continuous case, with slightly modified conditions and objective functions.

As the first IP, instead of using Eqs. (7) and (8), we can enforce the cutoff score being zero for each unfilled college c_j with the following constraint.

$$(1 - t_j^1) \cdot u_j \le \sum_{(a_i,c_j)\in E: c_j \in C} x_{ij} \qquad \forall c_j \in C \tag{14}$$

The corresponding binary IP is then constructed from feasibility constraints (1), (2), cutoff score constraints (11), (12) and (13), non-wastefulness constraints (14), and the objective function (4) that enforces student-optimality. We abbreviate this IP as SO-NW-BIN-CUT, referring to *student-optimal non-wasteful binary cutoff scores*.

Alternatively, we can drop again the non-wastefulness constraints and enforce stability directly by obtaining the student-optimal envy-free matching by using either the following objective function

$$\max \sum_{c_j \in C, k = 1.|S_j|} t_j^k \tag{15}$$

or objective function (10). Combined with feasibility constraints (1), (2), and binary cutoff score constraints (11), (12) and (13) we obtain two IPs, the MIN-BIN-CUT and MSMR-BIN-CUT, referring to *minimum binary cutoff scores* and *maximum size minimum rank binary cutoff scores*, both resulting in the student-optimal stable matching.

Envy-free formulation

It is also possible to enforce envy-freeness without using cutoff scores, as explained in Ágoston & Biró (2017), by using the following constraints.

$$\sum_{k: r_{ik} \le r_{ij}} x_{ik} \ge x_{hj} \quad \forall (a_i, c_j), (a_h, c_j) \in E, s_{ij} \ge s_{hj} \tag{16}$$

The above constraint means that if a student a_h is allocated to college c_j then every student a_i, who has a score at c_j at least as high as a_h has, must also be allocated to c_j or to a better college of her preference. Combined with the feasibility constraints (1), (2), and objective function (10) the solution obtained is the student-optimal stable matching. We abbreviate this formulation as MSMR-EF, referring to *maximum size minimum rank envy-free*.

Summary of formulations

We summarise the constraints needed for all of the (mixed) IP formulations that we tested for the basic Gale-Shapley college admission model in Table 1.

2.3. Computational results

We took the 2008-Educatio instance after breaking the ties randomly, by considering only the faculty quotas and keeping only the highest ranked application of each student for every programme (i.e. the application for either the state funded or the privately funded seat). For the implementation we used AMPL with Gurobi.

As we can see in Table 2, the most efficient formulations used cutoff scores. Even though SO-NW-CUT needed twice as many variables as SO-BB, its runtime was smaller by a magnitude. Note that very similar findings were reported in Delorme et al. (2019). Comparing the continuous and binary cutoff score formulations, SO-NW-CUT and SO-NW-BIN-CUT, we can observe that the continuous version performed slightly better for this basic model. The simple MSMR-EF formulation did not terminate, so we excluded this formulation from further consideration for the more general models.

3. Models with ties

In many nationwide college admission programmes the students are ranked based on their scores, and ties may appear. In Hungary, for instance, the students can obtain integer points between 0 and 500 (the maximum was 144 until 2007), so ties do occur. When ties are present then one way to resolve this issue is to break ties by lotteries, as done in Ireland (so a lucky student with 480 point may be admitted to law school, whilst an unlucky student with the same score may be rejected). However, lotteries are often seen unfair, so in some countries, such as Hungary (Biró & Kiselgof, 2015) and Chile (Rios et al., 2014) equal treatment policies are used, meaning that students with the same score are either all accepted or all rejected. This policy gives way to two reasonable variants when deciding about the last group of students without whom the quota is unfilled and with whom the quota is violated. In the restrictive policy, used in Hungary, the quotas are never violated, so this last group of students is always rejected, whilst in Chile they use a permissive policy and they always admit this last group of students. For instance, if there are three students, a_1, a_2 and a_3, applying to a programme of quota 2 with scores 450,

Table 1
The summary of (mixed) integer programming formulations for the classical Gale-Shapley model.

IP formulations	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
SO-BB	√	√	√	√												
SO-NW-CUT	√	√		√	√	√	√	√								
MIN-CUT	√	√			√	√			√							
MSMR-CUT	√	√			√	√				√						
SO-NW-BIN-CUT	√	√		√							√	√	√	√		
MIN-BIN-CUT	√	√									√	√	√		√	
MSMR-BIN-CUT	√	√								√	√	√	√			
MSMR-EF	√	√								√						√

Table 2
The performance of (mixed) integer programming formulations for the classical Gale-Shapley model.

IP formulations	#variables	#constraints	#non-0 elem.	size(Kb)	run time(s)
SO-BB	287,035	381,115	73,989,595	1,319,663	1139
SO-NW-CUT	291,935	673,050	2,463,808	69,464	81
MIN-CUT	289,485	668,150	2,169,423	64,254	5062
MSMR-CUT	289,485	668,150	2,169,423	69,846	2318
SO-NW-BIN-CUT	574,070	955,185	3,028,078	75,810	107
MIN-BIN-CUT	574,070	952,735	2,738,593	65,657	871
MSMR-BIN-CUT	574,070	952,735	2,738,593	66,467	4325
MSMR-EF	n.a.	n.a.	n.a.	8,667,403	n.a.

443, and 443, respectively then in Hungary only a_1 is admitted, whilst in Chile all three of them. In Ireland, a_1 is admitted and they use a lottery to decide whether a_2 or a_3 will get the last seat.

3.1. Definitions and preliminaries

Stable matchings for the case of ties were defined through the cutoff scores in Biró & Kiselgof (2015). Using cutoff scores in case of ties makes the solution *envy-free*, meaning that no student a_i may be rejected from college c_j if this college admitted another student with a score less than or equal to the score of student a_i. This allocation concept is also called equal treatment policy, since the admission of a student to a programme implies the admission offer to all other students with the same score.[6] Here again, we have the same equivalence between envy-free matchings and matching induced by cutoff scores (Ágoston & Biró, 2017), that we prove here for being self-contained.

Proposition 1. *A matching is envy-free for a college admission problem with ties if and only if it is induced by cutoff scores.*

Proof. Given an envy-free matching M let us set the cutoff score of each college to be the score of the weakest admitted student. These cutoff scores will induce M. In the other direction, any matching induced by cutoff scores is obviously envy-free. □

In this paper we focus on the restrictive policy used in Hungary, where the stability of the matching can be defined by adding a non-wastefulness condition to envy-freeness. Namely, a matching induced by cutoff scores is *stable* if no college can decrease its cutoff score without violating its quota, assuming that the other cutoff scores remain the same. We note that the stability of a matching can equivalently be defined by the lack of a set of blocking applications involving one college and a set of applicants such that this set of applications would be accepted by all parties when compared to the applications of the matching considered. See more about this connection in Fleiner & Jankó (2014).[7]

More formally, for a college c_j and a set of applications $X \subset E$ to this college we define by $Ch_j(X) \subseteq X$ the set of applications selected by c_j. For the case of strict rankings the choice function is simple, if $|X| \le u_j$ then $Ch_j(X) = X$, and if $|X| > u_j$ when c_j selects the top u_j applicants according to their scores. For ties we consider two choice functions, Ch_j^H and Ch_j^C corresponding to the Hungarian restrictive and the Chilean permissive policies. First we note that for $|X| \le u_j$ we have $Ch_j^H(X) = Ch_j^C(X) = X$, the question is what happens for $|X| > u_j$. For cutoff score t_j let $X^{\ge t_j}$ denote the subset of applications in X where the students have score t_j or higher at c_j. In the Hungarian policy $Ch_j^H(X) = X^{\ge t_j}$ such that $|X^{\ge t_j}| \le u_j$ and $|X^{\ge t_j - 1}| > u_j$, thus the number of students selected is never more than the quota, but the cutoff is minimal, i.e., decreasing the cutoff would imply the violation of the quota. In the Chilean policy $Ch_j^C(X) = X^{\ge t_j}$ such that $|X^{\ge t_j}| \ge u_j$ and $|X^{\ge t_j + 1}| < u_j$, thus the number of students selected is at least as much as the quota, but the cutoff is maximal, i.e., increasing the cutoff would imply to have empty seats.

Biró & Kiselgof (2015) proved two main theorems about stable matchings for college admissions with ties. In their first theorem they showed that a student-optimal and a student-pessimal stable matching exist for both the restrictive policy (Hungary) and the permissive policy (Chile), where the cutoff scores are minimal / maximal, thus the matchings are the best / worst for all students, respectively. Furthermore, they also proved the intuitive results that if we compare the student-optimal cutoff scores (or the student-pessimal ones) with respect to the three reasonable policies, namely the Hungarian (restrictive), the Irish (lottery), and the Chilean (permissive), then the Hungarian cutoff scores are always as high for each college as the Chilean cutoff scores and the Irish cutoff scores are in between. When considering the student-

[6] Note that it is also possible to define envy-freeness and stability in a weaker form, where the rejection of a student is allowed when another student with the same score is accepted. These so-called *weakly stable* or *weakly envy-free* matchings are used in the Scottish resident scheme (Irving & Manlove, 2008), and in a project allocation application at CEMS universities (Ágoston et al., 2018b), respectively.

[7] There are many interesting properties that apply differently for the three policies, as demonstrated in Fleiner & Jankó (2014) and Biró & Kiselgof (2015). For instance, the corresponding choice functions are *substitutable* for all the three policies, but the *irrelevance of rejected contracts* property is violated for the Hungarian policy, and the *law of aggregate demand* property is violated for both the Hungarian and Chilean policies. That is why neither of the latter two policies is strategyproof for the students, even though student-optimal solutions do exist.

optimal stable matching, it turns out to be also the student-optimal envy-free matching, as the following proposition describes in Ágoston & Biró (2017).

Proposition 2. *For the college admission problem with ties the student-optimal stable matching is also student-optimal among the envy-free matchings with respect to the Hungarian (or Chilean) equal treatment policy.*

Proof. Assume by way of contradiction that there is an envy-free matching M, where one student gets a better assignment than in the student-optimal stable matching M^s. Without loss of generality we may also assume that M is not Pareto-dominated by another envy-free matching M' with the same property (i.e., where every student would get at least as good an assignment and somebody would get a strictly better assignment). By Proposition 1 we know that M is induced by some cutoff score $\mathbf{t}$. Note that M cannot be stable, since that would contradict to the student-optimality of M^s. Therefore there is at least one college where the cutoff score can be decreased so that new students will be admitted there, but its quota is not violated. Let the new cutoff score be $\mathbf{t}'$ and let the new matching implied be M'. But then M' is also envy-free and Pareto-dominates M, a contradiction. □

3.2. IP formulations

First we describe which of the (M)IP formulations for the classical model work unchanged for the case of ties, and then we give some alternative formulations.

Previous IP-s that work for ties

For the restrictive (Hungarian) equal treatment policy we have to keep the original feasibility constraints (1), (2). Constraints (16) ensure envy-freeness immediately, and we can also achieve envy-freeness by using the same constraints with cutoff scores ((5), (6)), or with binary cutoff scores ((11), (12) and (13)).

To secure stability, we can enforce the selection of the student optimal envy-free matching by using an appropriate objective function, as implied by Proposition 2. Essentially all the student-optimal IP formulations for envy-free matchings that were previously described for the Gale-Shapley model will lead to this solution, namely MIN-CUT, MSMR-CUT, MIN-BIN-CUT, MSMR-BIN-CUT and MSMR-EF (however, we leave out the latter from the simulations due to its bad performance for the basic model).

Alternative formulations

When we want to avoid the inclusion of objective functions, we may also enforce stability directly by adding new variables d_{ij} and constraints, as described in Ágoston et al. (2016). Here d_{ij} is a binary variable showing whether a_i would be admitted to c_j if the cutoff score decreased at c_j by one.

$$d_{ik} \le (1 - x_{ij}) \qquad \text{for each } (a_i, c_j) \in E, (a_i, c_k) \in E, r_{ik} \ge r_{ij} \quad (17)$$

Condition (17) implies that d_{ij} can only be one if student a_i prefers c_j to her current assignment.

$$t_j - 1 \le \left(1 - d_{ij}\right) \cdot (\bar{s} + 1) + s_{ij} \text{ for each } (a_i, c_j) \in E \qquad (18)$$

where (18) is a modification of (5), implying that d_{ij} can only be one if a_i reaches the cutoff score, when decreased by one.

Now, with these new variables we can also formulate the non-wastefulness condition, where f_j will again indicate whether c_j is essentially full, meaning that its cutoff score cannot be decreased without violating its quota. Besides keeping (8), we modify (7) into the following condition.

$$f_j \cdot (u_j + 1) \le \sum_{(a_i, c_j) \in E : c_j \in C} (x_{ij} + d_{ij}) \qquad \forall c_j \in C \qquad (19)$$

To summarise, together with the basic feasibility conditions (1), (2), and cutoff score constraints (5), (6), satisfaction of Eqs. (17), (18), (8), (19) result in a stable matching with respect to the Hungarian equal treatment policy. To find the student-optimal stable matching in this context, we may again use objective function (10). We denote this formulation by SO-H-NW-CUT.

Binary cutoffs

Finally, we can again use binary variables for the cutoffs. Keeping the feasibility constraints (1), (2), cutoff score constraints (11), (12) and (13), we modify the non-wastefulness constraints (14) as follows.

$$(1 - t_j^1) \cdot (u_j + 1) \le \sum_{(a_i, c_j) \in E : c_j \in C} x_{ij} + d_{ij} \qquad \forall c_j \in C \qquad (20)$$

We keep (17) and modify (18) to the following constraints.

$$d_{ij} \le t_i^{k+1} - t_i^k \text{ for each } (a_i, c_j) \in E, s_{ij} = s_j^k \qquad (21)$$

These constraints mean that if $d_{ij} = 1$ then the score of the applicant is just below of the cutoff score. Using objective function (10), with feasibility constraints (1), (2), cutoff score constraints (11), (12) and (13), non-wastefulness constraints (17), (20) and (21), we get the formulation denoted by SO-H-NW-BIN-CUT.

3.3. Simulations

We considered the 2008-Education instance with ties by taking into account only the faculty quotas and keeping only the highest ranked application of each student for every programme (i.e. the application for either a state funded or privately funded seat). The results are summarised in Table 3. We can observe that the best binary IP formulation has outperformed the best continuous formulation.

4. Policy comparison for ties

First we set up an IP formulation for the Chilean policy, and then we compare the solutions obtained by the Hungarian, Irish and Chilean policies in the 2008-Educatio instance.

4.1. IP for the Chilean policy

Recall that here we admit the last group of students, with whom the quota is violated, but without whom some seats remain unfilled. Alternatively, we may require that after decreasing the cutoff score at any college the number of admitted students would be strictly less than its quota. To achieve this in the most effective way, we use a similar formulation as SO-H-NW-CUT for the Hungarian policy.

The cutoff score formulation

Here $\bar{d}_{ij}$ is a binary variable showing that a_i is admitted to c_j, but would be rejected if the cutoff score increased at c_j by one.

$$\bar{d}_{ij} \le x_{ij} \qquad \text{for each } (a_i, c_j) \in E \qquad (22)$$

Conditions (22) imply that $\bar{d}_{ij}$ can only be one if student a_i is admitted to c_j in the actual matching.

$$\left(\bar{d}_{ij} - 1\right) \cdot (\bar{s} + 1) + s_{ij} \le t_j \text{ for each } (a_i, c_j) \in E \qquad (23)$$

where (23) implies that $\bar{d}_{ij}$ will only be one if a_i is rejected if the cutoff increases by one.

Now, with these variables we can also formulate the non-wastefulness condition, where f_j will again indicates whether c_j is essentially full, meaning that there would be empty seats if

Table 3
The performances of (mixed) integer programming formulations for the case of ties.

IP formulations	#variables	#constraints	#non-0 elem.	size(Mb)	run time(s)
MIN-CUT	289,485	668,150	2,169,423	59,694	5247
MSMR-CUT	289,485	668,150	2,169,423	65,286	1460
MIN-BIN-CUT	428,513	807,178	2,447,479	53,548	982
MSMR-BIN-CUT	428,513	807,178	2,447,479	57,106	1362
SO-H-NW-CUT	578,970	1,694,333	4,793,409	114,882	1310
SO-H-NW-BIN-CUT	861,105	1,813,840	5,352,772	118,828	165

Table 4
The comparison of student-optimal (A-opt) and student-pessimal (C-opt) stable matchings for the case of ties under the Hungarian, Irish and Chilean policies.

	size of matching		average rank		average cutoffs		# rejections	
policies	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.
Hungarian	86,195	86,195	1.2979	1.2979	58.3931	58.3931	37,698	37,698
Irish	86,410	86,410	1.2916	1.2916	58.2090	58.2106	36,802	36,804
Chilean	86,614	86,614	1.2844	1.2844	57.2000	57.5200	35,901	35,901

the cutoff score increased by one. Besides keeping (7) and (8), we modify (2) into the following sets of constraints.

$$\sum_{(a_i,c_j)\in E: c_j\in C} (x_{ij} - \overline{d}_{ij}) \le u_j - 1 \qquad \forall c_j \in C \tag{24}$$

In summary, together with the basic applicant-feasibility conditions (1) and cutoff score constraints (5), (6), satisfaction of Eqs. (7), (8), (22), (23), (24) result in a stable matching with respect to the Chilean equal treatment policy. To find the student-optimal stable matching in this context, we may again use objective function (10). Denote this formulation by SO-C-NW-CUT.

Binary cutoffs
As follows we also describe the alternative IP formulation with binary cutoff variables, abbreviated as SO-C-NW-BIN-CUT.

We keep feasibility conditions (1) and (24), as in the SO-C-NW-CUT model, and we also impose (22). Furthermore, we have cutoff score constraints (11), (12) and (13), just like in SO-NW-BIN-CUT and SO-H-NW-BIN-CUT. However, (23) will be now replaced with the following set of conditions.

$$\overline{d}_{ij} \le t_i^k - t_i^{k-1} \text{ for each } (a_i, c_j) \in E, s_{ij} = s_j^k \tag{25}$$

These conditions ensure that $\overline{d}_{ij}$ can only be one if s_{ij} is equal to the current cutoff score.

4.2. Computational results

We conducted the simulation on the 2008-Educatio instance, where we compared the results for the Hungarian, Irish and Chilean policies, as summarised in Table 4. For the student-pessimal solutions, we minimised the objective function of (10). The results indeed follow the theorems of Biró & Kiselgof (2015) regarding the cutoff scores for the three different policies. The improvements from the Hungarian to the Irish and from the Irish to the Chilean matchings are significant for the students. Another interesting fact of the simulation is that for the Hungarian and Chilean policies we observed no difference between the student-optimal and student pessimal solutions, so the stable solutions are unique for both cases. Regarding the Irish policy the difference between the student-optimal and student-pessimal solutions is minor. These findings are in line with previous results on large markets, such as the case of NRMP described in Roth & Peranson (1999).

5. Models with common quotas

Here we consider the case of common quotas for strict rankings first.

5.1. Definitions and preliminaries

In the Hungarian university admission scheme the students can apply for so-called *studies*, where the study programme and the associated financial contract are both specified. The contract can be either state-funded, where the students can study free of charge (but under some strict conditions over the length of the study and their future employment in Hungary), or it can be a privately financed contract, where the students pay a tuition. For each programme, there is a common *faculty quota* for the number of admitted students irrespective of their contract terms, and there is also a *nationwide quota* for the number of students in each subject area that can be admitted to any programme with state-fund. For instance, the government may decide that they cover the studies of 3000 computer science students in Hungary, whilst BME (Budapest University of Technology and Economics) can have a faculty quota of 500 for CS students, implying for both state funded and privately funded contracts together.

The rejection of a student a_i to a state-funded study c_j is considered fair, if either the faculty quota is filled with higher ranked students applying for the programme of c_j, or the nationwide quota for the subject of c_j is filled with higher ranked students. Regarding the applications to privately-funded studies, only the binding faculty quotas can result in rejections.

This feature of the Hungarian application scheme motivated the study of the college admission problem with common quotas, defined and studied first in Biró et al. (2010). In this model, for each set of colleges $C_p \subseteq C$, there may be an upper quota u_p for the number of applications accepted by colleges belonging to C_p. Let C^c denote the sets of colleges that have common upper quotas (which also includes every individual college, since they also have upper quotas). A matching is *feasible* if no common upper quota is violated. A feasible matching is *stable* if for each rejected application by applicant a_i to college c_j there exists at least one common quota for a set of colleges C_p, such that $c_j \in C_p$ and C_p is filled with applicants ranked higher than a_i. In this definition we must also assume that a set of colleges C_j with common quota has a *master* ranking over the students applying to any college in that set. Therefore, considering the scores of the applicants, if two colleges c_j and c_h both belong to a common set C_p then the score of an applicant must be the same at both c_j and c_h, so essentially the students have scores for the sets of colleges with common quotas.

For a set of colleges C_p we can also define the choice function over a set of applications $X \subset E$, that we denote by Ch_p for strict rankings, and Ch_p^H and Ch_p^C for ties under the Hungarian and Chilean policies, respectively. We suppose that X contains only the best application of each student (if not then we remove all but the most preferred one by the student). Again, we define $X^{\geq t_p}$ as the set of applications with score at least t_p, where t_p is the cutoff score for common quota C_p. Just as for single colleges, the choice function selects here the best feasible set of applications from X, where feasibility is defined according to the Hungarian and the Chilean policies in an analogous way as for single colleges. Thus if $|X| > u_p$ then for the Hungarian policy $Ch^H_p(X) = X^{\geq t_p}$ such that $|X^{\geq t_p}| \leq u_p$ and $|X^{\geq t_j - 1}| > u_p$, and in the Chilean policy $Ch^C_p(X) = X^{\geq t_p}$ such that $|X^{\geq t_p}| \geq u_p$ and $|X^{\geq t_p + 1}| < u_p$.

An alternative definition of stability can be given by introducing cutoff scores for sets of colleges with common quotas. The correspondence between the envy-free matchings and cutoff scores are similar as in the basic model, every student is admitted to the best college according to her preference, where she achieved the cutoff score for every set of colleges with common quota including this college. Non-wastefulness can also be described in a similar way as before, no set of colleges with common quota can decrease its cutoff score without violating its common quota.

Biró et al. (2010) showed that a stable matching may not exist in this model[8], and the problem to decide the existence is NP-hard, even in the special case of the Hungarian university admission system introduced in 2007. However, they also showed that if the set system is *nested* (also called *laminar* or *hierarchic* in the literature), that is if for every pair of sets $C_p, C_q \subseteq C$ either $C_p \subseteq C_q$ or $C_q \subseteq C_p$, then a stable matching always exists and can be computed efficiently. This was the case until 2007 in the Hungarian university admission.

Ágoston et al. (2016) gave an IP formulation for the problem, but they were not able to solve the problem on the 2008-Educatio instance. Below, we will give a slightly modified IP formulation with binary cutoff scores, and demonstrate its applicability for the 2008-Educatio and 2009-KRTK instances, and in addition we further generalise the IP formulations for the case when both ties and common quotas are present.

5.2. IP formulations

First we recall the IP formulation suggested in Ágoston et al. (2016) and then extend it to binary variables. For all formulations in order to take into account the common upper quotas we need the following set of feasibility constraints

$$\sum_{(a_i,c_j)\in E;c_j\in C_p} x_{ij} \leq u_p \qquad \forall C_p \subseteq C \tag{26}$$

Cutoff formulation for common quotas

Introduce a continuous cutoff variable t_p for each common quota C_p. As before, we ensure that an applicant a_i is admissible to a college c_j belonging to C_p if she achieved the cutoff score t_p for every set of colleges C_p college c_j belongs to. That is, conditions (5) change as follows.

$$t_p \leq \left(1 - x_{ij}\right) \cdot (\bar{s} + 1) + s_{ij} \text{ for each } (a_i, c_j) \in E, c_j \in C_p \tag{27}$$

To ensure envy-freeness the following sets of constraints are introduced.

$$s_{ij} + \epsilon \leq t_p + \left(\sum_{k:r_{ik} \leq r_{ij}} x_{ik} + (1 - b^p_j) \right) \cdot (\bar{s} + 1) \quad \forall (a_i, c_j) \in E, c_j \in C_p \tag{28}$$

where b^p_j is a binary variable for each pair of college c_j and common quota C_p such that $c_j \in C_p$. Finally, we require that

$$\sum_{p:c_j\in C_p} b^p_j \geq 1, \qquad \forall c_j \in C. \tag{29}$$

The latter two sets of conditions imply that, if the application (a_i, c_j) is rejected, then there must be at least one common quota for set C_p such that $c_j \in C_p$ and the cutoff score at C_p is higher than the score of a_i there.

For stability, besides envy-freeness we also need to provide non-wastefulness. This can be reached by enforcing the cutoff scores for sets of colleges to be zero if the quota is not fully filled. This can be implemented with the following sets of constraints, where f_p is a binary variable for every set of colleges C_p with common quota.

$$f_p \cdot u_p \leq \sum_{(a_i,c_j)\in E:c_j\in C_p} x_{ij} \qquad \forall C_p \in C^c \tag{30}$$

and

$$t_p \leq f_p(\bar{s} + 1) \qquad \forall C_p \in C^c \tag{31}$$

To summarise, besides the feasibility conditions (1) and (26), the stability is achieved by conditions (27), (28), (29), (30) and (31). Together with objective function (10) we abbreviate this IP as COM-SO-CUT.

Binary cutoffs for common quotas

Instead of the continuous variables t_p for the cutoff of C_p now we introduce binary variables. The corresponding constraints (11), (12) and (13) will be generalised for the case of common quotas as follows.

For a set of colleges C_p, let S_p denote the set of scores the students have, i.e. $S_p = \{s_{ij} : (a_i, c_j) \in E, c_j \in C_p\}$. Suppose also that the elements of S_p are sorted in an increasing order, so $S_p = \{s^1_p, s^2_p, \dots, s^m_p\}$, where $s^k_p < s^{k+1}_p$. For each set of colleges C_p with a common quota, introduce $|S_p|$ binary cutoff variables: $t^1_p, t^2_p, \dots, t^m_p$ with the following constraints.

$$x_{ij} \leq t^k_p \text{ for each } (a_i, c_j) \in E, c_j \in C_p, s_{ij} = s^k_p \tag{32}$$

and

$$t^k_p \leq t^{k+1}_p \text{ for each } k = 1.(|S_p| - 1), C_p \in C^c \tag{33}$$

Here again, $t^k_p = 0$ means that the cutoff score at C_p is greater than s^k_p. Furthermore, (33) ensures the monotonicity of the binary cutoff variables and (32) ensures that an application (a_i, c_j) can only be accepted if the cutoff score is met for each set of colleges C_p college c_j belongs to. Finally, envy-freeness is achieved with the following conditions.

$$1 \leq \sum_{h:r_{ih} \leq r_{ij}} x_{ih} + \sum_{p:c_j\in C_p, s_{ij}=s^k_p} (1 - t^k_p) \text{ for each } (a_i, c_j) \in E, c_j \in C_p \tag{34}$$

The latter set of constraints imply that if a_i is rejected from c_j (i.e., when the first term of the right hand side is zero), then there must be at least one set of colleges C_p such that t^k_p is zero, where $s_{ij} = s^k_p$, meaning that a_i has not reached the cutoff score at C_p, so her rejection is fair.

[8] For clarity we give their example here. Assume there are applicants, a_1, a_2 and a_3, and four colleges c_1, c_2, c_3 and c_4, where c_1 and c_2 have a common quota of 1 and c_2, c_3 have a common quota of 1, whilst c_4 has a simple upper quota of 1. For colleges c_1, c_2 and c_3, the ranking order is a_1, a_2, a_3, while for c_4 it is the opposite. Finally, a_1 has the preference list c_4, c_1, a_2 only applies to c_2, and a_3 has preference list c_3, c_4. One can easily check that none of the possible matchings is stable, as $\{(a_1, c_1), (a_3, c_3)\}$ is blocked by (a_1, c_4), $\{(a_1, c_4), (a_3, c_3)\}$ is blocked by (a_2, c_2), $\{(a_1, c_4), (a_2, c_2)\}$ is blocked by (a_3, c_4), $\{(a_2, c_2), (a_3, c_4)\}$ is blocked by (a_1, c_1), and $\{(a_1, c_1), (a_3, c_4)\}$ is blocked by (a_3, c_3).

For stability, we also have to ensure non-wastefulness, which can be achieved with constraints similar to (14) used in SO-NW-BIN-CUT, as follows.

$$(1 - t_p^1) \cdot u_p \leq \sum_{(a_i,c_j)\in E: c_j \in C_p} x_{ij} \qquad \forall C_p \in C^c \tag{35}$$

To summarise, the binary formulation is composed of feasibility conditions (1) and (26) and stability conditions (32), (33), (34) and (35) with (10) as objective function. For simplicity, denote this IP formulation by COM-Irish, since the further characteristics, denoted by SO-NW-BIN-CUT, will be shared in all of the IP models that we will consider for the case of ties and common quotas in the next section.

This problem variant represents the situation where the ties in the rankings are broken by lottery (as done in Ireland). However, in Hungary and Chile the ties are not broken, which motivates the next, most general setting, where both ties and common quotas are present.

6. Models with ties and common quotas

We learned in our simulations for the case of common quotas with no ties that the best performing IPs are the ones with binary cutoff scores (the IP with continuous variables did not terminate in a reasonable time), therefore, we will only use the latter IP formulation in this section.

When ties and common quotas are both present in the application then it is unclear what would be the most suitable stability definitions. First of all, we have to differentiate between the more restrictive approach used in Hungary and the more permissive method used in Chile, in case of ties with no common quotas. We will also try to ensure that in case of no ties both concepts are equivalent with the model described in the previous section.

6.1. Definitions and preliminaries

We will define the desired solution with cutoff scores, as this provides envy-freeness automatically. The cutoffs are defined for the common quotas, and an application to college c_j may be accepted if the cutoff score is met for every set of colleges C_p with common quota college c_j belongs to. The question is how to define non-wastefulness. By generalising the concepts for the case of ties (with no common quotas) define a solution to be non-wasteful for the Hungarian policy if no cutoff score can be decreased without violating the corresponding common quota. Regarding the Chilean policy, it is required that common quotas are violated only by the last group of students with the same score. This is equivalent to the assumption that after increasing any cutoff score of a set of colleges C_p with common quota, the number of students admitted to C_p would be strictly less than its common quota. We can see that both the Hungarian and the Chilean definition are equivalent to the previous stable matching definition for common quotas if no ties occur.

Since we have seen that the binary cutoff formulations outperformed the continuous formulations, it is enough to consider binary formulations for the Hungarian and Chilean policies here.

6.2. Binary IP formulations for the Hungarian policy

Just as in the previous section, we have binary cutoff variables $\{t_p^1, t_p^2, \ldots, t_p^m\}$ for each set of colleges C_p. The feasibility of the solution is satisfied due to conditions (26). Also, we can establish the correspondence between the cutoff scores and the matching by the same conditions as before, i.e. with (32), (33) and (34).

To achieve non-wastefulness new binary variables are needed, d_i^p for each applicant a_i and set of colleges $C_p \in C^c$, that have value one if a_i is not yet admitted to a college in C_p, but after decreasing the cutoff at C_p by one step (i.e. to the subsequent score group), a_i would be admitted to a college in C_p. For the precise formulations, for each application by a_i to c_j, where $c_j \in C_p$, we also need to introduce a binary variable e_{ij}^p, which can have value one, if a_i is neither admitted to c_j nor to a preferred place, but she would be admitted to c_j if the cutoff at C_p decreased by one.

Non-wastefulness in this model means that no cutoff score can be decreased without violating a common quota. Instead of (17), we have the following constraints.

$$e_{ik}^p \leq (1 - x_{ij}) \qquad \text{for each } (a_i,c_j) \in E,\ (a_i,c_k) \in E,\ c_k \in C_p,\ r_{ik} \geq r_{ij} \tag{36}$$

This implies that e_{ik}^p can only be one if a_i is not admitted to c_k or to a more preferred college. We replace (21) with

$$e_{ij}^p \leq t_p^{k+1} - t_p^k \text{ for each } (a_i,c_j) \in E,\ c_j \in C_p,\ s_{ij} = s_p^k \tag{37}$$

and add

$$e_{ij}^p \leq t_q^k \text{ for each } (a_i,c_j) \in E,\ c_j \in C_q,\ p \neq q,\ s_{ij} = s_q^k \tag{38}$$

These two sets of constraints imply that e_{ij}^p can only be one if a_i is not admitted to c_j, but she would be admitted to c_j if the cutoff at C_p decreased by one, since she also meets the cutoffs of every other set of colleges C_q that include c_j. Finally, we link e_{ij}^p with d_i^p as follows.

$$d_i^p \leq \sum_{j: c_j \in C_p} e_{ij}^p \text{ for each } (a_i,c_j) \in E,\ c_j \in C_p \in C^c \tag{39}$$

The latter conditions make sure that d_i^p can only be one if there exists some college c_j, where a_i would be admitted if the cutoff of C_p decreased by one. Finally, just as in (20), we require the following constraints by setting $x_i^p = \sum_{(a_i,c_j)\in E: c_j \in C_p} x_{ij}$.

$$(1 - t_p^1) \cdot (u_p + 1) \leq \sum_{a_i \in A} x_i^p + d_i^p \text{ for each } C_p \in C^c \tag{40}$$

These constraints are in line with our definition of non-wastefulness, by ensuring that no cutoff score of a set of colleges C_p can be decreased, since otherwise the common quota of C_p would be violated.

To summarise, the IP formulation for the Hungarian variant for the most general case of ties and common quotas, denoted by COM-Hungarian, consists of feasibility conditions (1) and (26) and stability conditions (32), (33), (34), (36), (37), (38), (39) and (40) with (10) as objective function.

6.3. Binary IP formulations for the Chilean policy

The IP formulation for the Chilean policy is similar to the Hungarian policy but somewhat simpler. Now we do not need all variables e_{ij}^p, only those of type $\overline{d}_i^p$. This is so because $\overline{d}_i^p$ can be one if a_i is admitted to a college in C_p and she just met the cutoff, so she would be rejected if the cutoff of C_p decreased. The common quotas may be violated, but only with the last score group admitted. Without them some seats would remain empty in the colleges belonging to the common quota.

To summarise, there are binary cutoff variables $\{t_p^1, t_p^2, \ldots, t_p^m\}$ for each set of colleges C_p, and for each applicant a_i and common quota C_p there is a binary variable $\overline{d}_i^p$ if a_i is applying to any college in C_p. The correspondence between the cutoff scores and the matching is established by conditions (32), (33) and (34). The feasibility of the solution is ensured with a condition similar to (24), as follows

$$\sum_{i:(a_i,c_j)\in E, c_j \in C_p} (x_i^p - \overline{d}_i^p) \leq u_p - 1 \text{ for each } C_p \in C^c \tag{41}$$

Table 5
The performance of the binary cutoff formulations for the case of ties and common quotas under the Hungarian, Irish and Chilean policies.

IP formulations	#variables	#constraints	#non-0 elem.	size(Kb)	run time(s)
COM-Hungarian	1,566,052	3,311,220	11,423,605	191,222,474	158,949.17
COM-Irish	804,576	1,490,331	5,360,107	121,110,852	22,320.28
COM-Chilean	1,114,913	2,389,382	8,331,816	192,978,328	52,962.27

Table 6
The comparison of student-optimal (A-opt) and student-pessimal (C-opt) stable matchings for the case of ties and common quotas under the Hungarian, Irish and Chilean policies for the 2008-Educatio instance.

	size of matching		average rank		# rejections	
policies	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.
Hungarian	81,581	81,427	1.6496	1.6534	81,369	82,043
Irish	81,825	81,825	1.6413	1.6414	80,237	80,239
Chilean	82,082	82,082	1.6337	1.6337	79,176	79,176

These conditions allow the common quotas to be violated, but only for the last score group admitted. Without them the number of students admitted is strictly less than the common quota. For d_i^p we have a condition similar to (22):

$$\overline{d}_i^p \le x_i^p \qquad \text{for each } a_i \in A, C_p \in C^c \text{ s.t. } \exists c_j : (a_i, c_j) \in E, c_j \in C_p \tag{42}$$

Conditions (42) imply that $\overline{d}_i^p$ can only be one if student a_i is admitted to a college c_j belonging to common quota C_p in the matching. Finally, similarly to (25), we need to add the following constraints

$$\overline{d}_i^p \le t_p^k - t_p^{k-1} \text{ for each } (a_i, c_j) \in E, c_j \in C_p, s_{ij} = s_p^k \tag{43}$$

These conditions imply that $\overline{d}_i^p$ can only be one if the score of a_i is equal to the current cutoff score at C_p.

To summarise, the IP formulation for the Chilean variant for the most general case of ties and common quotas, denoted by COM-Chilean, is composed of feasibility conditions (1) and (26) and stability conditions (32), (33), (34), (41), (42) and (43) with (10) as objective function.

6.4. Computational results for the 2008-Educatio instance

We compared the three above described IP formulations first for the 2008-Educatio instance, where the ties and common quotas were included in the data in their actual form. The performance of the three formulations are summarised in Table 5. Although the solver needed about a whole day to terminate[9], the IP was solved for the Hungarian case as well.

The final student-optimal and student-pessimal solutions are shown in Table 6, where the former solutions used (10) as objective function and the latter used the same with minimisation (instead of maximisation). Similarly to the case of ties, it can be observed that there were significant differences between the results across the three policies, the Hungarian providing the highest cutoff scores (and the worst assignments for the students), followed by the Irish policy, and the Chilean turned out to be the most favourable for the students, as expected. Also, we can see that the differences between the student-optimal and student-pessimal solutions are minor (if any) for the Irish and Chilean policies, but a little more significant for the Hungarian policy.[10]

6.5. Improvement by preprocessing

When solving the 2009-KRTK instances we experienced even longer computation times than for the 2008-Educatio instance, so we looked for possibilities for fixing some of the variables. The new preprocessing algorithm was based on the following Lemmas. In the first one we gave a natural condition for the certain acceptable of a first application.

Lemma 3. *Suppose that (a_i, c_j) is the first application of student a_i. For the Hungarian policy, if this application is selected with respect to each relevant common quota from all the applications, i.e., if $(a_i, c_j) \in Ch_p^H(E)$ for each common quota C_p, where $c_j \in C_p$, then (a_i, c_j) is a part of all stable solutions with respect to this policy. Likewise, for the Chilean policy, if $(a_i, c_j) \in Ch_p^C(E)$ for each common quota C_p, where $c_j \in C_p$, then (a_i, c_j) is contained in all stable solutions with respect to this policy.*

Proof. Consider the Hungarian policy first, with choice function Ch_p^H for each common quota C_p, where $c_j \in C_p$. According to our assumption we have $(a_i, c_j) \in Ch_p^H(E)$ for every C_p, where $c_j \in C_p$. Suppose on the contrary that a stable solution would not include (a_i, c_j). It must then be the case that cutoff score t_p is greater than s_{ij} for some common quota C_p, where $c_j \in C_p$. However, this contradicts the minimality of the cutoff scores of a stable solution since t_p can be decreased to s_{ij} because even for E, $|E^{\ge s_{ij}}| \le u_p$ according to our assumption. The proof for the Chilean policy is analogous. □

In the next lemma we give a condition for rejecting an application.

Lemma 4. *Let F_j denote the set of first applications submitted to college c_j. For an application (a_i, c_j), let $X = F_j \cup \{(a_i, c_j)\}$. If $(a_i, c_j) \notin Ch_p^H(X)$ for some C_p, where $c_j \in C_p$, then (a_i, c_j) cannot be included in any stable solution with respect to the Hungarian policy. Likewise, if $(a_i, c_j) \notin Ch_p^C(X)$ for some C_p, where $c_j \in C_p$, then (a_i, c_j) cannot be included in any stable solution with respect to the Chilean policy.*

Proof. Considering the Hungarian policy first, suppose for a contradiction that $(a_i, c_j) \notin Ch_p^H(X)$ for some C_p, where $c_j \in C_p$, but (a_i, c_j) is part of a stable solution. This implies that $t_{p'}$ is less than or equal to s_{ij} for every $C_{p'}$, where $c_j \in C_{p'}$. This means that all the

[9] We used a normal PC with the following parameters: Intel Core i3-8100 CPU, 3,6 Gzh, 8 GB memory, Windows 10 Enterprise operation system, Gurobi 8.1.0 solver.

[10] To give an intuition for the latter result, we give a simple example, where the same phenomenon can be observed. Let us have three students, s_1, s_2 and s_3 with scores 1010 and 9 respectively, and two colleges A and B with individual quota 1 for each, and also with a common quota of 1 for both A and B. Suppose that s_1 and s_2 are applying to A and s_3 is applying to B. For the Irish policy, we run a lottery, and whichever student between s_1 and s_2 is luckier gets admitted to A and both of the other students are both rejected. This is the unique stable solution. For the Chilean case, the unique stable solution is when both s_1 and s_2 are admitted to A, and supported by the cutoff scores of 10 for both A and $\{A, B\}$. However, when considering the Hungarian policy, we have two significantly different stable matchings. In the first, student-optimal one, s_3 is admitted to B and the stable cutoff scores are 11,0, and 0 for A, B and $\{A, B\}$. Meanwhile in the student-pessimal solution everybody is rejected, and the cutoff scores are 0,0 and 11 for A, B and $\{A, B\}$, respectively.

Table 7
The comparison of student-optimal (A-opt) and student-pessimal (C-opt) stable matchings for the case of ties and common quotas under the Hungarian, Irish and Chilean policies for the 2009-KRTK-previous instance.

	size of matching		average rank		# rejections	
policies	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.
Hungarian	78,851	78,778	1.6591	1.6605	120,141	120,360
Irish	79,279	79,279	1.6519	1.6519	118,878	118,878
Chilean	79,764	79,755	1.6412	1.6415	117,117	117,162

Table 8
The comparison of student-optimal (A-opt) and student-pessimal (C-opt) stable matchings for the case of ties and common quotas under the Hungarian, Irish and Chilean policies for the 2009-KRTK-admitted instance.

	size of matching		average rank		# rejections	
policies	A-opt.	C-opt.	A-opt.	C-opt.	A-opt.	C-opt.
Hungarian	73,579	73,560	1.6910	1.6917	131,964	132,013
Irish	74,272	74,272	1.6818	1.6818	130,224	130,224
Chilean	74,823	74,823	1.6715	1.6715	128,405	128,405

first applications at c_j with score t_p or greater should also be included in the matching. However, this is in contradiction with the assumption that $(a_i, c_j) \notin Ch_p^H(X)$. The same argument applies for the Chilean policy. □

The preprocessing algorithm can use the above two lemmas iteratively. When some of the first applications are set to be rejected based on Lemma 4 then we can consider the second applications of these rejected students as their first applications and apply Lemma 3 again, and so forth.

For the 2008-Educatio instance by doing one round of the above preprocessing algorithm we could already fix 23,138 applications to be accepted (and thus the same number of students assigned), and 14,991 applications to be rejected. As a result the run time of the longest running Hungarian version improved from 158,949 seconds to 10,740. This improvement also enabled us to compute the results for the 2009-KRTK instances to be presented next.

6.6. Simulation results for the 2009-KRTK instances

As we described in the Introduction, we created two instances using different assumptions on the constraints, "2009-KRTK-previous" uses the quotas of the 2008-Educatio instance, and "2009-KRTK-admitted" uses the number of admitted students for the quotas. In the one round preprocessing phase we could fix 14,828 applications as accepted and 25,336 applications as rejected for the "2009-KRTK-previous" instance, whereas we could fix 18,902 applications as accepted and 40,404 applications as rejected for the "2009-KRTK-admitted" instance. The main statistics of the computed solutions can be seen in Tables 7 and 8, respectively.

7. Conclusion

Following up Ágoston et al. (2016), we developed new (mixed) IP formulations for the classical Gale-Shapley college admissions model, for the case of ties and for the case of common quotas. We also considered the most challenging case when both ties and common quotas are present. We found that the most efficient formulations with binary cutoff scores can terminate in a reasonable time for a real instance from 2008. Furthermore we also compared the three possible policies allowing ties, used in Hungary, Ireland and Chile.

One can consider to further improve our formulations and IP solution technique, perhaps also with more careful preprocessing tailored for the application. Another important line of research would be to provide more insight into the qualities of the models, what can be the reason of the different performances, whether some models are tighter than others, and can give better bounds in the branching.

For future work, one can also try to include lower quotas as the third special feature that is present in the application. This special feature makes the problem NP-hard, but it has turned out to be tractable in practice with careful preprocessing and IP techniques by Ágoston et al. (2016). Although it will be even more complicated to define the concepts of fairness, stability and non-wastefulness for this real scenario. The ultimate goal of our research project is going to be resolving this case as well designing an alternative (possibly better) solution technique to the currently used heuristic approach.

In addition, we also propose the usage of our formulations for other two-sided matching problems with distributional constraints. We believe that the flexibility of the robust IP technique can create a new perspective in solving complex matching problems under preferences with special objectives and constraints even for large markets, as demonstrated in this paper.

Acknowledgement

We thank Bence Kapás for linking the two sources of the 2008 Hungarian university admission data, that is our 2008 Educatio instance with the 2008 KRTK Databank instance, and also the 2008 and 2009 KRTK Databank instances under a BSc thesis project supervised by Péter Biró and Rita Fleiner. This work made it possible to do computational simulations on the 2009-KRTK instance as well.

References

Abdulkadiroğlu, A. (2005). College admissions with affirmative action. *International Journal of Game Theory, 33*(4), 535–549.
Abdulkadiroğlu, A., & Ehlers, L. (2007). Controlled school choice. Working paper.
Abdulkadiroğlu, A., Pathak, A., & Roth, A. E. (2005a). The New York City high school match. *American Economic Review, Papers and Proceedings, 95*(2), 364–367.
Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., & Sönmez, T. (2005b). The Boston public school match. *American Economic Review, Papers and Proceedings, 95*(2), 368–371.
Ágoston, K. C., & Biró, P. (2017). Modelling preference ties and equal treatment policy. In *Proceedings of ECMS 2017: 31st European conference on modelling and simulation* (pp. 516–522).
Ágoston, K. C., Biró, P., & McBride, I. (2016). Integer programming methods for special college admissions problems. *Journal of Combinatorial Optimization, 32*(4), 1371–1399.
Ágoston, K. C., Biró, P., & Szántó, R. (2018b). Stable project allocation under distributional constraints. *Operations Research Perspectives, 5*, 59–68.
Aygün, O., & Bo, I. (2013). College admission with multidimensional reserves: The Brazilian affirmative action case. Working paper.
Azevedo, E. M., & Leshno, J. D. (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy, 124*(5), 1235–1268.
Baïou, M., & Balinski, M. (2000). The stable admissions polytope. *Mathematical Programming, 87*(3), 427–439.
Baswana, S., Chakrabarti, P. P., Chandran, S., Kanoria, Y., & Patange, U. (2019). Centralized admissions for engineering colleges in India. In *Proceedings of EC-2019: the 2019 ACM conference on economics and computation* (pp. 323–324).
Biró, P. (2014a). Matching practices for secondary schools – Hungary. matching-in-practice.eu, accessed on 23 August.
Biró, P. (2014b). University admission practices - Hungary. matching-in-practice.eu, accessed on 23 August.
Biró, P. (2017). Applications of matching models under preferences. In U. Endriss (Ed.), *Trends in computational social choice* (pp. 345–373). chapter 18, AI Access
Biró, P., Fleiner, T., Irving, R. W., & Manlove, D. F. (2010). The college admissions problem with lower and common quotas. *Theoretical Computer Science, 411*, 3136–3153.
Biró, P., & Kiselgof, S. (2015). College admissions with stable score-limit. *Central European Journal of Operations Research, 23*(4), 727–741.
Biró, P., van de Klundert, J., Manlove, D., et al. (2019). Modelling and optimisation in European kidney exchange programmes. *European Journal of Operational Research, 291*(2), 447–456.
Biró, P., McBride, I., & Manlove, D. F. (2014). The hospitals/residents problem with couples: Complexity and integer programming models. In *Proceedings of SEA 2014: the 13th international symposium on experimental algorithms*. In *LNCS: vol. 8504* (pp. 10–21). Springer.

Bo, I. (2016). Fair implementation of diversity in school choice. *Games and Economic Behavior, 97*, 54–63.
Cao, N. V., Fragniére, E., Gautier, J.-A., Sapin, M., & Widmer, E. D. (2010). Optimizing the marriage market: An application of the linear assignment model. *European Journal of Operational Research, 202*(2), 547–553.
Chen, L. (2012). University admission practices Ireland. MiP Country Profile 8.
Delorme, M., García, S., Gondzio, J., Kalcsics, J., Manlove, D., & Pettersson, W. (2019). Mathematical models for stable matching problems with ties and incomplete lists. *European Journal of Operational Research, 277*(2), 426–441.
Diebold, F., & Bichler, M. (2017). Matching with indifferences: A comparison of algorithms in the context of course allocation. *European Journal of Operational Research, 260*(1), 268–282.
Echenique, F., & Yenmez, M. B. (2015). How to control controlled school choice. *American Economic Review, 105*(8), 2679–2694.
Ehlers, L., Hafalir, I. E., Yenmez, M. B., & Yildirim, M. A. (2014). School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory, 153*, 648–683.
Fleiner, T., & Jankó, Z. (2014). Choice function-based two-sided markets: Stability, lattice property, path independence and algorithms. *Algorithms, 7*(1), 32–59.
Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly, 69*(1), 9–15.
Garg, N., Kavitha, T., Kumar, A., Mehlhorn, K., & Mestre, J. (2010). Assigning papers to referees. *Algorithmica, 58*(1), 119–136.
Gonczarowski, Y. A., Kovalio, L., Nisan, N., & Romm, A. (2019). Matching for the Israeli "mechinot" gap-year programs: Handling rich diversity requirements. arXiv preprint arXiv:1905.00364.
Goto, M., Kojima, F., Kurata, R., Tamura, A., & Yokoo, M. (2017). Designing matching mechanisms under general distributional constraints. *American Economic Journal: Microeconomics, 9*(2), 226-262.
Irving, R. W., & Manlove, D. F. (2008). Approximation algorithms for hard variants of the stable marriage and hospitals/residents problems. *Journal of Combinatorial Optimization, 16*(3), 279–292.
Kamada, Y., & Kojima, F. (2014). Efficient matching under distributional constraints: Theory and applications. *American Economic Review, 105*(1), 67–99.
Kamada, Y., & Kojima, F. (2017a). Recent developments in matching with constraints. *American Economic Review, 107*(5), 200–204.
Kamada, Y., & Kojima, F. (2017b). Stability concepts in matching under distributional constraints. *Journal of Economic Theory, 168*, 107–142.
Kwanashie, A., & Manlove, D. F. (2014). An integer programming approach to the hospitals/residents problem with ties. In *Proceedings of OR 2013: the international conference on operations research* (pp. 263–269). Springer.
Manlove, D. F. (2013). *Algorithms of matching under preferences*. World Scientific Publishing.
Pentico, D. V. (2007). Assignment problems: A golden anniversary survey. *European Journal of Operational Research, 176*, 774–793.
Rios, I., Larroucau, T., Parra, G., & Cominetti, R. (2014). College admissions problem with ties and flexible quotas. Working paper.
Roth, A. E. (1984). 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.
Roth, A. E., & Peranson, E. (1999). The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review, 89*, 748–780.
Roth, A. E., & Sotomayor, M. A. O. (1990). Two-sided matching: a study in game-theoretic modeling and analysis. Cambridge: Econometric Society monographs.
Sönmez, T., & Yenmez, B. (2019a). Affirmative action in india via vertical and horizontal reservations. Working paper.
Sönmez, T., & Yenmez, B. (2019b). Constitutional implementation of vertical and horizontal reservations in India: A unified mechanism for civil service allocation and college admissions. Working paper.
Wu, Q., & Roth, A. E. (2018). The lattice of envy-free matchings. *Games and Economic Behavior, 109*, 201–211.
Yokoi, Y. (2020). Envy-free matchings with lower quotas. *Algorithmica, 82*, 188–211.