

NBER WORKING PAPER SERIES

SCHOOL ASSIGNMENT BY MATCH QUALITY

Atila Abdulkadiroglu  
Umut M. Dur  
Aram Grigoryan

Working Paper 28512  
<http://www.nber.org/papers/w28512>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
February 2021

Abdulkadiroglu: Department of Economics, Duke University and National Bureau of Economic Research (NBER), [aa88@duke.edu](mailto:aa88@duke.edu); Dur: Department of Economics, North Carolina State University, [udur@ncsu.edu](mailto:udur@ncsu.edu), Grigoryan: Department of Economics, Duke University, [ag404@duke.edu](mailto:ag404@duke.edu). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Atila Abdulkadiroglu, Umut M. Dur, and Aram Grigoryan. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

School Assignment by Match Quality  
Atila Abdulkadiroglu, Umut M. Dur, and Aram Grigoryan  
NBER Working Paper No. 28512  
February 2021  
JEL No. D47,I20

**ABSTRACT**

Proponents of school choice argue that it improves educational outcomes by allowing parents to self-select into schools that are most effective for their children. Contrary to these arguments, empirical evidence suggests that parents may not incorporate school effectiveness or match quality when choosing schools. The findings potentially impugn proponents' effectiveness arguments of choice-based assignment. We develop novel solutions that restore effectiveness by maximizing match quality subject to stability constraints. Maximization algorithms are provided for both small and large school districts. Simulations reveal substantial match quality gains from our solutions compared to the celebrated Deferred Acceptance mechanism with a random tie-breaker. Our methodology can be used to optimize for other policy objectives in school choice or other priority-based matching problems.

Atila Abdulkadiroglu  
Department of Economics  
Duke University  
213 Social Sciences Building  
Durham, NC 27708  
and NBER  
atila.abdulkadiroglu@duke.edu

Aram Grigoryan  
Duke University  
aram.grigoryan@duke.edu

Umut M. Dur  
North Carolina State University  
College of Management  
umutdur@gmail.com

# School Assignment by Match Quality

Atila Abdulkadiroglu, Umut Dur and Aram Grigoryan\*

February 24, 2021

## Abstract

Proponents of school choice argue that it improves educational outcomes by allowing parents to self-select into schools that are most effective for their children. Contrary to these arguments, empirical evidence suggests that parents may not incorporate school effectiveness or match quality when choosing schools. The findings potentially impugn proponents' effectiveness arguments of choice-based assignment. We develop novel solutions that restore effectiveness by maximizing match quality subject to stability constraints. Maximization algorithms are provided for both small and large school districts. Simulations reveal substantial match quality gains from our solutions compared to the celebrated Deferred Acceptance mechanism with a random tie-breaker. Our methodology can be used to optimize for other policy objectives in school choice or other priority-based matching problems.

## 1 Introduction

Parental choice over public schools has become an integral education reform tool around the world. Market Design for school choice (Abdulkadiroğlu and Sönmez, 2003)<sup>1</sup> has led to creation and implementation of efficient and transparent admissions processes in school choice programs. It is argued that parental choice may boost educational outcomes for students by creating competitive pressure on schools (Friedman, 1962; Tweedie et al., 1990; Hoxby, 2003), by allowing students to sort into schools with better match quality (Hoxby, 2000), and by allowing parents to act on local information, in turn providing better incentives for schools to invest in educational effectiveness than a centralized accountability system would do (Peterson and Campbell, 2001). These mechanisms would be at work only if parents choose schools partly based on school effectiveness and match quality. Contrary to these claims, parents do not seem to incorporate school effectiveness when forming preferences. For example, Abdulkadiroğlu et al. (2020) find that parents do not prefer schools that are especially effective for their own children and do not enroll their children in schools that are a better-than-average match for them. These findings raise a new challenge for the Market

---

\*Abdulkadiroğlu: Department of Economics, Duke University and National Bureau of Economic Research (NBER), aa88@duke.edu; Dur: Department of Economics, North Carolina State University, udur@ncsu.edu, Grigoryan: Department of Economics, Duke University, ag404@duke.edu.

<sup>1</sup>School choice problem is closely related to college admission problem (Gale and Shapley, 1962; Roth, 1985) and student placement problem (Balinski and Sönmez, 1999). It differs from college admission problem as schools are considered as objects, and it differs from student placement problem as respecting school priorities is not forced by the central authority.

Design agenda: Can school admissions process be designed to attain higher educational outcomes even in the absence of parental preferences for effectiveness and match quality?

Given some measure of match quality of students at schools, finding an assignment of students to schools that maximizes sum of match qualities (hereafter, aggregate match quality) is a trivial linear programming problem. However, admissions to schools are constrained by parental preferences and by admissions priorities granted to students at schools. Preferences and priorities introduce to this otherwise standard linear programming problem the so-called *stability constraint*. Namely, an applicant can only be assigned to a school listed in her application form, and if an applicant prefers a school to her assignment, the school must be fully assigned to applicants with better or equal admissions priorities. Stability, also more aptly called justified-envy-freeness in the context of school assignment, has become critical in the design of school admissions processes (Abdulkadiroğlu et al., 2009). Finding a stable assignment that maximizes aggregate match quality, which we will refer to as match quality optimal assignment, is an NP-hard problem when priorities are weak. Hence, in general, this problem has no computationally tractable solution. We introduce an algorithm that is polynomial time in the number of students, but potentially exponential in the number of schools. Since the number of schools in a school districts is typically small, the algorithm is applicable for the majority of US school districts. Building on that solution, we also introduce a random search algorithm for large school districts. Our simulations show that even a single search produces an aggregate match quality that is close to that of the match quality optimal assignment.

When preferences and priorities are strict, a match quality optimal assignment can be computed in polynomial time by formulating it as a linear programming problem (Roth et al., 1993). However, real life choice programs typically feature weak priorities. In most applications, schools sort students into thick priority classes based on residential address, status of sibling enrollment etc. Our strategy builds on bounding the set of stable assignments and then solving the optimization problem within subsets of stable assignments. Admissions cutoff, or simply cutoff, at a school is the worst priority admitted by the school.<sup>2</sup> Our first building blocks are two new algorithms to compute lower and upper bounds for school cutoffs at stable assignments when priorities are weak.

The first algorithm, which we refer to as Deferred Rejection (DR), works like the student proposing deferred acceptance algorithm (DA) (Gale and Shapley, 1962) with the caveat that, when there are more students applying to a school than available seats, the school finds the worst priority that would fill up its seats, rejects all students with strictly worse priorities and provisionally holds the remaining students. The algorithm then identifies students that would be rejected by schools for any possible tie breaking among equal priority students. Unlike the student proposing DA, under DR a school may hold more students than available seats. Therefore, DR produces bounds for cutoffs: schools attain weakly better priority cutoffs in every stable assignment.

The second algorithm, which we refer to as Deferred Proposal (DP), works like the school proposing DA. Each school finds the worst priority that would fill up its seats, proposes to students with strictly better priority. Then, the algorithm identifies students that would reject schools for any possible tie breaking. Under DP schools do not necessarily fill up all of their seats. Therefore, DP too produces bounds for cutoffs: schools attain weakly worse priority cutoffs in every stable assignment.

---

<sup>2</sup>For example, consider a school that grants priorities A, B and C such that A is better than B and B is better than C. Suppose that the worst priority students assigned to the school have priority B. Then the cutoff at the school is B.

Finally, using the bounds identified by DR and DP, we find a match quality optimal assignment using our Match Quality Optimal Algorithm (MQO). For a fixed vector of admission cutoffs, or a cutoff profile, MQO finds a stable assignment that has weakly higher match quality than any stable assignment with the given cutoff profile. This step involves a novel formulation of the problem as a minimum-cost flow problem. Then, by exhaustively searching within the set of cutoff profiles identified by DR and DP, MQO yields a match quality optimal assignment. Such exhaustive search is infeasible without DR and DP procedures for many realistic instances.

MQO becomes computationally intractable as the number of schools increases. Therefore, we develop a random search algorithm for large districts, which we call Locally Match Quality Optimal Algorithm (L-MQO). Instead of searching within bounds, L-MQO first runs the student proposing DA with some random tie-breakers to identify cutoffs of some stable assignment. Then, using the minimum-cost flow formulation, it finds a stable assignment with a weakly higher match quality than any stable assignment with the given cutoffs. L-MQO finds a globally match quality optimal assignment if it uses the cutoffs of a globally match quality optimal assignment. Therefore, repeating L-MQO with distinct random tie-breakers yields a random search algorithm. Remarkably, even a single run of L-MQO yields locally optimal solutions that are very close to the global optimum.

We run simulations with 1,000 students and 20 schools, with 50 seats each for 75 environments with different parameter values. The match quality gains are substantial. MQO on average results in around 40% higher aggregate match quality compared to the student proposing DA with random tie-breaker. Match quality gains under L-MQO is only slightly smaller than that of the MQO, i.e., less than 1%. Additionally, we consider a version of student proposing DA, where schools break ties among equal priority students based on the match quality, instead of applying a random tie-breaker. Such quality-based tie-breaking in DA significantly improves on random tie-breaking, however the match quality gains are less than two thirds of those under MQO or L-MQO.

Data-driven effectiveness or match quality-based assignment is widely applied for many market design problems. Refugee resettlement is a major example. Hebrew Immigrant Aid Society (HIAS) uses a machine-learning based algorithm called *Annie MOORE* (Matching and Outcome Optimization for Refugee Empowerment), named after Annie Moore, the first immigrant on record at Ellis Island, New York in 1892, which, according to developers Trapp et al. (2018), is “... *the first software designed for resettlement agency pre-arrival staff to recommend data-driven, optimized matches between refugees and local affiliates while respecting refugee needs and affiliate capacities.*” Organ transplantation is another example of a match quality-based assignment. The algorithm of United Network for Organ Sharing (UNOS) prioritizes those patients who are in most urgent need of the transplant, and/or who are “... *most likely to have the best chance of survival if transplanted*”.<sup>3</sup> For school choice, on the other hand, assignment is typically determined only based on students preferences and priorities: match quality and effectiveness are completely ignored. To the best of our knowledge, ours is the first work to introduce a match quality-based algorithm for school choice. The practical value of such an algorithm is particularly enhanced by the recent advances in econometric techniques for assessing match quality or effectiveness in the school choice setting (Abdulkadiroğlu et al., 2017, 2020).

Although our emphasis in this work is on finding match quality optimal assignment, our optimization solution under stability constraints can be used to address other important policy objectives such as improving diversity (Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu, 2005; Hafalir et al.,

---

<sup>3</sup><https://unos.org/about/national-organ-transplant-system/>

2013; Dur et al., 2018) or maximizing student welfare (Erdil and Ergin, 2008, 2017; Abdulkadiroğlu et al., 2009). These extensions are briefly discussed in the Discussion section.

The remainder of this work is organized as follows. Section 2 reviews the related literature. Section 3 describes the school choice problem and discusses the NP-hardness of finding match quality optimal assignment. Section 4 introduces a solution based on a minimum-cost flow formulation. Section 5 develops DA-based algorithms that bound the set of stable assignments, which is an important prerequisite for the tractability of the solution in Section 4. Section 6 introduces a partial solution for large school districts. Section 7 reviews simulations results and Section 8 concludes. Proofs are in Appendix A.

## 2 Related Literature

As mentioned above, our work contributes to the literature that advocates match quality- or effectiveness-based assignment. Slauch et al. (2016) develop a data-driven algorithm to match children to families for adoption. The algorithm evaluates the probability of a successful adoption based on child’s traits and family’s preferences and incorporates this information to match them more effectively. Bansak et al. (2018) and Trapp et al. (2018) introduce similar data-driven solutions for refugee resettlement to maximize employment outcomes. Grigoryan (2020) provides match quality maximization algorithm for pandemic rationing problems. Our environment is different from all the works above, and therefore, match quality maximization in our problem requires an original solution. We provide new algorithms for bounding cutoffs and a minimum-cost flow formulation for maximization under stability constraints.

Our work is related to the widely studied problem of finding a stable assignment of largest size, i.e., one that assigns the most number of students. This is a special case of our match quality maximization problem where match quality is binary and equals zero if and only if a student is unassigned. In general, this problem is NP-hard (Manlove et al., 2002). Several heuristic and approximation algorithms have been discussed in the computer science literature. For example, Kwanashie and Manlove (2014) show that in practice the problem may be solved through integer programming. However, there are no theoretical guarantees that this approach is successful. Other papers provide  $\alpha$ -approximation algorithms (McDermid, 2009; Király, 2011; Iwama et al., 2014), which guarantee that the value of the optimal solution is no more than  $\alpha$  times larger than that of the approximating algorithm. These approximation bounds are typically not tight and the solutions may be far from the global optimum.

Another strand of papers provide approximate solutions to NP-hard optimization and matching problems by utilizing the ‘sparsity’ of the relevant constraint matrix. For example, Nguyen and Vohra (2013) study a multi-unit allocation problem and maximize social-welfare under fairness and incentive constraints. They assume that the number of objects each agent can receive is small relative to the market size. Therefore, the number of fractional solutions in the linear programming relaxation is guaranteed to be small. After applying a rounding algorithm, the resultant solution is ‘approximately feasible’, in a sense that constraints are violated only ‘slightly’ violated. Similar approximately feasible solutions are obtain stable allocations in environments with complementarities (Nguyen and Vohra, 2018) and complex distributional constraints (Nguyen et al., 2020; Nguyen and Vohra, 2019). These last papers use versions of the Scarf’s lemma and rounding algorithms to obtain approximate solutions. Unlike the works above, our main solution offers exact optimization.

Methodologically, our work is related to matching theory papers that use network flows, such Katta and Sethuraman (2006); Kim and Mierendorff (2013); Chandramouli and Sethuraman (2020) and Grigoryan (2020). The network flow approach allows to solve various optimization problem with integer solutions. This makes the method highly useful for matching problems with indivisible goods. Our setting, and therefore, the corresponding network flow formulation is different from all the works above.

Our work is potentially closest to papers that solve optimization problems in the context of school choice. Despite the NP-hardness result, Bodoh-Creed (2020) (hereafter, BC) provides a computationally feasible optimization algorithm under stability constraints using the fact that the number of schools and priority classes are typically small. Our solution is different from that of BC in several important ways. First, BC studies a continuum model and his method finds fractional solutions. In finite markets his fractional solutions may be interpreted as ex-ante stable stochastic assignments (Kesten and Ünver, 2015). We, on the other hand, study a finite market setting and use the standard definition of stability. Thus, our optimization problem requires an integral solution. We use a novel method based on minimum-cost flow formulation to solve the problem with integrality constraints. Second, BC’s algorithm relies on the assumption that if a school has empty seats under student proposing DA with some tie-breaker, it also has empty seats in all stable assignments. Our solution does not require such an assumption. Instead, we develop procedures, namely DR and DP, that reduce the computational burden of our algorithm so that it can handle realistic problems without additional assumptions. Moreover, the reduction procedures allow us to work with larger problems, with potentially more priority classes at each school. Finally, we also introduce a partial solution for large school districts, which makes our contribution applicable to any school choice problem. Two other papers that study optimization problems in the school choice setting are Ashlagi and Shi (2016) and Shi (2019). These papers however ignore exogenously given priority rankings, which allows them to relax stability constraints.

### 3 The Problem

A typical school choice problem with weak priorities consists of a nonempty set of students  $S$ , a nonempty set of schools  $C$ , a profile of strict preferences of students  $P = (P_s)_{s \in S}$ , a vector of school capacities  $\kappa = (\kappa_c)_{c \in C}$  and a profile of priorities granted to students at schools  $\rho = (\rho_{sc})_{s \in S, c \in C}$ . Let  $c P_s c'$  denote that  $s$  prefers  $c$  to  $c'$ , and let  $R_s$  denote the weak preference relation, i.e.,  $c R_s c'$  if and only if  $c P_s c'$  or  $c = c'$ . Capacity  $\kappa_c \in \mathbb{N}$  denotes the maximum number of students that can be assigned to  $c$ . We assume that  $\sum_{c \in C} \kappa_c \geq |S|$ , i.e., the total number of school seats exceed the total number of students.<sup>4</sup> We model priorities via integer numbers:  $\rho_{sc} \in \{1, 2, \dots, K\}$  denotes student  $s$ ’s priority at school  $c$ . Without loss of generality, we assume that a smaller number indicates better priority. That is, if  $\rho_{sc} < \rho_{s'c}$ , then  $s$  has better priority than  $s'$  at  $c$ , and if  $\rho_{sc} = \rho_{s'c}$ , then  $s$  and  $s'$  have equal priority at  $c$ . Note that we allow for weak priorities by allowing students to have the same priority number at schools. For any  $k \in \mathbb{N}$ , we say student  $s$  has the  $k$ -th best priority at school  $c$  among students in  $\tilde{S} \subseteq S$ , if the number of students in  $\tilde{S}$  with strictly better priorities at  $c$  is strictly  $k - 1$ , i.e.,  $|\{s' \in \tilde{S} : \rho_{s'c} < \rho_{sc}\}| = k - 1$ . We add **match quality** to this standard model:  $q(s, c) \in \mathbb{R}$  denotes the quality of match between  $s$  and  $c$ . A higher value of

---

<sup>4</sup>Assuming that students rank all school as acceptable, and that there are enough seats for all students is without loss of generality as we can add a school that represents being ‘unassigned’ and that has enough capacity for every student.

$q(s, c)$  indicates a better match quality. Let  $q = (q(s, c))_{s \in S, c \in C}$  be the match quality profile. We represent a school choice problem, or simply a problem, with tuple  $(S, C, \kappa, P, \rho, q)$ . We will fix the problem and omit references to it in the rest of the text.

An assignment is a mapping  $\mu : S \cup C \rightarrow C \cup 2^S$ , such that for all  $s \in S$  and  $c \in C$ ,

- $\mu(s) \in C$ ,
- $\mu(c) \subseteq 2^S, |\mu(c)| \leq \kappa_c$ ,
- $c = \mu(s) \Leftrightarrow s \in \mu(c)$ .

An assignment  $\mu$  is **stable** if there is no **blocking pair**  $(s, c) \in S \times C$  such that  $c P_s \mu(s)$  and either  $|\mu(c)| < \kappa_c$  or  $\rho_{sc} < \rho_{s'c}$  for some  $s' \in \mu(c)$ . Let  $\mathcal{A}$  denote the set of all stable assignments. It is well-known that  $\mathcal{A}$  is not empty for any problem (Gale and Shapley (1962), Irving (1994)).

Our objective is to find a stable assignment that maximizes the aggregate match quality among all stable assignments. Namely, a **match quality optimal** assignment  $\mu^*$  is a stable assignment such that

$$\sum_{s \in S} q(s, \mu^*(s)) \geq \sum_{s \in S} q(s, \mu(s)),$$

for any stable  $\mu \in \mathcal{A}$ .

Since the set of stable assignments is non-empty and finite, there always exists at least one match quality optimal assignment. When preferences and priorities are strict, which is a special case of our setting, a match quality optimal assignment can be found in polynomial time by formulating the set of stable assignments as linear programming constraints via Roth et al. (1993). In general, finding a match quality optimal assignment is an NP-hard problem. Therefore, it is unlikely to be polynomial time solvable.

**Proposition 1.** *When priorities are weak, finding a match quality optimal assignment is an NP-hard problem.*

Proposition 1 has been proved by Manlove et al. (2002) for the special case of our problem of finding the stable matching of maximum size.

## 4 Our Solution

The NP-hardness of finding match quality optimal assignment prohibits a general computationally tractable solution for all problems. In this section, we develop a solution that finds a match quality optimal assignment for problems with 30 and less schools in reasonable time. Thus, the solution is applicable for most school districts in the US (National Center for Education Statistics 2002).<sup>5</sup> We discuss our solution for large districts in Section 6.

<sup>5</sup>Only 323 of around 13,000 school districts in the US had more than 30 schools in the school year 2000-2001 (National Center for Education Statistics 2002). Moreover, even if the school district is large, the number of schools serving the same grade may be below 30, in which case the district can be handled by our algorithm. For example, Boston Public Schools has around 130 schools, but only around 30 of those are high schools. So our solution will be tractable for assigning high school students in Boston. Moreover, not all schools in a school district admits students



In a nutshell, stable assignments can be characterized by admissions cutoffs at schools (Abdulkađirođlu et al., 2015; Azevedo and Leshno, 2016). Given a cutoff profile, we provide a linear programming formulation that finds a stable assignment with an aggregate match quality that is weakly higher than that at every stable assignment with the given cutoff profile, whenever one exists. This allows us to find a locally optimal assignment in polynomial time. Equipped with this solution, we search for a global optimum by tracing all cutoff profiles. However, this strategy may not generally be computationally tractable. We introduce two new algorithms to bound the set of cutoff profiles supporting a stable assignment (see Section 5). This allows us to restrict the search of a match quality optimal assignment only within these bounds. Simulations in Section 7 show that the global optimum, i.e., match quality optimal stable assignment, can be computed in reasonable time for realistic parameter values for problems with 30 schools and smaller.

#### 4.1 Locally Match Quality Optimal Assignment

We first define (admissions) cutoffs at schools. Given a stable assignment  $\mu \in \mathcal{A}$ , let cutoff at school  $c$ ,  $\rho_c(\mu) \in \{1, 2, \dots, K + 1\}$ , be

$$\rho_c(\mu) := \begin{cases} \max_{s \in \mu(c)} \rho_{sc} & \text{if } |\mu(c)| = \kappa_c, \\ K + 1 & \text{otherwise.} \end{cases}$$

In words, under assignment  $\mu$ , if school  $c$  fills its capacity, then its cutoff is equal to the priority of the assignee with the worst priority at the school, otherwise the cutoff is set to  $K + 1$ . Let  $\rho(\mu) = (\rho_c(\mu))_{c \in C}$  denote the cutoff profile. Then, the following result is immediate.

**Proposition 2.** *An assignment  $\mu$  is stable if and only if there is no pair  $(s, c) \in S \times C$  such that  $c P_s \mu(s)$  and  $\rho_{sc} < \rho(\mu)$ .*

For a vector  $r \in \{1, 2, \dots, K + 1\}^{|C|}$ , let  $\mathcal{A}_r \subseteq \mathcal{A}$  be the set of stable assignments with cutoff profile equal to  $r$ , i.e.,  $\mu \in \mathcal{A}_r$  if and only if  $\mu$  is stable and  $\rho(\mu) = r$ . We say that  $r \in \{1, 2, \dots, K + 1\}^{|C|}$  supports a stable assignment if  $\mathcal{A}_r \neq \emptyset$ .

For a given  $r \in \{1, 2, \dots, K + 1\}^{|C|}$ , we construct a set of stable assignments that is a superset of  $\mathcal{A}_r$  as follows. First, we define two disjoint subsets of schools:  $C^+(r) := \{c \in C : r_c < K + 1\}$  and  $C^-(r) := C \setminus C^+(r)$ . Next, for every student  $s \in S$ , define  $C_s(r) \subseteq C$  as

$$C_s(r) = \{c \in C : \rho_{sc} \leq r_c \text{ and } \rho_{sc'} \geq r_{c'} \text{ for all } c' \in C \text{ such that } c' P_s c\}.$$

Here,  $C_s(r)$  can be interpreted as the set of schools that  $s$  can potentially be assigned given cutoffs  $r$ . Let  $\tilde{\mathcal{A}}_r$  denote the the set of assignments satisfying, for all  $\mu \in \tilde{\mathcal{A}}_r$ ,

1.  $\mu(s) \in C_s(r)$  for all  $s \in S$ ,
2.  $|\mu(c)| = \kappa_c$  for all  $c \in C^+(r)$ .

---

via school choice. For instance, in Wake County public school system, which is the 15th largest school system in US, there are more than 170 schools (only magnet schools are admitting students via choice-based assignment). There are there 36 magnet schools, 23 elementary schools, 9 middle schools and 4 high schools.

That is, under any assignment in  $\bar{\mathcal{A}}_r$  each student  $s$  is assigned a school in  $C_s(r)$  and each school  $c \in C^+(r)$  fills its capacity. It is immediate from the definition of  $\bar{\mathcal{A}}_r$  that  $\rho(\mu) \leq r$  for all  $r \in \{1, 2, \dots, K + 1\}^{|C|}$ . The following lemma justifies our search for match quality maximizing assignments within  $\bar{\mathcal{A}}_r$ .

**Lemma 1.** *Each assignment in  $\mathcal{A}_r$  is a member of  $\bar{\mathcal{A}}_r$  and each assignment in  $\bar{\mathcal{A}}_r$  is stable. That is,*

$$\mathcal{A}_r \subseteq \bar{\mathcal{A}}_r \subseteq \mathcal{A}.$$

Thus, maximizing match quality in  $\bar{\mathcal{A}}_r$  yields a stable assignment that has a weakly higher match quality than any assignment in  $\mathcal{A}_r$ . We refer to this solution as a **locally match quality optimal assignment** for the vector  $r$ .

Our final step formulates this optimization problem as a **minimum-cost flow** problem, which finds a locally match quality optimal assignment, whenever one exists, i.e., whenever  $\bar{\mathcal{A}}_r \neq \emptyset$ ,<sup>6</sup> in polynomial time. The minimum-cost flow problem has numerous practical applications, including the design of optimal electrical network, transportation system or house allocation (Ahuja et al., 1993). To the best of our knowledge, ours is the first application of minimum-cost flow method for optimization under stability constraints.

To formulate the minimum-cost flow problem, consider the following components:

- a set of vertices  $V = S \cup C \cup \{t\}$ , for some  $t \notin S \cup C$ ,
- a set of edges

$$E = \{(c, t) : c \in C^-(r)\} \cup \{(s, c) : c \in C_s(r)\} \subseteq V \times V,$$

- a capacity  $u(e)$  for each edge  $e \in E$ , given by

$$u(e) = \begin{cases} 1 & \text{if } e \in S \times C \\ \kappa_c & \text{if } e = (c, t) \in C^-(r) \times \{t\} \end{cases},$$

- a cost  $l(e)$  for each edge  $e \in E$ , given by

$$l(e) = \begin{cases} -q(e) & \text{if } e \in S \times C \\ 0 & \text{if } e \in C^-(r) \times \{t\} \end{cases},$$

- a value  $b(v)$  for each vertex  $v \in V$ , given by

$$b(v) = \begin{cases} 1 & \text{if } v \in S \\ 0 & \text{if } v \in C^-(r) \\ -\kappa_v & \text{if } v \in C^+(r) \\ -|S| + \sum_{c \in C^+(r)} \kappa_c & \text{if } v = t \end{cases}.$$

Figure 1 illustrates the constructed minimum-cost flow graph using for an example with three students  $S = \{s_1, s_2, s_3\}$  and two schools  $C^+ = \{c_1\}$  and  $C^- = \{c_2\}$ .

<sup>6</sup>Otherwise, our minimum-cost flow algorithm/solver outputs that  $\bar{\mathcal{A}}_r = \emptyset$ .

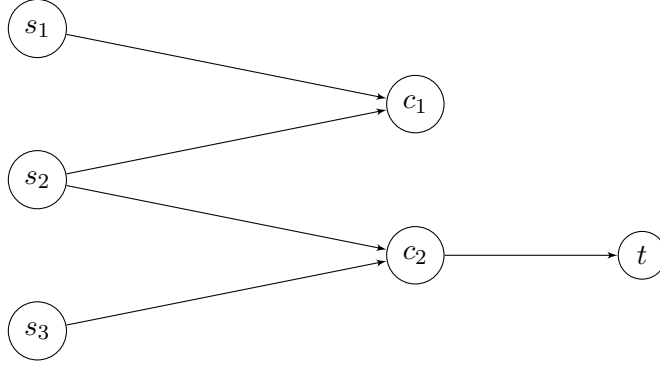


Figure 1: Minimum-cost flow graph

A positive value of  $b(v)$  indicates that  $v$  is a supply vertex and a negative value of  $b(v)$  indicates that  $v$  is a demand vertex. A vertex with  $b(v) = 0$  is a transshipment vertex. We represent a flow function with a mapping  $f : E \rightarrow \mathbb{R}$ . Our goal is to find a least costly way of transferring values from supply vertices to demand vertices without exceeding the capacity of edges. Formally, we solve the following linear program:

$$\min_{f: E \rightarrow \mathbb{R}} \sum_{e \in E} l(e)f(e)$$

subject to

$$\sum_{v' \in V: (v, v') \in E} f(v, v') - \sum_{v' \in V: (v', v) \in E} f(v', v) = b(v), \forall v \in V, \quad (1)$$

$$0 \leq f(e) \leq u(e), \forall e \in E. \quad (2)$$

A flow function  $f$  is feasible if it satisfies constraints 1 and 2. A flow function is integral if  $f(e) \in \mathbb{Z}$  for all  $e \in E$ . By condition 2, for any feasible and integral flow function  $f$ ,  $f(e)$  takes values 0 or 1 for each edge  $e \in S \times C$ . We establish a connection between the set of feasible and integral solutions of the minimum-cost flow problem and that set of locally match quality optimal assignments for vector  $r$ . Formally, with each feasible and integral flow function  $f$  we associate an assignment  $\mu_f$  such that for each edge  $(s, c) \in S \cup C$ ,

$$\mu_f(s) = c \text{ if and only if } f(s, c) = 1.^7$$

Conversely, with each assignment  $\mu$  we associate a flow function  $f_\mu$ , where

$$f_\mu(e) = \begin{cases} 1 & \text{if } e = (s, c) \in S \times C, \mu(s) = c \\ 0 & \text{if } e = (s, c) \in S \times C, \mu(s) \neq c \\ |\mu(c)| & \text{if } e = (c, t) \in C^-(r) \times \{t\} \end{cases}$$

Then,

**Lemma 2.** *If  $f$  is feasible and integral, then  $\mu_f \in \bar{\mathcal{A}}_r$ . Conversely, if  $\mu \in \bar{\mathcal{A}}_r$ , then  $f_\mu$  is feasible and integral.*

<sup>7</sup>Note that for  $\mu_f$  to be an assignment we need that for each  $s \in S$ ,  $f(s, c) = 1$  for exactly one  $c \in C$ . As we verify in Lemma 2, this is indeed the case.

We briefly discuss the intuition behind Lemma 2. A non-zero flow from a student to a school is interpreted as the student being assigned to the school. Constraint 1 requires that each supply vertex has an outgoing flow that equals its value, and each demand vertex has an incoming flow that equals its value. Each student  $s \in S$  has a value  $b(s) = 1$ , therefore, she is assigned to exactly one school in  $C_r(s)$  in any integral solution of the minimum-cost flow problem. By definition, this is required for every assignment in  $\bar{\mathcal{A}}_r$ . Each school  $c \in C^+(r)$  has a value  $b(c) = -\kappa_c$ , meaning that  $\kappa_c$  students are assigned to  $c$  in any integral solution of the minimum-cost flow problem. Again, by definition, this is required for every assignment in  $\bar{\mathcal{A}}_r$ . There are no such requirements for schools in  $C^-(r)$ . However, these schools are connected to vertex  $t$ , and  $b(t) = -|S| + \sum_{c \in C^+(r)} \kappa_c$ . This means that schools in  $C^-(r)$  cumulatively accommodate all students who are not assigned to a school in  $C^+(r)$  in any integral solution of the minimum-cost flow problem. Finally, constraint 2 guarantees that no school  $c \in C^-(r)$  is assigned more than  $\kappa_c$  students.

Given the equivalence established in Lemma 2, we conclude that there is a feasible and integral flow function if and only if  $\bar{\mathcal{A}}_r \neq \emptyset$ . Moreover, if  $f^*$  is an optimal integral solution to the minimum-cost flow problem, then  $\mu_{f^*} \in \bar{\mathcal{A}}_r$  maximizes match quality in  $\bar{\mathcal{A}}_r$ . There are known polynomial time algorithms that find an optimal integral solution to the minimum-cost flow problem, whenever one exists. For example, the problem can be solved by a cycle-canceling algorithm (Ahuja et al., 1993; Sokkalingam et al., 2000). Hence, the following result is an immediate consequence of Lemma 2 and polynomial time solvability of the minimum-cost flow problem.

**Proposition 3.** *There are polynomial time algorithms that find a locally match quality optimal assignment for a vector  $r$ , whenever one exists.*

In Example 1 in Appendix B we illustrate how a locally match quality optimal assignment is found using the minimum-cost flow formulation.

## 4.2 Match Quality Optimal Assignment

Equipped with these results, we are ready to introduce an algorithm for match quality optimal assignment.

### *Match Quality Optimal Algorithm (MQO)*

**Step 0:** Set  $Q^* = -\infty$ .

**Step  $t \geq 1$ :** Select an  $r \in \{1, 2, \dots, K + 1\}^{|C|}$  which has not been selected in previous steps. If such an  $r$  does not exist, the algorithm terminates. Otherwise, using the minimum-cost flow formulation described in Section 4.1 we find a locally match quality optimal assignment  $\mu_r$  for vector  $r$ , if one exists. If  $\sum_{s \in S} q(s, \mu_r(s)) > Q^*$ , we set  $Q^* = \sum_{s \in S} q(s, \mu_r(s))$  and  $\mu^* = \mu_r$ . We continue with Step  $t + 1$ .

The outcome of MQO algorithm is assignment  $\mu^*$ . In other words, the above algorithm find a locally match quality optimal assignment for every  $r \in \{1, 2, \dots, K + 1\}^{|C|}$ , whenever they exist, and picks the one that gives the largest aggregate match quality. The fact that MQO algorithm considers all possible vectors that can support a stable assignment implies the following result.

**Proposition 4.** *MQO algorithm gives a match quality optimal assignment.*

The time complexity of MQO algorithm is

$$\mathcal{P}(|S| + |C|) \times (K + 1)^{|C|},$$

where  $\mathcal{P}(|S| + |C|)$  is some polynomial function of  $|S| + |C|$ . Without further reductions in computation, the algorithm becomes intractable for moderately large problems, as its time complexity grows exponentially in the number of schools. Therefore, in Section 5 we provide algorithms that reduce the computational burden of MQO, by eliminating many cutoff profiles that do not support a stable assignment. With these reductions, the solution becomes applicable for most of the real-life student assignment systems.

## 5 Upper and Lower Bounds for Stable Assignment Cutoffs

The method developed in the previous section is not computationally tractable for problems with moderately large number of schools and priority classes. To lower the computational burden, in this section we introduce two algorithms, Deferred Rejection and Deferred Proposal, that identify upper and lower bounds for cutoff profiles that support a stable assignment. This helps us eliminate a significant number of cutoff profiles that do not support a stable assignment. Then, match quality optimal assignment is found by computing the minimum-cost flow solution only for cutoff profiles within the identified bounds.

The new algorithms are based on student proposing and school proposing versions of DA. Applying DA with an arbitrary tie-breaker results in a cutoff profile that supports a stable assignment. Identifying all cutoff profiles that support a stable assignment by applying different tie-breakers is computationally intractable. In contrast, our algorithms utilize no tie-breaker and operate in polynomial time. Although they do not eliminate all cutoff profiles not supporting a stable assignment, simulations in Section 7 show that the amount of eliminations is substantial.

We first describe the conventional DA algorithms. At every step of the student proposing DA, each student applies to her most preferred school that has not rejected her yet. Each school  $c$  provisionally accepts from all of its applicants up to  $\kappa_c$  in the order of priorities and tie-breakers, and rejects the rest. The algorithm terminates when there is no rejection, the provisional acceptances at that point are finalized. At every step of the school proposing DA, each school proposes to up to  $\kappa_c$  students who have not rejected the school in the order of priorities and tie-breakers. Students provisionally accept the proposal of their most preferred school and reject the rest. The algorithm terminates when there is no rejection, the provisional acceptances at that point are finalized.

Unlike the student proposing DA, our Deferred Rejection algorithm only rejects students who cannot be assigned to the school at any stable assignment. Consequently, the number of students tentatively held by each school may exceed the school’s capacity. The resulting assignment provides upper bounds for cutoff profiles that support a stable assignment.

DR runs through multiple rounds until it cannot reject any more students. In each round, it goes through two stages. The first stage identifies “immediate unstable demand” at every school by students who cannot be assigned to the school at any stable assignment.<sup>8</sup> The second stage

---

<sup>8</sup>Kwanashie and Manlove (2014) introduce algorithms analogous to the first stages of our algorithms. The authors use the algorithms to reduce the problem’s size by shortening students’ preference lists. Our goal is to bound the stable assignment cutoffs.

identifies “future unstable demand” at each school by students who have not been considered at the school yet but will apply to the school under DA with any tie-breaker.

**Deferred Rejection Algorithm (DR):**

The following two stages are run in each round until the algorithm terminates.

**Stage 1: Immediate Unstable Demand**

**Step  $t \geq 1$ :** Each student  $s$  applies to her most preferred school which has not rejected her yet. Let  $A_c$  be the set of students applying to school  $c$  in this step. For every student  $s \in A_c$ , her demand for  $c$  is stable among  $A_c$  if  $|\{s' \in A_c : \rho_{s'c} < \rho_{sc}\}| < \kappa_c$ ; otherwise her demand is unstable. Each school  $c$  rejects students in  $A_c$  with unstable demand and provisionally holds the remaining students in  $A_c$ .

Stage 1 proceeds to the next step and terminates when no student is rejected.

By the end of Stage 1, the number of students a school holds may exceed its capacity. Some of these students with the worst priority need to be rejected, who then will try to apply to their next best choice. Stage 2 identifies such students and their unstable demand at their next best choice.

**Stage 2: Future Unstable Demand**

Let  $\mu : S \cup C \rightarrow C \cup 2^S$  denote the outcome of Stage 1.<sup>9</sup>

**Step 0:** Let  $D_c$  denote the students who prefer  $c$  to their tentative assignment at  $\mu$ , i.e.,  $D_c = \{s \in S : c P_s \mu(s)\}$ . For each school  $c$ , we define the school’s threshold priority  $\bar{r}_c(\mu)$  at  $\mu$  as follows:

1. if  $|\mu(c)| \geq \kappa_c$ , then  $\bar{r}_c(\mu) = \max_{s \in \mu(c)} \rho_{sc}$ ,
2. if  $|\mu(c)| < \kappa_c$  and  $D_c = \emptyset$ , then  $\bar{r}_c(\mu) = K + 1$ ,
3. if  $|\mu(c)| < \kappa_c$  and  $D_c \neq \emptyset$ , then  $\bar{r}_c(\mu) = \min_{s \in S_c^c} \rho_{sc}$ .<sup>10</sup>

Find students who will be rejected by  $c'$  and whose next best alternative is  $c$ . To this end, let  $M_{c'}$  be the set of students in  $\mu(c')$  who have priorities equal to  $\bar{r}_c(\mu)$ . Refer to them as marginal students at  $c'$ . Let  $D_{c'}^c \subseteq M_{c'}$  be the subset of marginal students whose next best alternative is  $c$ , i.e., each  $s \in D_{c'}^c$  weakly prefers  $c$  to any school  $c'' \in C \setminus \{c'\}$  which has not rejected her yet. Also define  $g_{c'}^1$  as the number of students in  $\mu(c')$  who have priorities strictly better than  $\bar{r}_c(\mu)$ .

**Step  $t \geq 1$ :** The best priority  $g_{c'}^t$  students in  $\mu(c')$  are guaranteed to be held by  $c'$  at this step. This leaves  $\kappa_{c'} - g_{c'}^t$  seats available for marginal students. Recall that  $D_{c'}^c$  is the set of marginal students at  $\mu(c')$  and whose next best alternative is  $c$ . If  $|D_{c'}^c| > \kappa_{c'} - g_{c'}^t$ , then some of these students must be rejected by  $c'$ . Let  $\tilde{D}_{c'}^c \subseteq D_{c'}^c$  be such that  $|\tilde{D}_{c'}^c| = \max\{0, |D_{c'}^c| - (\kappa_{c'} - g_{c'}^t)\}$  and no student in  $\tilde{D}_{c'}^c$  has strictly better priority than the ones in  $D_{c'}^c \setminus \tilde{D}_{c'}^c$  for school  $c$ . Let  $\tilde{D}^c = \cup_{c' \in C} \tilde{D}_{c'}^c$ . For each school  $c$ , let  $\hat{\rho}_c$  be the priority of  $\kappa_c$ -th best priority student in  $\mu(c) \cup \tilde{D}^c$  if  $|\mu(c) \cup \tilde{D}^c| \geq \kappa_c$ . If  $|\mu(c) \cup \tilde{D}^c| < \kappa_c$ , let  $\hat{\rho}_c = \bar{r}_c(\mu)$ .

<sup>9</sup>Note that  $\mu$  is not an assignment as the number of students at schools may exceed its capacity.

<sup>10</sup>This last case is not relevant in the initial round.

We consider each school  $c \in C$  one by one. If  $|\mu(c) \cup \tilde{D}^c| \geq \kappa_c$ , we set  $g_c^{t+1}$  to the number of students in  $\mu(c) \cup \tilde{D}^c$  who have a strictly better priority than the  $\kappa_c$ -th priority student in  $\mu(c) \cup \tilde{D}^c$ . Otherwise, we set  $g_c^{t+1}$  to  $|\mu(c) \cup \tilde{D}^c|$ . If  $\hat{\rho}_c < \bar{r}_c(\mu)$ , then school  $c$  rejects all students with priorities strictly greater than  $\hat{\rho}_c$  and continue with Stage 1 of the next round. If  $\hat{\rho}_c \geq \bar{r}_c(\mu)$  for all  $c \in C$  and  $g_c^{t+1} > g_c^t$  for some  $c' \in C$ , then we continue with Step  $t + 1$  of Stage 2. Otherwise, the procedure terminates and the final outcome is  $\mu$ .

Given the final outcome  $\mu$  of DR, we calculate  $\bar{r}(\mu)$  as described in Stage 2 (see Step 0) of DR. As stated in our next result,  $\bar{r}(\mu)$  gives an upper bound on cutoff profiles that support a stable assignment.

**Proposition 5.** *If  $\bar{\mu}$  is a stable assignment, then  $\rho(\bar{\mu}) \leq \bar{r}(\mu)$ .*

Here is a brief intuition behind the algorithm and the result. In Stage 1 of the DR algorithm, we run student proposing DA algorithm by only rejecting students who cannot be assigned to the schools they are applying to for any tie-breaker. Among the set of students who are tentatively held by school  $c$  in the outcome of Stage 1, we can determine the number of students who cannot be assigned to  $c$  for any tie-breaker and whose next achievable choice is  $c'$ . By using this information, we further update the set of schools that cannot be achieved by each students in any stable matching in Stage 2 of DR algorithm. This information allows us to rerun Stage 1 and repeat the procedure.

Next, we define the Deferred Proposal Algorithm (DP), which is the schools proposing analog of the of DR. Unlike the school proposing DA, our Deferred Proposal algorithm only proposes to students who clear the school's (admissions) cutoff at any stable assignment. Consequently, the number students tentatively held by each school may be less than the school's capacity. The resulting assignment provides lower bounds for cutoff profiles that support a stable assignment.

DP runs through multiple rounds until no more school is rejected by students. In each round, it goes through two stages. Again, first stage identifies "immediate unstable demand" of every school for students who cannot be assigned to the school at any stable assignment. The second stage identifies "future unstable demand" at each school by students who have not been proposed by the school yet but will have a guaranteed proposal under DA with any tie-breaker.

### ***Deferred Proposal Algorithm (DP)***

*The following two stages are run in each round until the algorithm terminates.*

#### ***Stage 1: Immediate Unstable Demand***

*In the first stage, we run a modified version of school proposing DA. For each school  $c \in C$ , let  $A_c$  denote the set of students who have not rejected school  $c$  yet.*

**Step  $t \geq 1$ :** *Each school  $c$  considers students in  $A_c$  one by one. Each school  $c$  proposes to student  $s \in A_c$  if the number of students in  $A_c$  with weakly better priority than  $s$  is less than or equal to  $\kappa_c$ , i.e.,  $|\{s' \in A_c : \rho_{s'c} \leq \rho_{sc}\}|$ . A school's demand for a student is unstable, if the student receives a proposal from a more preferred school. Each student  $s$  provisionally accepts her most preferred proposing school. All schools with unstable demand are rejected by corresponding students. If there are no more rejection, Stage 1 terminates.*

#### ***Stage 2: Future Unstable Demand***

Let  $\mu : S \cup C \rightarrow C \cup 2^S$  denote the outcome of Stage 1. In this stage, given the outcome of Stage 1, we identify schools that each student is guaranteed under any tie-breaker. For each  $s \in S$  define  $D_s$  as the set of schools preferred to  $\mu(s)$ , i.e.,  $D_s = \{c \in C : c P_s \mu(s)\}$ . Each student  $s$  rejects any school  $c \notin D_s \cup \mu(s)$  and  $s$  is removed from  $A_c$ . We proceed with the following steps:

- (i) Select a student  $s \in S$  such that  $|D_s| \geq 2$  and who has not been considered before. Define  $E_s = (D_s \times D_s) \setminus \cup_{c \in D_s} \{(c, c)\}$ .
- (ii) We consider each pair in  $E_s$  one by one. Let  $(c_1, c_2)$  be the pair under consideration. Let  $\bar{S}$  be the set of students such that for all  $s' \in \bar{S}$  we have  $s' \in A_c$  and  $\rho_{s'c} < \rho_{sc}$  for some  $c \in \{c_1, c_2\}$ .
- (iii) If  $|\bar{S}| \leq \kappa_{c_1} + \kappa_{c_2}$ , then move  $s$  from  $A_c$  for any school  $c$  less preferred than  $c_1$  and  $c_2$ . If  $|\bar{S}| > \kappa_{c_1} + \kappa_{c_2}$  and there is a school pair in  $D_s$  not considered, then go back to bullet (ii). Otherwise, go back to bullet (i).<sup>11</sup>

If  $A_c$  is updated during the Stage 2, we continue with Stage 1 of the next round. Otherwise, the algorithm terminates. Let  $\mu$  denote the outcome of DP.

Due to finiteness of the sets of  $S$  and  $C$ , the algorithm terminates in finite number of rounds. For each  $c \in C$ , we define threshold priority  $\bar{r}_c(\mu)$  as follows:

1. if  $|\mu(c)| = \kappa_c$ , then  $\bar{r}_c(\mu) = \max_{s \in \mu(c)} \rho_{sc}$ ,
2. if  $\mu(c) = \emptyset$ , then  $\bar{r}_c(\mu) = 0$ ,
3. if  $\mu(c) \neq \emptyset$  and  $|\mu(c)| < \kappa_c$ , then  $\bar{r}_c(\mu) = \max_{s \in \mu(c)} \rho_{sc} + 1$ .

As stated in the next result,  $\bar{r}(\mu)$  gives a lower bound on cutoff profiles that support a stable assignment.

**Proposition 6.** *If  $\bar{\mu}$  is a stable matching, then  $\rho(\bar{\mu}) \geq \bar{r}(\mu)$ .*

Here is a brief intuition behind the algorithm and the result. In Stage 1 of the DP algorithm, each school  $c$  proposes to the set of highest ranked students in its priority order up to its capacity (possibly fewer than its capacity) without using a tie-breaker. As a result, when Stage 1 terminates, some school  $c$ 's proposals might be accepted by strictly less than  $\kappa_c$  students. Some students who have not been proposed during Stage 1 might prefer  $c$  to their match. Such students can be considered as the students in wait lists of the schools. In Stage 2, we determine the worst school that a student can guarantee by considering the overlaps in schools' wait lists. Hence, we can eliminate some schools from students consideration and repeat the procedure by using this updated information.

We illustrate DR and DP through an example in Appendix B.

<sup>11</sup>We can do this stage for any subset of schools, instead of pairs only, to determine the guaranteed schools for the students. The cost of increasing the number of schools is increasing the computational burden.



## 6 Random Search Algorithm for Large School Districts

Our solution in Section 4 traces all possible cutoff vectors to find a match quality optimal assignment. Doing so is a computationally demanding task. Although the algorithms introduced in Section 5 diminish the set of cutoff vectors to be considered, for large schools districts we may still face problems due to computational tractability. In this section, we introduce a polynomial time algorithm that maximizes aggregate match quality in a subclass of stable assignments. We start with a cutoff profile resulting from student proposing DA with a random tie-breaker. Then, we find a locally match quality optimal assignment for cutoff profile using the minimum-cost flow algorithm of Section 4.1. The cutoff profile of the resulted stable assignment may be different from the original one. In that case, we repeat the local optimization, and check the new cutoff profile. When the cutoff profile does not change at some step, we terminate the procedure. Here is the formal description of the algorithm.

### *Locally Match Quality Optimal Algorithm (L-MQO)*

**Step 0:** Let  $\mu^*$  be the outcome of student proposing DA with a random tie-breaker. Let  $r_0$  be the cutoff profile of  $\mu^*$ , i.e.,  $r_0 = \rho(\mu^*)$ .

**Step  $t \geq 1$ :** Let  $\mu_{r_{t-1}}$  be the locally match quality optimal assignment for  $r_{t-1}$ , found by the minimum-cost flow algorithm described in Section 4.1.<sup>12</sup> and let  $r_t = \rho(\mu_{r_{t-1}})$  be its cutoff profile. If  $\sum_{s \in S} q(s, \mu_{r_{t-1}}(s)) > \sum_{s \in S} q(s, \mu^*(s))$ , we set  $\mu^* = \mu_{r_{t-1}}$ . If  $r_t = r_{t-1}$ , then the algorithm terminates and its outcome is assignment  $\mu^*$ . Otherwise, continue with Step  $t + 1$ .

By construction of  $\bar{\mathcal{A}}_r$ ,  $r_t = \rho(\mu_r^t) \leq r_{t-1}$  for each  $t > 1$ , i.e., the cutoff profile increases throughout the algorithm. Therefore, the number of steps it takes the algorithm to terminate is bounded by  $|K| \times |C|$ , which means that L-MQO terminates in polynomial time.

Let  $\bar{R}$  denote the set of vectors that have been considered during the implementation of L-MQO, i.e.,  $\bar{R} = \{r_t\}_{t=1}^T$ , where  $T$  is the last step of L-MQO. The outcome  $\mu^*$  of the L-MQO is not necessarily match quality optimal. However, it is immediate from its description that the algorithm creates weakly higher aggregate match quality than any assignment in  $\cup_{r \in \bar{R}} \bar{\mathcal{A}}_r$ . Therefore, we can state the following result.

**Proposition 7.** *Suppose there is a match quality optimal assignment  $\mu$  such that  $\rho(\mu) \in \bar{R}$ . Then, L-MQO gives a match quality optimal assignment.*

In simulations of Section 7 we report the results for a simpler version of the procedure that only includes Steps 0 and 1, i.e., locally match quality optimal assignment is only found in the cutoffs identified by the DA. Even for this simple version, the aggregate match quality is remarkably close to the (globally) match quality optimal assignment.

<sup>12</sup>A locally match quality optimal assignment exists since  $\bar{\mathcal{A}}_{r_{t-1}} \neq \emptyset$  for any  $t \geq 1$ . In fact,  $\mu^* \in \bar{\mathcal{A}}_{r_0}$  and  $\mu_{r_{t-2}} \in \bar{\mathcal{A}}_{r_{t-1}}$  for  $t \geq 2$ .

## 7 Simulations

### 7.1 Setting

The algorithm described in Section 4.2 finds a match quality optimal assignment by searching in the set of all cutoff profiles. Without bounding the set of cutoff profiles to be searched within, we may need to consider  $(K + 1)^{|C|}$  cutoff profiles. Section 5 introduces two algorithms that decrease the number of cutoff profiles supporting stable matchings. In this section, by using computer simulations we first measure the possible reduction in the number of cutoff profiles supporting a stable assignment can be done by DR and DP. Then, we find a match quality optimal stable assignment and compare it with the student proposing DA outcome with a random tie-breaker.

In our simulations, we consider an environment that mimics a standard school choice setting. In particular, we consider environments with 20 and 30 schools each with 50 available seats. As explained in Introduction, these numbers are representative for a moderately large school district.<sup>13</sup> The number of students is equal to the total capacities of the schools.

In the construction of school priorities we use two criteria: sibling status and distance between students and schools. We set the fraction of students with sibling status to 0.4 and determine the students with sibling priority and at which school they have the sibling priority randomly. Let  $sib : S \times C \rightarrow \{0, 1\}$  be an indicator function such that  $sib(s, c) = 1$  means student  $s$  has sibling priority at  $c$  and  $\sum_{c \in C} sib(s, c) \leq 1$  for all  $s \in S$ .

In order to determine the neighborhood (walk-zone) priority, we randomly distribute schools and students on an  $1 \times 1$  unit map. In particular, we represent the location of agent  $i \in C \cup S$  with  $\ell_i = (\ell_i^1, \ell_i^2)$  and both  $\ell_i^1$  and  $\ell_i^2$  are i.i.d. standard uniformly distributed random variables. Given the locations of each student  $s$  and school  $c$ , we calculate the euclidean distance and denote it with  $d(s, c)$ , i.e.,  $d(s, c) = \sqrt{(\ell_s^1 - \ell_c^1)^2 + (\ell_s^2 - \ell_c^2)^2}$ . We set the neighborhood radius of each school to 0.2. If  $d(s, c) \leq 0.2$ , then student  $s$  has neighborhood priority at school  $s$ .

Given the sibling and neighborhood status, we group students into four groups for each school. For each school  $c \in C$ , students having sibling priority at school  $c$  are ranked higher than students without sibling priority at  $c$ . Then, we subgroup students based on the neighborhood priority. That is, first priority group is composed of students with sibling and neighborhood priority. Second priority group is composed of students with sibling priority but not neighborhood priority. Third priority group is composed of students with neighborhood priority but not sibling priority. The remaining students constitute the fourth priority group.

Preferences of students are constructed by taking various criteria into account including: students' common and individual taste over schools, having siblings, and distance. To construct the preference of each student  $s$  we calculate her utility from being assigned to each school  $c$ , denoted by  $U_{sc}$ , as follows:

$$U_{sc} = \alpha X_c + (1 - \alpha) Y_{sc} + \beta sib(s, c) - \gamma dist(s, c),$$

where  $X_c \in (0, 1)$  and  $Y_{sc} \in (0, 1)$  represents all students common taste and student  $s$ 's in-

---

<sup>13</sup>For example, Boston Public Schools has around 30 high schools.

dividual taste over school  $c$ , respectively. Both  $X_c$  and  $Y_{sc}$  are i.i.d. standard uniformly distributed random variables. The level of correlation in the preferences of students is captured by  $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$ , i.e., as  $\alpha$  increases preferences become more correlated. Variable  $sib(s, c)$  takes value 0 or 1 and coefficient  $\beta \in \{0, 0.25, 0.5, 0.75, 1\}$  is the additional utility received from attending the same school with sibling. Finally, travel time is usually considered as disutility for the students and we capture this by multiplying distance between students and schools with coefficient  $-\gamma$ .<sup>14</sup> The utility values of students are used to construct the ordinal preferences of students over the schools. Match quality of each student-school pair is drawn i.i.d. uniformly from the  $(0, 1)$  interval. Thus, in our simulations preferences are uncorrelated with the idiosyncratic match quality, which is consistent with the empirical evidence in Abdulkadiroğlu et al. (2020).

## 7.2 Reductions in Cutoff Profiles by DR and DP

In simulations, under each scenario we calculate the outcomes of the DR and DP algorithms 100 times by using different draws for  $X$ ,  $Y$ , and locations. Table 1 presents the average number of schools having a unique cutoff at every stable assignment when the number of schools is 20. In Appendix C we report results for 30 schools.

In our setting there are enough seats to accommodate all students, therefore all schools fill up their seats in any stable assignment. This means that there are at most  $4^{20}$  cutoff profiles that our optimization algorithm need to consider to find a match quality optimal assignment. However, in our setup, for each school  $c$ , the number of students having sibling priority does not exceed the capacity of the school. Therefore, it is immediate that for any problem DR and DP eliminate all cutoff profiles where a school has an admission cutoff equal to 1 or 2. This reduces the number of cutoff profiles to consider to  $2^{20}$ . Moreover, whenever a school is guaranteed to have a unique cutoff under any stable assignment, that school does not cause any additional computational burden. Our simulations show that for parameter values  $\alpha \geq 0.5, \beta \geq 0.5, \gamma \geq 0.25$ , DR and DP eliminate at least around 97% of the remaining cutoff profiles. Comparing columns (1)-(3) in Table 1, we can see that as  $\gamma$  increases, the reduction algorithms give tighter bounds in most of the cases.

In our simulations, all students rank all schools acceptable and the total number of seats at schools have enough seats to accommodate all students. This fact allows us not to consider cutoffs equal to five, and potentially reduces the computational burden. Although, assuming more school seats than applicants is a realistic assumption in the public school choice setting, school districts commonly restrict the number of choices in the students' preference lists, therefore cutoffs at some schools can equal to five at some stable assignments. However, in the environment with restricted lists computational burden is not necessarily heavier than with unrestricted ones. The reason behind this is that when preference lists are restricted many schools do not fill their seats at any stable assignment, and DR and DP algorithms partially identify some of those schools. As a result, the number of cutoff profiles that can support stable assignments in a setting with restricted lists may decrease compared to our environment.

---

<sup>14</sup>When  $\gamma = 0$ , travel distance does not have affect on preferences.

$\alpha$	$\beta$	$\gamma = 0.00$	$\gamma = 0.25$	$\gamma = 0.50$
		(1)	(2)	(3)
0.00	0.00	0	0	0
	0.25	0	0.01	0
	0.50	0	0.52	0
	0.75	0	1.49	0.02
	1.00	0	2.93	0.03
0.25	0.00	0	0	0.14
	0.25	0	0.22	0.28
	0.50	0	1.91	0.67
	0.75	0	3.06	0.77
	1.00	0	3.33	0.39
0.50	0.00	0.15	1.31	5.30
	0.25	0.59	2.77	6.57
	0.50	0.44	5.18	6.41
	0.75	0.20	5.42	5.56
	1.00	0.21	5.31	4.69
0.75	0.00	2.63	4.63	7.54
	0.25	4.42	5.88	7.85
	0.50	2.96	7.05	8.03
	0.75	2.26	8.28	7.87
	1.00	2.21	8.68	7.80
1.00	0.00	7.11	5.65	7.81
	0.25	7.42	6.35	8.09
	0.50	7	6.98	8.28
	0.75	6.55	8.06	8.12
	1.00	6.66	9.31	8.02

Table 1: Average number of schools with a unique stable assignment cutoff, 20 schools

### 7.3 Comparing Match Quality

We compare aggregate match quality across stable algorithms when the number of schools is 20. In Table 2 we report the results for  $\gamma = 0.25$ , and all combinations of  $\alpha$  and  $\beta$ . Results for  $\gamma = 0$  and  $\gamma = 0.5$  are reported in Appendix C.

Numbers in Table 2 denote percentage gains in match quality compared to student proposing DA algorithm with a random tie-breaking. Columns (1) and (2) report the results for MQO and L-MQO algorithms, respectively. The last row denotes the average percentage gains in match quality across all possible values of  $\alpha$  and  $\beta$  parameters. The match quality gains from our optimization algorithms are remarkable: both MQO and L-MQO create around 40% higher match quality compared to DA with random tie-breaking. Moreover, the performance of L-MQO is almost indistinguishable from that of MQO. Additionally, in Column (3) we report the result for DA algorithm when ties are broken according to the school specific match quality. We call this algorithm DA with quality-based tie-breaking. Such a heuristic solution significantly improves on DA with a random tie-breaking. However, the match quality gains are less than two thirds of those of MQO and L-MQO.

$\alpha$	$\beta$	MQO (1)	L-MQO (2)	DA with quality-based tie-breaking (3)
0.00	0.00	50.044	50.044	22.746
	0.25	44.816	44.816	20.134
	0.50	39.682	39.682	14.414
	0.75	35.708	35.708	16.040
	1.00	32.496	32.426	16.046
0.25	0.00	44.224	44.224	24.498
	0.25	38.978	38.978	21.406
	0.50	34.510	33.860	18.134
	0.75	30.194	29.868	16.754
	1.00	29.906	29.540	16.024
0.50	0.00	49.878	49.728	28.766
	0.25	43.326	43.326	24.994
	0.50	36.890	36.890	21.486
	0.75	34.570	34.570	20.824
	1.00	34.154	34.154	19.738
0.75	0.00	50.252	50.066	31.142
	0.25	44.682	44.682	29.690
	0.50	41.390	40.948	26.892
	0.75	37.620	36.070	23.922
	1.00	37.154	37.154	24.398
1.00	0.00	50.236	46.780	31.112
	0.25	46.632	46.632	29.480
	0.50	43.698	43.698	29.984
	0.75	40.018	39.612	26.390
	1.00	38.070	36.806	25.124
Average gain		40.365	40.010	23.206

Table 2: Percentage gains compared to student proposing DA with a random tie-breaking,  $\gamma = 0.25$

## 8 Discussion

Student proposing DA with a random tie-breaker assigns students to schools solely based on preference reports and priorities. Given the empirical evidence that parents' preferences may not reflect match quality (Abdulkadiroğlu et al., 2020), those solutions are likely to lead to ineffective assignments. We provide a novel match quality-based solution to the school choice problem that maximizes aggregate match quality among all stable assignments. Simulations show that our algorithms substantially improve match quality compared to the widely used DA algorithm.

Our solution partially addresses another potential challenge in school choice. There is an ongoing academic debate on whether parents form their preferences based on peer composition and achievement, or whether they align with effectiveness (Hanushek, 1981; Jacob and Lefgren, 2007; Abdulkadiroğlu et al., 2020). Moreover, even when parents value effectiveness, informational and cognitive barriers may preclude separation of a school's effectiveness from achievement of its student body (Kane and Staiger, 2002). Then, higher demand for schools that recruit higher-achieving students may create incentives for school principals to devote resources to screening and selection rather than better instruction (Ladd, 2002; MacLeod and Urquiola, 2015). Our solution does not fully address this issue, as preferences still play a role in determining the final assignment. However, the issue is likely to be alleviated: match quality- or effectiveness-based assignment gives schools incentives to improve and become more effective.

Although our primary motivation is finding a match quality optimal assignment, the methodology can be easily applied to optimize other policy objectives. For instance, student welfare has been an important consideration in school choice. Typically, there are multiple stable assignments to choose from, and there is no obvious selection rule for a policy maker who cares for both stability and student welfare. Erdil and Ergin (2008) provide an algorithm that finds a stable assignment that is not Pareto dominated by other stable assignment. Our method allows to find stronger results. For example, with an appropriate choice of match quality matrix, our method can be used to find stable assignment that maximizes the number of students that are assigned to their first choices. Another common policy objective is achieving desirable distributional outcomes, such as a more diverse student body or more students assigned to neighborhood schools. Such goals, too, can be achieved using our optimization method with the appropriate choice of the match quality matching. Finally, it is worth mentioning that the applicability of our solution is broader than the school choice problem. Our optimization algorithms under stability constraints may be applied to an arbitrary two-sided or priority-based matching problem.

## References

- ABDULKADIROĞLU, A. (2005): “College Admission with Affirmative Action,” *International Journal of Game Theory*, 33, 535–549.
- ABDULKADIROĞLU, A., J. D. ANGRIST, Y. NARITA, AND P. A. PATHAK (2017): “Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation,” *Econometrica*, 85, 1373–1432.
- ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2009): “Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match,” *American Economic Review*, 99(5), 1954–1978.
- ABDULKADIROĞLU, A. AND T. SÖNMEZ (2003): “School Choice: A Mechanism Design Approach,” *American Economic Review*, 93, 729–747.
- ABDULKADIROĞLU, A. A., P. A. PATHAK, J. SCHELLENBERG, AND C. R. WALTERS (2020): “Do Parents Value School Effectiveness?” *American Economic Review*.
- ABDULKADIROĞLU, A., Y.-K. CHE, AND Y. YASUDA (2015): “Expanding “Choice” in School Choice,” *American Economic Journal: Microeconomics*, 7, 1–42.
- AHUJA, R. K., T. L. MAGNANTI, AND J. B. ORLIN (1993): *Network Flows: Theory, Algorithms and Application*.
- ASHLAGI, I. AND P. SHI (2016): “Optimal Allocation without Money: An Engineering Approach,” *Management Science*, 62, 1078–1097.
- AZEVEDO, E. M. AND J. D. LESHNO (2016): “A Supply and Demand Framework for Two-sided Matching Markets,” *Journal of Political Economy*, 124, 1235–1268.
- BALINSKI, M. AND T. SÖNMEZ (1999): “A Tale of Two Mechanisms: Student Placement,” *Journal of Economic Theory*, 84, 73–94.
- BANSAK, K., J. FERWERDA, J. HAINMUELLER, A. DILLON, D. HANGARTER, D. LAWRENCE, AND J. WEINSTEIN (2018): “Improving Refugee Integration through Data-driven Algorithmic Assignment,” *Science*, 359, 325–329.
- BODOH-CREED, A. L. (2020): “Optimizing for Distributional Goals in School Choice Problems,” *Management Science*.
- CHANDRAMOULI, S. AND J. SETHURAMAN (2020): “A Note on the Rationing of Divisible and Indivisible Goods in a General Network,” *Working Paper*.
- DUR, U., S. D. KOMINERS, P. A. PATHAK, AND T. SÖNMEZ (2018): “Reserve Design: Unintended Consequences and the Demise of Boston’s Walk Zones,” *Journal of Political Economy*, 6, 2457–2479.
- ERDIL, A. AND H. ERGIN (2008): “What’s the Matter with Tie-Breaking? Improving Efficiency in School Choice,” *American Economic Review*, 98, 669–689.
- (2017): “Two-sided Matching with Indifferences,” *Journal of Economics Theory*, 171, 268–292.



- FRIEDMAN, M. (1962): “Capitalism and Freedom: With the Assistance of Rose D. Friedman,” *University of Chicago Press*.
- GALE, D. AND L. S. SHAPLEY (1962): “College Admissions and the Stability of Marriage,” *American Mathematical Monthly*, 69, 9–15.
- GRIGORYAN, A. (2020): “Effective, Fair and Equitable Pandemic Rationing,” *Working Paper*.
- HAFALIR, I. E., M. B. YENMEZ, AND M. A. YILDIRIM (2013): “Effective Affirmative Action in School Choice,” *Theoretical Economics*, 8, 325–363.
- HANUSHEK, E. A. (1981): “Throwing Money at Schools,” *Journal of Policy Analysis*, 1, 19–41.
- HOXBY, C. (2000): “Does Competition among Public Schools Benefit Students and Taxpayers?” *American Economic Review*, 90(5), 1209–1238.
- (2003): “School Choice and School Productivity (Or, Could School Choice be a Rising Tide that Lifts All Boats),” in *The Economics of School Choice*, ed. by C. Hoxby, Chicago: University of Chicago Press.
- IRVING, R. W. (1994): “Stable Marriage and Indifference,” *Discrete Applied Mathematics*, 48, 261–272.
- IWAMA, K., S. MIYAZAKI, AND H. YANAGISAWA (2014): “A 25/17-approximation Algorithm for the Stable Marriage Problem with One-sided Ties,” *Algorithmica*, 68, 758–775.
- JACOB, B. A. AND L. LEFGREN (2007): “What do Parents Value in Education? An Empirical Investigation of Parents’ Revealed Preferences for Teachers,” *The Quarterly Journal of Economics*, 122, 1603 – 1637.
- KANE, T. J. AND D. O. STAIGER (2002): “The Promise and Pitfalls of Using Imprecise School Accountability Measures,” *Journal of Economic Perspectives*, 16(4), 91–114.
- KATTA, A.-K. AND J. SETHURAMAN (2006): “A Solution to the Random Assignment Problem on the Full Preference Domain,” *Journal of Economic Theory*, 131(1), 231–250.
- KESTEN, O. AND U. ÜNVER (2015): “A Theory of School Choice Lotteries,” *Theoretical Economics*, 10(2), 543–595.
- KIM, Y. C. J. AND K. MIERENDORFF (2013): “Generalized Reduced-Form Auctions: A Network-Flow Approach,” *Econometrica*, 81, 2487–2520.
- KIRÁLY, Z. (2011): “Better and Simpler Approximation Algorithms for the Stable Marriage Problem,” *Algorithmica*, 60, 3–20.
- KWANASHIE, A. AND D. F. MANLOVE (2014): “An Integer Programming Approach to the Hospitals/residents Problem with Ties,” *In Operations Research Proceedings*, 263–269.
- LADD, H. F. (2002): *Market-Based Reforms in Urban Education*, Economic Policy Institute.
- MACLEOD, B. W. AND M. URQUIOLA (2015): “Reputation and school competition,” *American Economic Review*, 105, 3471 – 3488.
- MANLOVE, D. F., R. W. IRVING, K. IWAMA, S. MIYAZAKI, AND Y. MORITA (2002): “Hard Variants of Stable Marriage,” *Theoretical Computer Science*, 276, 261–279.

- MCDERMID, E. (2009): “A  $3/2$ -approximation Algorithm for General Stable Marriage,” *In International Colloquium on Automata, Languages, and Programming*, 689–700.
- NGUYEN, T., H. NGUYEN, AND A. TEYTELBOYM (2020): “Stability in Matching Markets with Complex Constraints,” *Management Science*.
- NGUYEN, T. AND R. VOHRA (2013): “The Allocation of Indivisible Objects via Rounding,” *Working Paper*.
- (2018): “Near-feasible Stable Matchings with Couples,” *American Economic Review*, 108, 3154–3169.
- (2019): “Stable Matching with Proportionality Constraints,” *Operations Research*, 67, 1503–1519.
- PETERSON, P. E. AND D. E. CAMPBELL (2001): *Charters, Vouchers, and Public Education*, Brookings Institution Press.
- ROTH, A. E. (1985): “The College Admission Problem is not Equivalent to the Marriage Problem,” *Journal of Economic Theory*, 36, 277–288.
- ROTH, A. E., U. G. ROTHBLUM, AND J. H. V. VATE (1993): “Stable Matchings, Optimal Assignments, and Linear Programming,” *Mathematics of Operations Research*, 18, 803–828.
- SHI, P. (2019): “Optimal Priority-Based Allocation Mechanisms,” *Working Paper*.
- SLAUGH, V. W., M. AKAN, O. KESTEN, AND M. U. ÜNVER (2016): “The Pennsylvania Adoption Exchange Improves its Matching Process,” *Interfaces*, 46, 133–153.
- SOKKALINGAM, P. T., R. K. AHUJA, AND J. B. ORLIN (2000): “New Polynomial-time Cycle-canceling Algorithms for Minimum-cost Flows,” *Networks: An International Journal*, 36, 53–63.
- TRAPP, A. C., A. TEYTELBOYM, A. MARTINELLO, T. ANDERSSON, AND N. AHABI (2018): “Placement Optimization in Refugee Resettlement,” *Working Paper*.
- TWEEDIE, J., D. D. RILEY, J. E. CHUBB, AND T. M. MOE (1990): “Should Market Forces Control Educational Decision Making?” *The American Political Science Review*, 549–567.

## A Proofs of Main results

### A.1 Proof of Lemma 1

First we prove  $\mathcal{A}_r \subseteq \bar{\mathcal{A}}_r$ . Suppose  $\mu \in \mathcal{A}_r$ . We show that  $\mu$  satisfies conditions in the definition of  $\bar{\mathcal{A}}_r$  one by one.

We start with the first condition. On the contrary, suppose the first condition does not hold. That is, there exists a student  $s$  such that  $c = \mu(s) \notin C_s(r)$ . By definition of  $\rho(\mu)$ ,  $\rho_{sc} \leq \rho_c(\mu) = r_c$ . Therefore,  $c \notin C_s(r)$  implies that there is a school  $c' \in C$  such that  $c' P_s c$  and  $\rho_{sc'} < r_{c'} = \rho_{c'}(\mu)$ . By Proposition 2, this contradicts the stability of  $\mu$ .

We continue with the second condition. If  $C^+(r) = \emptyset$ , then second condition holds trivially. Suppose  $C^+(r) \neq \emptyset$  and  $c \in C^+(r)$ . By definition of  $C^+(r)$ ,  $\rho_c(\mu) = r_c < K + 1$ . Hence, by definition of cutoffs,  $|\mu(c)| = \kappa_c$ . This completes the proof of the first part, i.e.,  $\mathcal{A}_r \subseteq \bar{\mathcal{A}}_r$ .

Next we prove  $\bar{\mathcal{A}}_r \subseteq \mathcal{A}$ . Suppose  $\mu \in \bar{\mathcal{A}}_r$ . We first show that  $\rho(\mu) \leq r$ . By definition of  $C^-(r)$ ,  $r_c = K + 1$  for any  $c \in C^-(r)$ . Hence,  $\rho_c(\mu) \leq K + 1 = r_c$  for any  $c \in C^-(r)$ .

Now consider a school  $c \in C^+(r)$ . Recall that,  $|\mu(c)| = \kappa_c$ . Therefore, by definition of  $\rho_c(\mu)$ ,

$$\rho_c(\mu) = \max_{s \in \mu(c)} \rho_{sc}. \quad (3)$$

By definition of  $\bar{\mathcal{A}}_r$ ,

$$\rho_{sc} \leq r_c, \forall s \in \mu(c). \quad (4)$$

Equations 3 and 4 imply that  $\rho_c(\mu) \leq r_c$ .

We now show that  $\mu$  is stable. Consider an arbitrary  $s \in S$  and  $c = \mu(s)$ . By definition of  $\bar{\mathcal{A}}_r$ , there is no  $c' \in C$  such that  $c' P_s c$  and  $\rho_{sc'} < r_{c'}$ . Since  $\rho_{c'}(\mu) \leq r_{c'}$ , there is no  $c' \in C$  such that  $c' P_s c$  and  $\rho_{sc'} < \rho_{c'}(\mu)$ . Hence, by Proposition 2,  $\mu$  is stable, i.e.,  $\mu \in \mathcal{A}$ .

### A.2 Proof of Lemma 2

First, suppose  $f$  is feasible and integral. By constraint 1,

$$\sum_{c \in C_r(s)} f(s, c) = b(s) = 1 \text{ for any student } s \in S.$$

Thus,  $\mu_f$  is indeed an assignment, and  $\mu_f(s) \in C_s(r)$  for any  $s \in S$ . This establishes the first condition in the definition of  $\bar{\mathcal{A}}_r$ . Also, for any  $c \in C^+(r)$ ,

$$|\mu_f^{-1}(c)| = - \sum_{s \in S} f(s, c) = b(c) = -\kappa_c,$$

where the second equality follows from constraint 1. This establishes the second condition in the definition of  $\bar{\mathcal{A}}_r$ .

Now suppose  $\mu \in \bar{\mathcal{A}}_r$ . Integrality of  $f_\mu$ , as well as equation 2 are immediate from the construction of  $f_\mu$ . We now verify feasibility constraint 1. We check feasibility for each vertex type one by one.

- Let  $v \in S$ . By definition, each student is assigned to exactly one school. Therefore,

$$\sum_{c \in C} f_{\mu}(v, c) = 1 = b(v).$$

- Let  $v \in C^-(r)$ . Then,

$$\sum_{s \in S} f_{\mu}(s, v) - f_{\mu}(v, t) = |\mu(v)| - |\mu(v)| = 0 = b(v).$$

- Let  $v \in C^+(r)$ . Then,

$$-\sum_{s \in S} f_{\mu}(s, v) = -|\mu(v)| = -\kappa_c = b(v),$$

where second equality follows from  $v \in C^+(r)$ .

- Finally, let  $v = t$ . Then,

$$\begin{aligned} -\sum_{c \in C^-(r)} f(c, v) &= -\sum_{c \in C^-(r)} |\mu(c)| = -\sum_{c \in C} |\mu(c)| + \sum_{c \in C^+(r)} |\mu(c)| \\ &= -|S| + \sum_{c \in C^+(r)} \kappa_c = b(v). \end{aligned}$$

This completes the proof of Lemma 2.

### A.3 Proof of Proposition 5

We first show that if a student  $s$  is rejected from school  $c$  during DR, then there is no stable assignment  $\bar{\mu}$  such that  $\bar{\mu}(s) = c$ . By contradiction, suppose our claim does not hold. That is, there exists a stable assignment  $\bar{\mu}$ , a student  $s$  and a school  $c$  such that  $c$  rejects  $s$  during DR and  $\bar{\mu}(s) = c$ . Without loss of generality, we assume  $s$  is the first such student who is rejected from her assignment under  $\bar{\mu}$  during DR, i.e., if  $s'$  is rejected by  $c'$  before  $s$  is rejected by  $c$  during DR, then  $\bar{\mu}(s') \neq c'$ . Let  $c$  rejects  $s$  in some Round  $m$ . We consider the following possible cases.

*Case 1:  $c$  rejects  $s$  in Stage 1 of Round  $m$ .* Suppose that  $s$  is rejected by  $c$  in Step  $t$ . Then,  $s \in A_c$  and by our supposition  $c$  rejects  $\bar{\mu}(s')$  for all  $s' \in A_c$ . Since  $s$  is rejected by  $c$  in Step  $t$  of Stage 1, we have  $|\{s' \in A_c : \rho_{s'c} < \rho_{sc}\}| \geq \kappa_c$ . Stability of  $\bar{\mu}$  implies that any student  $s'' \in \{s' \in A_c : \rho_{s'c} < \rho_{sc}\}$  is assigned to  $c$  under  $\bar{\mu}$ . Then,  $|\bar{\mu}^{-1}(c)| \geq \kappa_c + 1$ , which contradicts the feasibility of  $\bar{\mu}$ .

*Case 2:  $c$  rejects  $s$  in Step 0 of Stage 2 of Round  $m$ .* Let  $\mu$  be the outcome achieved at the end of Stage 1 in Round  $m$ . Since  $c$  rejects  $s$  in Step 0, then  $\rho_{sc} > \bar{r}_c(\mu)$ . Then, either  $|\mu(c)| \geq \kappa_c$  and  $\bar{r}_c(\mu) = \max_{s' \in \mu(c)} \rho_{s'c}$  or  $D_c \neq \emptyset$  and  $\bar{r}_c(\mu) = \min_{s' \in D_c} \rho_{s'c}$ . If the former case holds, our supposition and stability of  $\bar{\mu}$  imply that all students in  $\mu(c)$  and  $s$  are assigned to  $c$ . This requires at least  $\kappa_c$  students to be assigned to  $c$ , violation of feasibility. Suppose the latter case holds. Our supposition implies that any student in  $D_c$  who has been rejected by  $c$  before  $s$  cannot be assigned to a school weakly better than  $c$ . Since  $\bar{r}_c(\mu) = \min_{s' \in D_c} \rho_{s'c} < \rho_{sc}$ ,  $\bar{\mu}$  cannot be stable, a contradiction.

*Case 3:  $c$  rejects  $s$  in some Step  $t \geq 1$  of Stage 2 of Round  $m$ .* Then, there exists at least  $\kappa_c$  students in  $\mu(c) \cup \tilde{D}^c$  who have strictly better priority than  $s$ . By our supposition at least  $\kappa_c$  students having better priority than  $\rho_{sc}$  cannot be assigned to schools weakly better than  $c$  under  $\bar{\mu}$ . Stability of  $\bar{\mu}$  requires more than  $\kappa_c$  students to be assigned to  $c$ , a contraction.

Hence, if a school  $c$  has rejected student  $s$  during DR, student  $s$  cannot be assigned to school  $c$  in any stable matching. Thus, no student  $s$  with  $\rho_{s'c} > \bar{r}_c(\mu)$  can be assigned to  $c$  in any stable assignment. This concludes the proof.

## A.4 Proof of Proposition 6

In order to prove this statement, we show that if a student  $s$  rejects a school  $c$  during DP, then there does not exist a stable matching  $\bar{\mu}$  such that  $c R_s \bar{\mu}(s)$ .

By contradiction, suppose a student  $s$  who rejects school  $c$  during this procedure is assigned to a school weakly worse than  $c$  under  $\bar{\mu}$ . Without loss of generality, let  $s$  be the first such a student who has rejected some school  $c$  and is assigned to a worse school under  $\bar{\mu}$  when we apply DP algorithm. We consider the following possible cases.

*Case 1:  $s$  rejects  $c$  in Stage 1 of Round  $m$ .* Suppose  $s$  rejects  $c$  in Step  $t$ . Then, by definition  $s \in A_c$  and she has received an offer from  $c'$  such that  $c' P_s c$  in Step  $t$  and the number of students with better priority than  $s$  for  $c'$  and who has not rejected  $c'$  yet is strictly less than  $\kappa_{c'}$ . By our supposition, any student who has rejected  $c'$  earlier cannot be assigned to  $c'$  under  $\bar{\mu}$ . Therefore, stability implies that  $s$  cannot be assigned to a school worse than  $c'$  under  $\bar{\mu}$ . This is a contradiction.

*Case 2:  $s$  rejects  $c$  in Stage 2 of Round  $m$ .* By definition of the procedure, there exists at least two schools  $c_1$  and  $c_2$  such that the number of students with better priority than  $s$  either for  $c_1$  or  $c_2$  who has not rejected them yet is strictly less than the total capacity of  $c_1$  and  $c_2$ . By our supposition, any student who has rejected either school before cannot be assigned to these schools. Therefore, by our supposition,  $s$  and either  $c_1$  or  $c_2$  would form a blocking pair at assignment  $\bar{\mu}$  which contradicts stability of  $\bar{\mu}$ .

Then, our claim implies that if a student  $s \in \mu(c)$ , then at most  $\kappa_c - 1$  students with weakly better priority than  $s$  can be assigned to  $c$  in stable matching  $\bar{\mu}$ . Therefore,  $\bar{\mu}(s) R_s \mu(s)$ . This completes the proof.

## B Omitted Examples

**Example 1.** Let  $C = \{c_1, c_2\}$ ,  $S = \{s_1, s_2, s_3\}$  and  $\kappa = (2, 2)$ . Preferences of students are:

$s_1$	$s_2$	$s_3$
$c_1$	$c_1$	$c_2$
$c_2$	$c_2$	$c_1$

There are 2 priority classes and school priorities are:

Priority Points	$c_1$	$c_2$
1	$s_1$	$s_1, s_3$
2	$s_2, s_3$	$s_2$

Match qualities are:

$$\begin{aligned}
 q(s_1, c_1) &= 3 & q(s_2, c_1) &= 5 & q(s_3, c_1) &= 2 \\
 q(s_1, c_2) &= 4 & q(s_2, c_2) &= 2 & q(s_3, c_2) &= 5
 \end{aligned}$$

Consider the vector  $r = (2, 3)$ . Then,

$$\begin{aligned}
 C_{s_1}(r) &= \{c_1\} & C^+(r) &= \{c_1\} \\
 C_{s_2}(r) &= \{c_1, c_2\} & C^-(r) &= \{c_2\} \\
 C_{s_3}(r) &= \{c_2\}
 \end{aligned}$$

The corresponding minimum-cost flow graph is depicted in Figure 1.

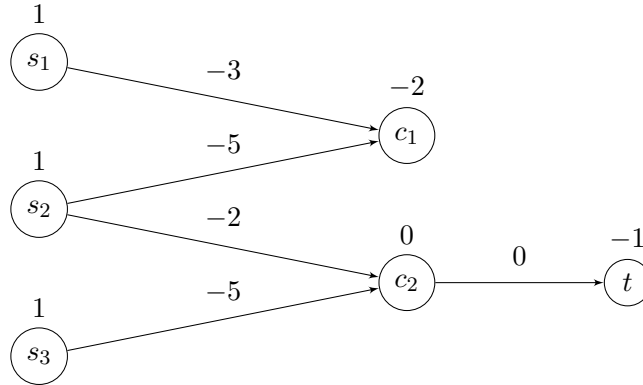


Figure 2: Minimum-cost flow graph

In Figure 2, numbers above the vertices denote their values (supply/demand), numbers above the edges denote their costs.

In this example, there is a unique feasible flow function  $f$ , given by

$$f(e) = \begin{cases} 0 & \text{if } e = (s_2, c_1) \\ 1 & \text{otherwise} \end{cases} .$$

Therefore,  $f$  is the desired solution.

**Example 2.** Let  $C = \{c_1, c_2, c_3, c_4, c_5\}$ ,  $S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$  and  $\kappa = (1, 1, 1, 1, 2)$ . Preferences of students are:

$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$
$c_1$	$c_1$	$c_2$	$c_2$	$c_1$	$c_2$
$c_2$	$c_2$	$c_3$	$c_1$	$c_3$	$c_3$
$c_4$	$c_3$	$c_5$	$c_3$	$c_2$	$c_1$
$c_3$	$c_5$	$c_1$	$c_4$	$c_4$	$c_4$
$c_5$	$c_4$	$c_4$	$c_5$	$c_5$	$c_5$

There are 4 priority classes and school priorities are:

Priority Points	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
1	$s_4$	$s_1$	$s_5, s_6$	$s_1, s_5, s_6$	$s_2$
2	$s_1, s_2$	$s_2$	$s_4, s_1$	$s_2, s_3$	$s_3, s_4, s_5$
3	$s_3, s_5, s_6$	$s_3, s_4$	$s_3$	$s_4$	$s_1, s_6$
4		$s_5, s_6$	$s_2$		

We first apply DR to this problem.

### Round 1.

#### Stage 1.

We illustrate the steps of the first stage below. In each step, tentatively held students are given in bold.

	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
Step 1:	<b><math>s_1, s_2, s_5</math></b>	<b><math>s_3, s_4, s_6</math></b>			
Step 2:	<b><math>s_1, s_2</math></b>	<b><math>s_3, s_4</math></b>	<b><math>s_5, s_6</math></b>		

When Stage 1 terminates school  $c_1$  holds  $s_1$  and  $s_2$ ,  $c_2$  holds  $s_3$  and  $s_4$ , and  $c_3$  holds  $s_5$  and  $s_6$ . We denote this outcome with  $\mu$ .

#### Stage 2.

Step 0: We first determine  $\rho_c(\mu)$  for all  $c \in C$ :  $\rho_{c_1}(\mu) = 2$ ,  $\rho_{c_2}(\mu) = 3$ ,  $\rho_{c_3}(\mu) = 1$ ,  $\rho_{c_4}(\mu) = \rho_{c_5}(\mu) = 5$ . Then, in addition to the students rejected in Stage 1,  $c_1$  rejects  $s_3$  and  $s_6$ ,  $c_2$  rejects  $s_5$  and  $c_3$  rejects  $s_4$ ,  $s_1$ ,  $s_3$  and  $s_2$ . For each school  $c$  we construct  $M_c, g_c^1$  as follows:

	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
$M$	$s_1, s_2$	$s_3, s_4$	$s_5, s_6$	$\emptyset$	$\emptyset$
$g^1$	0	0	0	0	0

Step 1: Given  $M$  and  $g^1$ , we have  $\tilde{D}_{c_1}^{c_2} = \{s_2\} = \tilde{D}^{c_2}$  and  $\tilde{D}_{c_3}^{c_4} = \{s_5\} = \tilde{D}^{c_4}$ . For all  $c \notin \{c_2, c_4\}$  we have  $\tilde{D}^c = \emptyset$ . Then, we calculate  $\hat{\rho}_{c_2} = 2$  and  $\hat{\rho}_{c_4} = 1$ . Since  $\hat{\rho}_{c_2} < \rho_{c_2}(\mu)$ ,  $c_2$  rejects  $s_3$  and  $s_4$ . Since  $\hat{\rho}_{c_4} < \rho_{c_4}(\mu)$ ,  $c_4$  rejects  $s_2, s_3$  and  $s_4$ . We go Round 2.

### Round 2.

Stage 1. We illustrate the steps of the first stage below. In each step, tentatively held students are given in bold.

	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
Step 1:	<b><math>s_1, s_2, s_4</math></b>		<b><math>s_5, s_6</math></b>		<b><math>s_3</math></b>
Step 2:	<b><math>s_4</math></b>	<b><math>s_1, s_2</math></b>	<b><math>s_5, s_6</math></b>		<b><math>s_3</math></b>
Step 3:	<b><math>s_4</math></b>	<b><math>s_1</math></b>	<b><math>s_5, s_6</math></b>		<b><math>s_2, s_3</math></b>

When Stage 1 terminates school  $c_1$  holds  $s_4$ ,  $c_2$  holds  $s_1$ ,  $c_3$  holds  $s_5$  and  $s_6$ , and  $c_5$  holds  $s_2$  and  $s_3$ . We denote this outcome with  $\mu$ .

**Stage 2.**

Step 0: We determine  $\rho_c(\mu)$  for all  $c \in C$ :  $\rho_{c_1}(\mu) = 1$ ,  $\rho_{c_2}(\mu) = 1$ ,  $\rho_{c_3}(\mu) = 1$ ,  $\rho_{c_4}(\mu) = 1$ , and  $\rho_{c_5}(\mu) = 2$ . Then,  $c_5$  rejects  $s_1$  and  $s_6$ . For each school  $c$  we construct  $M_c, g_c^1$  as follows:

	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
$M$	$s_4$	$s_1$	$s_5, s_6$	$\emptyset$	$s_3$
$g^1$	0	0	0	0	1

Step 1: Given  $M$  and  $g^1$ , we have  $\tilde{D}_{c_3}^{c_4} = \{s_5\} = \tilde{D}^{c_4}$ . For all  $c \neq c_4$  we have  $\tilde{D}^c = \emptyset$ . Then, we calculate  $\hat{\rho}_{c_4} = 1$ . Since  $\hat{\rho}_{c_4} = \rho_{c_4}(\mu)$  and  $g_{c_4}^2 = g_{c_4}^1$ , the algorithm terminates here.

Final outcome of DR is

$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
$s_4$	$s_1$	$s_5, s_6$	$\emptyset$	$s_2, s_3$

Next, we apply DP to the problem.

**Round 1.**

**Stage 1.** We illustrate the steps of the first stage below. In each step, tentatively held colleges are given in bold.

	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$
Step 1:	<b><math>c_2</math></b>	<b><math>c_5</math></b>		<b><math>c_1</math></b>		

Stage 1 terminates and the outcome is  $\mu(s_1) = c_2$ ,  $\mu(s_2) = c_5$ ,  $\mu(s_4) = c_1$ , and  $\mu(s_3) = \mu(s_5) = \mu(s_6) = \emptyset$ .

**Stage 2.**

We first construct  $D_s$  for each  $s \in S$ :  $D_{s_1} = \{c_1\}$ ,  $D_{s_2} = \{c_1, c_2, c_3\}$ ,  $D_{s_4} = \{c_2\}$ , and  $D_{s_3} = D_{s_5} = D_{s_6} = C$ . We update  $A_c$  for each  $c \in C$ :  $A_{c_1} = A_{c_2} = S$ ,  $A_{c_3} = \{s_2, s_3, s_5, s_6\}$ ,  $A_{c_4} = \{s_3, s_5, s_6\}$  and  $A_{c_5} = \{s_2, s_3, s_5, s_6\}$ .

Once we follow the steps of Stage 2, we can see that both  $s_5$  and  $s_6$  are guaranteed to be assigned to a school not worse than both  $c_3$  and  $c_4$ . Then, we remove them from  $A_{c_5}$  and set it to be  $A_{c_5} = \{s_2, s_3\}$ . We continue with Round 2 with updated  $A_c$  for all  $c \in C$ .

**Round 2.**

**Stage 1.** We illustrate the steps of the first stage below. In each step, tentatively held colleges are given in bold.

	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$
Step 1:	<b><math>c_2</math></b>	<b><math>c_5</math></b>	<b><math>c_5</math></b>	<b><math>c_1</math></b>		



Stage 1 terminates and the outcome is  $\mu(s_1) = c_2$ ,  $\mu(s_2) = \mu(s_3) = c_5$ ,  $\mu(s_4) = c_1$ , and  $\mu(s_5) = \mu(s_6) = \emptyset$ .

**Stage 2.**

We first construct  $D_s$  for each  $s \in S$ :  $D_{s_1} = \{c_1\}$ ,  $D_{s_2} = \{c_1, c_2, c_3\}$ ,  $D_{s_3} = \{c_2, c_3\}$ ,  $D_{s_4} = \{c_2\}$ , and  $D_{s_5} = D_{s_6} = C$ . We update  $A_c$  for each  $c \in C$ :  $A_{c_1} = \{s_1, s_2, s_4, s_5, s_6\}$ ,  $A_{c_2} = S$ ,  $A_{c_3} = \{s_2, s_3, s_5, s_6\}$ ,  $A_{c_4} = \{s_5, s_6\}$  and  $A_{c_5} = \{s_2, s_3\}$ .

Once we follow the steps of Stage 2, we can see that both  $s_5$  and  $s_6$  are guaranteed to be assigned to a school not worse than both  $c_3$  and  $c_4$ . We continue with Round 3 with updated  $A_c$  for all  $c \in C$ .

**Round 3.**

**Stage 1.**

We illustrate the steps of the first stage below. In each step, tentatively held colleges are given in bold.

$$\text{Step 1: } \begin{array}{cccccc} & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ \hline & \mathbf{c_2} & \mathbf{c_5} & \mathbf{c_5} & \mathbf{c_1} & & \end{array}$$

Stage 1 terminates and the outcome is  $\mu(s_1) = c_2$ ,  $\mu(s_2) = \mu(s_3) = c_5$ ,  $\mu(s_4) = c_1$ , and  $\mu(s_5) = \mu(s_6) = \emptyset$ .

**Stage 2.**

We first construct  $D_s$  for each  $s \in S$ :  $D_{s_1} = \{c_1\}$ ,  $D_{s_2} = \{c_1, c_2, c_3\}$ ,  $D_{s_3} = \{c_2, c_3\}$ ,  $D_{s_4} = \{c_2\}$ , and  $D_{s_5} = D_{s_6} = C$ . We do not update  $A_c$  for any  $c \in C$ .

Once we follow the steps of Stage 2, we can see that both  $s_5$  and  $s_6$  are guaranteed to be assigned to a school not worse than both  $c_3$  and  $c_4$ . Algorithm terminates since  $A_c$  stays the same for all  $c \in C$ .

## C Omitted Simulation Results

$\alpha$	$\beta$	$\gamma = 0.00$	$\gamma = 0.25$	$\gamma = 0.50$
		(1)	(2)	(3)
0.00	0.00	0	0	0
	0.25	0	0	0
	0.50	0	0	0
	0.75	0	0	0.03
	1.00	0	0	0.06
0.25	0.00	0	0.24	2.15
	0.25	0	0.53	2.69
	0.50	0.02	0.85	3.21
	0.75	0.01	0.81	3.13
	1.00	0	0.5	2.36
0.50	0.00	0.82	6.23	9.23
	0.25	1.71	8.61	10.42
	0.50	1.57	8.37	10.60
	0.75	0.68	5.87	9.55
	1.00	0.61	5.32	8.81
0.75	0.00	5.41	9.04	9.89
	0.25	6.89	9.63	10.49
	0.50	5.46	9.51	10.62
	0.75	4.26	8.85	10.17
	1.00	4.03	8.65	9.90
1.00	0.00	8.76	9.41	10.41
	0.25	8.86	9.92	10.44
	0.50	8.40	9.91	10.67
	0.75	7.86	9.38	10.34
	1.00	7.68	9.15	10.07

Table 3: Average number of schools with a unique stable assignment cutoff, 30 schools

$\alpha$	$\beta$	MQO (1)	L-MQO (2)	DA with quality-based tie-breaking (3)
0.00	0.00	57.926	57.926	12.410
	0.25	51.122	51.122	11.860
	0.50	45.384	45.384	11.736
	0.75	39.984	39.984	14.370
	1.00	36.878	36.878	15.042
0.25	0.00	52.692	52.692	29.796
	0.25	45.276	45.276	25.504
	0.50	36.202	36.202	18.918
	0.75	33.404	33.404	16.960
	1.00	33.194	33.148	18.240
0.50	0.00	45.538	45.188	24.816
	0.25	38.786	38.518	23.152
	0.50	32.826	32.546	19.608
	0.75	31.070	31.012	19.102
	1.00	30.936	30.890	19.066
0.75	0.00	45.318	45.256	27.758
	0.25	39.732	39.732	23.736
	0.50	34.844	34.782	22.592
	0.75	31.734	30.904	20.946
	1.00	31.634	30.734	20.860
1.00	0.00	44.856	44.328	32.248
	0.25	40.710	40.096	28.804
	0.50	36.002	35.412	25.664
	0.75	33.062	32.490	23.984
	1.00	30.900	30.014	22.410
Average gains		39.200	38.956	21.183

Table 4: Percentage gains compared to DA with random tie-breaking,  $\gamma = 0.00$

$\alpha$	$\beta$	MQO (1)	L-MQO (2)	DA with quality-based tie-breaking (3)
0.00	0.00	35.238	33.976	16.880
	0.25	32.504	31.060	14.956
	0.50	28.522	27.598	17.588
	0.75	24.882	23.722	14.750
	1.00	24.696	21.486	13.616
0.25	0.00	32.108	30.370	15.544
	0.25	29.002	26.422	14.390
	0.50	26.744	25.194	14.062
	0.75	23.928	20.978	11.300
	1.00	23.606	21.776	11.488
0.50	0.00	37.644	36.980	22.916
	0.25	33.362	33.144	21.376
	0.50	29.482	29.482	18.628
	0.75	25.174	25.174	15.724
	1.00	24.350	24.344	14.944
0.75	0.00	41.004	41.004	25.688
	0.25	37.382	36.956	23.976
	0.50	32.752	32.350	21.202
	0.75	28.030	28.030	18.088
	1.00	26.096	25.666	17.134
1.00	0.00	43.196	42.678	29.000
	0.25	39.286	39.286	27.132
	0.50	35.948	35.932	25.262
	0.75	31.874	31.874	22.236
	1.00	28.224	27.442	19.248
Average gain		31.001	30.117	18.685

Table 5: Percentage gains compared to DA with random tie-breaking,  $\gamma = 0.50$