Implementing quotas in university admissions: An experimental analysis

Sebastian Braun*, Nadja Dwenger†, Dorothea Kübler‡, Alexander Westkamp§

Abstract

This paper studies the implementation of quotas in matching markets. In a controlled laboratory environment, we compare the performance of two university admissions procedures that both initially reserve a significant fraction of seats at each university for a special subgroup of students. The first mechanism mimics the sequential procedure currently used by the central clearinghouse for university admissions in Germany. This procedure starts by allocating reserved seats among eligible students and then allocates all remaining seats among those who were not already assigned one of the reserved seats in the first part of the procedure. The second mechanism is based on a modified student-proposing deferred acceptance algorithm in which all seats are allocated simultaneously. In theory, the two mechanisms should lead to similar outcomes. Our experimental results, however, suggest that, relative to the sequential procedure, the simultaneous mechanism significantly improves the match outcome for the beneficiaries of reserved seats. The reason is that these students often accept suboptimal matches in the first part of the sequential procedure.

Keywords: College admissions; experiment; quotas; Gale-Shapley mechanism; Boston mechanism

JEL Classification: C78; C92; D78; I20

*Kiel Institute for the World Economy, email: sebastian.braun@ifw-kiel.de.

†Max Planck Institute for Tax Law and Public Finance. Email: nadja.dwenger@tax.mpg.de.

‡Corresponding author: WZB Berlin and Technical University Berlin. Email: dorothea.kuebler@wzb.eu.

§University of Maastricht, email: a.westkamp@maastrichtuniversity.nl.

1 Introduction

Affirmative action constraints play an important role in school choice and university admissions. For example, students from minority groups in the population often receive preferential treatment in school choice programs. This can be implemented by introducing *quotas* for minority students, that is, reserving some seats at each school for students belonging to this group.[1] Such minority seats must be opened up to majority students in case there is insufficient demand from minority students. Without such redistribution, minority quotas may make all students, including those belonging to the minority, worse off compared to a situation without affirmative action constraints (Kojima (2012)). But how can it be determined whether there is insufficient demand from minority students? Or, put differently, *how* are quota systems and affirmative action constraints best implemented in practice? To the best of our knowledge, this paper is the first experimental study of this problem. It complements a growing theoretical literature on matching mechanisms with flexible quota systems, including Echenique and Yenmez (2013), Ehlers et al. (2012), Hafalir et al. (2013), Kamada and Kojima (2010), Kominers and Sönmez (2013), and Westkamp (2013).

In our experiment, we implement two matching mechanisms with quotas in a controlled laboratory environment with complete information and compare their performance. The first mechanism is a simplified version of the centralized procedure currently used to allocate seats in medicine and related subjects at German universities.[2] The German mechanism is *sequential*: in the first step, a significant number of seats are allocated among applicants who have performed exceptionally well in high school (henceforth: *top-grade students*) on the basis of these students' preferences and exogenously given priorities using the well known *Boston mechanism*;[3] in the second step, all remaining seats are allocated among remaining

[1]Equivalently, one can limit the number of majority students allowed to attend each school, as in Abdulkadiroglu and Sönmez (2003). See Abdulkadiroglu and Sönmez (2003), Abdulkadiroglu et al. (2005), and Abdulkadiroglu et al. (2006) for several examples of affirmative action constraints in matching mechanisms.

[2]The admissions market for medical disciplines is characterized by a dramatic shortage of available seats: in the winter term 2010/2011, there were 56 000 applicants for 13 000 places.

[3]In the actual German procedure, the priority based part consists of two completely separate sub-steps: one in which up to 20% of available capacity is allocated among top-grade applicants, and another in which up to 20% is allocated among waiting-time applicants. Waiting-time applicants typically have almost no chance of being admitted in the two-sided part of the procedure, see Braun et al. (2010). In the experiment we therefore

applicants on the basis of students' preferences and criteria determined by the universities using the *deferred acceptance* algorithm of Gale and Shapley (1962).[4] Students are allowed to submit two lists - one for the first, and one for the second part of the procedure.

The second mechanism we implemented in our experiment uses a modified version of the student-proposing deferred acceptance algorithm that was recently proposed by Westkamp (2013). This mechanism allocates all seats *simultaneously* while respecting the quota structure of the German system. In each round of the algorithm, a "regular" student can claim a seat reserved for a top-grade student only if no such student has demanded that seat yet. This is radically different from the sequential mechanism, where a regular student can claim a seat initially reserved for top-grade students if it has not been allocated in a reduced problem in which *only* these seats are available to top-grade students. This difference has important consequences: the simultaneous mechanism is strategy-proof for top-grade students, while the sequential mechanism induces a difficult trade-off between securing a match in the first part, but possibly at a lower ranked university, and competing without priority for a seat in the regular quota. Our goal is to evaluate experimentally how students resolve such trade-offs and how this affects the performance of a matching mechanism with quotas.

The relevance of this analysis is not limited to the specific context of the German university admissions system, since similarly structured strategic problems can be expected to arise in most sequential allocation procedures: Dur and Kesten (2012) show that if a set of objects is allocated by means of a two-step procedure such that (i) the mechanism used for the first step is non-wasteful and individually rational, and (ii) the mechanism used for the second step is non-wasteful, then the two-step procedure as a whole may be neither strategyproof, nor efficient, nor non-wasteful. In particular, even if two strategyproof mechanisms are pasted together, we cannot expect to obtain a strategyproof sequential mechanism. Despite this deficiency, sequential mechanisms are used in practice to allocate goods in essentially static

take the group of students that receives preferred treatment to consist only of top-grade students.

[4]The actual German admissions procedure employs the university-proposing deferred acceptance algorithm in the second stage of the procedure. In the experiment, we chose to change this part of the procedure to the student-proposing variant in order to focus on the problems caused by the sequential assignment of seats. See Section 3.2 for details on the differences between the actual German admissions procedure and the version we implemented in our experiment.

allocation problems. Apart from the German university admissions mechanism, examples are the allocation of teaching jobs in Turkey, where tenured positions are always allocated before fixed term contractual positions, and school choice systems in the US, where seats at so called exam schools are often allocated before seats at regular schools.[5]

Our main result is that the simultaneous mechanism significantly improves the matching outcome of top-grade students relative to the current sequential mechanism. The sequential mechanism, designed to work in favor of top-grade students,[6] actually harms them. The reason is that top-grade students often fail to use the sequential system to their benefit. In particular, these participants often accept a relatively undesirable match in the first part of the procedure although a better match could have been obtained in the last part of the procedure. The differences between the two mechanisms persist also in later rounds of the experiment although they become smaller over time due to learning effects. Our findings suggest that the modified student-proposing deferred acceptance algorithm could be a valuable tool for redesigning university admissions in Germany – and seems also well suited to address matching problems with complex constraints in other contexts.[7]

Our experiments shed light on implementation issues in a matching environment. An interesting feature of our environment is that essentially all equilibria of the sequential mechanism yield the same outcome as the simultaneous mechanism under truth-telling. The crucial difference, however, is that top-grade students often have to misrepresent their preferences in order to reach this outcome. In our experiments we explained the basic strategic properties of the two mechanisms to participants (see Section 3.1 for details). We find that despite this advice, a significant share of participants fail to choose optimal manipulation strategies in the

[5] See Dur and Kesten (2012) for a theoretical analysis of the mechanisms used in these two examples.

[6] The quota for top-grade students was implemented in order to allow top-grade students to freely choose their university: "The ratio legis for this quota [the quota for top-grade students] is that the top-grade students can pick "their" university." (statement by the director of the German central clearing house, http://latnrw.de/lat-blog/wp-content/uploads/2012/11/doc20120827154027.pdf, accessed on February 2, 2013, translation by the authors).

[7] Westkamp (2013) shows that a generalization of the modified student-proposing deferred acceptance algorithm can handle much more complex constraints than those of the German system. For example, the mechanism can be used to implement the constraints that (1) x% of total capacity at a university/school should initially be reserved for a special group of applicants (e.g., siblings of existing students in case of school choice) and (2) any remaining capacity should be distributed equally among sexes. This is not possible with the type of affirmative action constraints considered in Abdulkadiroglu and Sönmez (2003).

sequential mechanism - even in situations where weakly dominant strategies exist. We thus provide experimental evidence on the question whether strategy-proofness is a desideratum (as it is easy to teach subjects the optimal strategy) or whether implementation in weakly dominant strategies, where the advice to subjects is more complicated, is sufficient to reach the desired outcome. A more general takeaway from our experiments is that the strategic complexity of the revelation game induced by a mechanism matters and that the high priority the market design literature assigns to strategy-proofness is justified.

Our paper is related to the growing experimental literature on matching mechanisms. Many of these papers share our basic experimental setup: all experimental subjects play the role of students and are asked to submit a rank-order list of their experimenter-assigned preferences to a centralized clearinghouse. In Chen and Sönmez (2006), experimental subjects play a one-shot game of incomplete information in which each participant is only informed about his own preferences, schools' capacities, and the matching mechanism. They find that from the perspective of students, the student-optimal mechanism outperforms both the Boston and the top-trading cycles mechanism.[8] Pais and Pintér (2008) compare the student-proposing deferred acceptance, Boston, and top-trading cycles mechanisms under various informational settings ranging from the zero information setting of Chen and Sönmez (2006) to the complete information setup that we employ in our experiment. For all three mechanisms the rate of truthful preference revelation is highest in the zero-information setting.[9] To the best of our knowledge, our paper is the first to study experimentally the performance of a two-stage matching mechanism combining the Boston and the student-optimal stable matching mechanism. The strategic properties of our sequential mechanism are very different from the so called proposal refusal mechanisms studied by Chen and Kesten (2011). In particular, truncations of preference lists can never be beneficial in a proposal refusal mechanism, whereas they are often optimal in the first stage of our sequential mechanism.

Our paper is also related to a small experimental literature on mechanism design and

[8]The top-trading cycles mechanism was introduced in Shapley and Scarf (1974) as a mechanism to find a core allocation in house exchange models. The mechanism was extended to school choice problems by Abdulkadiroglu and Sönmez (2003).

[9]Without being exhaustive, other experimental studies of matching mechanisms are Kagel and Roth (2000), Echenique et al. (2009), Calsamigla et al. (2010), Pais et al. (2011), as well as Guillen and Kesten (2012).

implementation. Implementation through Nash equilibrium can be successful in relatively simple environments (see Cabrales et al. (2003)), but most of the evidence on more complex mechanisms is rather negative (see Sefton and Yavas (1996) or Katok et al. (2001)). Moreover, it has been found that strategy-proofness is less valuable for implementation when the revelation game has Nash equilibria that are not dominant strategy equilibria (see Cason et al. (2006)). In contrast to these studies, our experiment focuses on matching algorithms, and the two mechanisms investigated should implement the same unique matching outcome, with one mechanism being strategy proof and the other requiring strategic manipulations.

The remainder of this paper is organized as follows: Section 2 introduces the two mechanisms we study and informally discusses their theoretical properties. In Section 3, we describe the experimental procedures before presenting the results from the experiments in Section 4. Section 5 summarizes the findings and concludes.

2 Mechanisms and theoretical background

In this section we introduce the two mechanisms used in our experiment and provide a mostly informal discussion of their theoretical properties. A full theoretical analysis can be found in the Online Appendix.

We consider a matching market consisting of a finite set of students S and a finite set of universities U. Each student $s \in S$ has a strict preference relation P_s over the available universities and an *average grade* $a(s) \in \mathbb{R}_+$. We assume throughout that there are no ties in average grades, so that $a(s) \neq a(s')$ for all $s \neq s'$, and follow the convention that *lower* average grades are associated with better performance in high school. Each university u has a fixed total number of available seats q_u and a strict ranking P_u of individual students.[10] What sets this market apart from previous experimental studies on matching is that each university u has to reserve $q_u^1 < q_u$ seats for students with "exceptional" average grades. More

[10]We will refer to P_u as u's preference (ranking of individual students) and not use the term priorities as is customary in the literature on school choice. The reason is that in the specific context of the German system, which motivates our experimental study, universities are free to choose how they evaluate and rank prospective students, within certain legal boundaries.

precisely, student s is a *top-grade student* if her average grade is one of the $q^1 = \sum_{u \in U} q_u^1$ best average grades among all students. An outcome of this market is a *matching* of students to universities. Since there are essentially two types of places at each university, those initially reserved for top-grade students and the rest, a matching has to specify not only to which university, if any, a student is matched, but also which "type" of place she receives. We assume throughout that students care only about the university they are assigned to and not about the particular type of place they get. The request that q_u^1 seats at university u should be reserved for top-grade students translates into the following constraint on matchings.

Constraint (A). Let μ be a matching, s be a top-grade student, and $v := \mu(s)$. There should not be a university u such that

(a) s strictly prefers u over v, and

(b) less than q_u^1 top-grade students were matched to u under μ.

The constraint guarantees that there is no top-grade student who would prefer a seat at a university which has not filled its top-grade quota over her actual matching. Note that Constraint (A) allows for a top-grade quota that is soft in the sense that seats initially reserved for top-grade students may be allocated among non top-grade students. This is in contrast to the "hard" quota systems that were first studied by Abdulkadiroglu and Sönmez (2003). A number recent theoretical studies of various matching problems focus on "soft" quotas, see Echenique and Yenmez (2013), Ehlers et al. (2012), Hafalir et al. (2013), Kamada and Kojima (2010), Kominers and Sönmez (2013), and Westkamp (2013).

Mechanisms. The main focus of our experiment is to compare two mechanisms implementing Constraint (A): (1) a stylized version of the current assignment procedure for seats in medical subjects at public universities in Germany, and (2) an alternative mechanism based on a modified version of the student-proposing deferred acceptance algorithm by Gale and Shapley (1962) that was recently proposed by Westkamp (2013). We now describe these mechanisms in detail.

Mechanism 1: Sequential Assignment (MSEQ)

To participate in the mechanism, each student submits two preference lists: one is used in the first, and the other in the second part of the procedure.

Part I: Assignment for top-grade students (Boston mechanism)

> In the first round, each top-grade student applies to her most preferred university (according to the ranking submitted for the first part). Each university admits students one at a time in increasing order of their average grades until either its top-grade quota is exhausted, or there are no more top-grade students who have ranked it first.
>
> In the kth round, each unassigned top-grade student applies to her kth most preferred university. Each university with remaining top-grade seats admits students one at a time in increasing order of their average grades until either its residual top-grade quota is exhausted, or there are no more top-grade students who have ranked it kth.

The first part ends when all unassigned top-grade students have applied to all universities they have declared acceptable for the first part. The second part allocates all remaining seats among all remaining students. Let μ^1 denote the matching produced in the first part of the procedure and $q_u^2 = q_u - |\mu^1(u)|$ denote the residual capacity of university u.

Part II: Two-sided part (student-proposing deferred acceptance algorithm)

> In the first round, each student applies to her most preferred acceptable university (with respect to the ranking submitted for Part II). Each university u temporarily admits students one at a time in order of their position in P_u until either its residual capacity is exhausted, or there are no more acceptable students, and rejects all other applicants.

> In the kth round, each unassigned student applies to her most preferred acceptable university among those that have not rejected her in previous rounds. Each university u temporarily admits students one at a time in order of their position in P_u until either its residual capacity is exhausted, or there are no more acceptable students, and rejects all other applicants.

The algorithm ends after a round in which no rejections are issued. Only at this point temporary assignments become final. □

In the following, we will refer to the above mechanism as MSEQ to emphasize its sequential structure. Note that if students submit preferences truthfully for both parts of the procedure, MSEQ satisfies constraint (A): in the first part, a top-grade student is rejected by a university u only if q_u^1 top-grade students have already been assigned to u. However, it is clearly not always beneficial for top-grade students to report preferences truthfully. On the one hand, a top-grade student may find it in her best interest to truncate, i.e., shorten, her true preference list for the first part of the procedure. A top-grade student who does not submit a truncated list always obtains a seat in the first part of the procedure and may thus forsake her chances of obtaining a better seat in the second part (where more capacity becomes available). On the other hand, a top-grade student may find it beneficial to overreport her preferences for some universities in the first part of the procedure if she cannot obtain a preferred assignment in the second part. The reasons are that (i) top-grade students lose their guaranteed priority over regular students if they are left unmatched in the first part of the procedure, and (ii) a top-grade student is guaranteed priority over other top-grade students with worse average grades only if she ranks a university first. Strategic misreporting may lead the outcome chosen by MSEQ to violate (A) with respect to students' true preferences. We study equilibria of the game induced by MSEQ below.

Mechanism 2: Simultaneous Allocation (MSIM)

To participate in the procedure, each student submits only one preference list.

In the first round, each student applies to her most preferred acceptable university. Out of the set of students applying to it, a university u

(1) temporarily admits top-grade students one at a time in increasing order of average grades until either its top-grade quota is exhausted, or there are no more top-grade students applying to it,

(2) temporarily admits remaining students one at a time in order of their position in P_u until either its residual capacity is exhausted, or there are no more acceptable students, and

(3) rejects all other students applying to it.

In the kth round, each student applies to her most preferred acceptable university among those that have not rejected her in any earlier round. Out of the set of students applying to it, a university u

(1) temporarily admits top-grade students one at a time in increasing order of average grades until either its top-grade quota is exhausted, or there are no more top-grade students applying to it,

(2) temporarily admits remaining students one at a time in order of their position in P_u until either its residual capacity is exhausted, or there are no more acceptable students, and

(3) rejects all other students applying to it.

The algorithm ends after a round in which no rejections are issued by universities.

$\square$

We will refer to this mechanism as MSIM to emphasize that it allocates all seats simultaneously. Note that this algorithm also implements Constraint (A) if students submit preferences truthfully: throughout the algorithm, a top-grade student is rejected by a university u only if at least q_u^1 other top-grade students with better average grades also apply to u. Before discussing the strategic properties of MSEQ and MSIM, we illustrate the two mechanisms by means of a simple example. The setting of the example corresponds to one of our experimental markets.

Example 1. *There are eight students $s_1, \ldots, s_8$ and four universities W, X, Y, Z. Students are indexed in increasing order of average grades, so that s_1 is the student with the best and s_8 the student with the worst average grade. Each university has a capacity of two seats. One seat at each university is reserved for top-grade students. Hence, students $s_1, \ldots, s_4$ are the top-grade students in this example.*

Students' and universities' preferences can be summarized by the following preference profiles:

$$
\begin{array}{rlr}
P_{s_i}: & W \succ X \succ Y \succ Z, & \forall i = 1, 2, ..., 8, \\
P_u: & s_1 \succ s_2 \succ s_3 \succ s_4 \succ s_5 \succ s_6 \succ s_7 \succ s_8, & \forall u = W, X, Y, Z.
\end{array}
$$

We now compute the outcomes of MSEQ and MSIM for this example under the assumption that all students (always) submit their preferences truthfully. The outcome of MSEQ is then given by

$$
\mu = \begin{pmatrix} & W & X & Y & Z \\ \textit{Top Grade} & s_1 & s_2 & s_3 & s_4 \\ \textit{Regular} & s_5 & s_6 & s_7 & s_8 \end{pmatrix},
$$

and the outcome of MSIM is given by

$$
\nu = \begin{pmatrix} & W & X & Y & Z \\ \textit{Top Grade} & s_1 & s_3 & \emptyset & \emptyset \\ \textit{Regular} & s_2 & s_4 & s_5, s_6 & s_7, s_8 \end{pmatrix},
$$

This notation means that, say, under μ, s_3 receives the place at Y that is initially reserved for top-grade students and s_7 receives the place that is open to all students. Under ν on the other hand, no top-grade student is assigned to Y and s_5 and s_6 both take regular seats. Note *that μ cannot be an equilibrium outcome of the revelation game induced by MSEQ. All top-grade students apart from s_1 could have obtained a strictly preferred assignment by ranking only their true first choice for the first and their full true preference ranking for the second part of the procedure (conditional on knowing universities' preferences, these strategies are actually weakly dominant in this example). For these reports the outcome of MSEQ coincides with the outcome of MSIM under truth-telling.*

Strategic properties and equilibrium outcomes. Next, we describe the strategic properties of the two mechanisms introduced above. Remember that in MSEQ, students submit two preference lists, while they submit only one list in MSIM. The next result describes the basic incentive properties of the two mechanisms.

Theorem 1. *(i) For MSEQ, it is a weakly dominant strategy for each student to submit her preferences truthfully for the second part of the mechanism.*

(ii) For MSIM, it is a weakly dominant strategy for each student to submit her preferences truthfully.

The first part of this result is due to Dubins and Freedman (1981) and Roth (1982). The second part follows from Theorem 2 in Westkamp (2013). This result shows that under certain conditions, which are satisfied in our application, matching mechanisms with flexible quotas satisfy Hatfield and Milgrom (2005)'s sufficient conditions for strategy-proofness.[11] The strategy-proofness of MSIM requires that final acceptances are deferred until the end of the procedure <u>and</u> that all places are allocated simultaneously. As shown by Dur and Kesten (2012), any sequential procedure that is comprised of reasonable mechanisms would violate strategy-proofness.

Of course, MSIM is not the only strategy-proof mechanism that satisfies Constraint (A). What sets MSIM apart from other such mechanisms are two things: First, it produces an

[11] The second part of Theorem 1 also follows from Proposition 1 in Hafalir et al. (2013) and Theorem 4 in Kominers and Sönmez (2013). Both of these papers were written subsequently to the first version of Westkamp (2013). It is worth mentioning that Theorem 2 in Westkamp (2013) applies under much more general conditions than those that we consider in our analysis. Proposition 1 in Hafalir et al. (2013) is a special case of, and Theorem 4 in Kominers and Sönmez (2013) is neither more nor less general than Theorem 2 in Westkamp (2013).

outcome that is “stable” in the sense that whenever a student s strictly prefers a university U over the university she was assigned to by MSIM, then (1) at least $q_U - q_U^1$ places have been assigned to students who rank higher than S according to P_U, and (2) if s is a top-grade student, q_U^1 places at U have been assigned to top-grade students with better average grades than s. Second, among all “stable” matchings, MSIM always produces the one that is best for all students (see the Online Appendix, part D, for a detailed discussion of the concept of stability we use). We view stability as an important property since in the German system, universities play an active role in the evaluation of students. Even in case that universities/schools have a completely passive role and allocate their places on the basis of priorities set by law, stability plays an important role in theory and applications (see Pathak and Sönmez (2013) and Sönmez and Unver (2011) for recent surveys). Hence, our analysis remains relevant for this case as well. Given that MSIM strikes an optimal compromise between stability and efficiency in the sense of producing <u>the</u> student optimal “stable” matching,[12] we believe that it is the right benchmark mechanism for our study.

By Theorem 1, students should always submit preferences truthfully for the second part of MSEQ. However, for the first part of MSEQ truth-telling is rarely an equilibrium. As shown by Pathak and Sönmez (2013) in a recent study, truth-telling is an equilibrium of the Boston mechanism if and only if (true) preferences are so dispersed that every student can be assigned her first choice.[13] The next result describes equilibrium outcomes of MSEQ.

Theorem 2. *Consider any complete information Nash-equilibrium of MSEQ such that*

(1) all students submit preferences truthfully for the second part of the mechanism, and

(2) no top-grade student who is matched to a university u in the second part of MSEQ could have been matched to u in the first part of MSEQ by unilaterally deviating to a strategy which ranks u as her top choice for that part.

Then this equilibrium outcome coincides with the outcome of MSIM under truth-telling.

[12]The results of Ergin (2002) imply that there may not exist an efficient and “stable” outcome in our setting.

[13]Ergin and Sönmez (2006) have shown that the set of (complete information Nash-)equilibrium outcomes of the revelation game induced by the Boston mechanism coincides with the set of matchings that are stable with respect to participants’ true preferences.

The proof of this result can be found in the Online Appendix, part A. The theory thus suggests that we should expect outcomes of MSEQ and MSIM to be similar to each other, provided applicants have sufficiently detailed information about the admissions environment.[14]

3 Experimental Design

We implemented the sequential assignment mechanism (treatment MSEQ) as well as the simultaneous assignment mechanism (treatment MSIM) in a laboratory experiment. In the experiment, eight students ($s_1, \ldots, s_8$) applied to four universities (W, X, Y, Z) with two seats each. One seat per university was reserved for top-grade students, the other seat was allocated according to the preferences of the university (regular quota). Applicants were ordered by their average grades so that student s_1 was the best student, s_2 the second best etc. Thus, students s_1, s_2, s_3, and s_4 were eligible under the quota for top-grade students in the experiment as half of the eight seats were reserved for this group.

Preferences, roles, and information. Participants in the experiment always took the role of students. Each student was assigned a strict ranking of available universities. Students received a payoff of EUR 22 when matched to their first choice, EUR 16 when matched to their second, EUR 10 when matched to their third, and EUR 4 when matched to their fourth choice. The universities were played by the computer, i.e., their strict preferences were exogenously given, and the computer acted truthfully according to these preferences. All relevant information was common knowledge among the students. In particular, participants were informed about the preferences of all other applicants and of universities.

Markets. In order to understand how the functioning of the two mechanisms depends

[14] The reason for focussing on equilibria satisfying constraint (2) above is that in the actual implementation of the German admissions procedure it can be expected that applicants have a strict preference for being assigned "early". The reason is that the second part of the procedure takes part about one month after the first, so that being assigned in the latter gives applicants more time to search for an apartment, prepare to move, etc. In our experiment, it makes no difference for applicants whether they are assigned in the first or the second part of the procedure, so that we cannot just impose restriction (2). However, in all of our experimental markets, all Nash-equilibria of MSEQ in which no applicant employs a weakly dominated strategy yield the same *matching* of applicants as MSIM under truth-telling (although the distribution across the different types of seats may be different). See Appendix A.1 for details.

on the preferences of students and universities, we designed four different markets. Table 1 provides an overview of the market characteristics. A detailed description of the markets and an analysis of the equilibrium outcomes can be found in Appendix A.1. In the four markets, we vary the degree of correlation of university and student preferences. The first market features perfectly correlated preferences of students and universities ('fully aligned' preferences). This market has already been analyzed in some detail in Example 1. Market 2 retains perfectly correlated student preferences ('student aligned') but reduces the correlation among university preferences. In particular, only two out of the four universities share the same preferences over students, while the other two universities have slightly different preferences. We refer to this preference pattern of the universities as 'split aligned'. Market 3 has perfectly aligned university preferences but split aligned student preferences ('university aligned'). Finally, market 4 features split aligned preferences on both sides of the market ('split aligned').[15] The specific market setting determines which top-grade students have an incentive to strategically misrepresent their preferences in order to improve their matching (see the second to last column of Table 1). Some but not all of these students have a weakly dominant strategy at hand (last column).

Table 1: Overview of market characteristics

	Preferences of		Students with	Of which with
	students	universities	incentive to misrepresent[1]	WD strategies[2]
Market 1: Fully aligned	aligned	aligned	s_2, s_3, s_4	s_2, s_3, s_4
Market 2: Student aligned	aligned	split aligned	s_2, s_3, s_4	
Market 3: University aligned	split aligned	aligned	s_2, s_4	s_2, s_4
Market 4: Split aligned	split aligned	split aligned	s_3, s_4	

Notes: [1] Top-grade students who can improve their payoffs by misrepresenting their true preferences in the first stage of MSEQ. [2] WD = weakly dominant. A detailed description of the four markets is provided in Appendix A.1.

Implementation, payoffs and observations. The experiment was conducted with students at the experimental lab of Technical University Berlin and on computers using z-

[15]In practice, preferences of universities are highly positively (and in part even perfectly) correlated as all universities have to use the final average grade from school as the main criterion (due to legal constraints) and some universities even base their ranking of applicants solely on the final grade. Because some universities use interviews, tests etc. as additional admission criteria, the 'student aligned' and 'split aligned' markets are also relevant benchmarks. From the students' perspective, some universities are regularly over-demanded. That is to say, a large number of applicants want to study at a famous university or in an attractive university town. We therefore consider both, markets with perfectly and split aligned student preferences, as important reference cases.

Tree (Fischbacher 2007). For each treatment MSIM and MSEQ, independent sessions were carried out. In the beginning of the experiment, printed instructions were given to participants (see Online Appendix, part E). Participants were informed that the experiment's aim was to analyze decision making processes in university admission procedures, that they took on the role of student applicants, and that their payoff depended on their own decisions and the decisions of the other participants. The instructions, which were identical for all participants of a session, explained in detail the experimental setting and the assignment mechanism. Questions were answered privately and all individuals answered a computerized quiz to make sure that everybody understood the main features of the particular mechanism.

Subjects played the four markets in changing roles, i.e., in each round they were randomly chosen to take on the role of one of the eight students. Each subject participated in a total number of 12 rounds and played each market three times. The ordering of markets was determined randomly, but each market had to occur exactly once in rounds 1-4, once in rounds 5-8, and once in rounds 9-12 (random draw without replacement).[16] Although real application decisions are typically made only once, we decided against running one-shot experiments with one separate treatment for each market. Such a design would not be able to account for real-life opportunities to learn about the mechanism, such as discussions with peers, counsellors, and parents. Moreover, our aim is to study how well students perform in the strategically more complex sequential mechanism. By not allowing for learning, we would have biased the design in favor of the simpler simultaneous mechanism.

In treatment MSIM, subjects had to submit one rank order list of universities in each round. In treatment MSEQ, subjects had to submit two lists, one for each part, in each round. Once all decisions were made, the matching was determined by the computer according to the algorithms described in Section 2. After each round, subjects were informed about their own

[16]The market ordering was only drawn once per mechanism and remained unchanged in all subsequent sessions. We ran a control experiment for MSIM (MSIM-control) to rule out that the different market orderings in the two treatments affect our results. In this control experiment for MSIM, participants played the exact same market ordering that was drawn for MSEQ in the main experiment. We ran 4 sessions, each hosting three groups of eight participants, yielding 4 x 3 x 8 = 96 participants overall (12 independent observations). To test for the robustness of our results, we compare the outcomes in MSIM-control to those obtained in MSEQ and find that the results obtained in the control experiment are essentially identical to those in our main experiment (seet part H of the Online Appendix for details).

matching and that of their co-players. At the end of the experiment, one round was chosen at random to determine the payoffs of the participants. The average payment for the matching was EUR 14.25 per participant (with a standard deviation of EUR 7.19). In addition, students received a fixed show-up fee of EUR 5 and a fixed bonus of EUR 5 for correctly answering the computerized quiz which queried the main principles of the mechanism.

For each of the two main treatments, 10 sessions were carried out, and each session hosted three groups of eight participants. The number of strictly independent observations is therefore 30 (ten sessions per treatment times three groups per session). Overall, 240 ($= 10 \times 3 \times 8$) subjects participated in each of the two main treatments of the experiment (or 480 in total). Each subject participated in only one session and played 12 rounds. Due to a computer problem, 24 observations were not recorded. This leaves us with a total of $(480 \times 12) - 24 = 5736$ observations.

3.1 Strategic coaching

We provided participants with as much information as possible about optimal strategies. Participants in treatment MSEQ were informed that truth-telling would not always be optimal for them in the first stage of the mechanism. They were told that, depending on the circumstances, it could be optimal for them to either truncate their submitted preference list, or to over-report their preference for a university by ranking it higher than another university which would have yielded a higher monetary payoff. The instructions illustrated these properties with the help of an exemplary market (which did not correspond to one of our experimental markets). Moreover, participants were informed that truth-telling would always be optimal for them in MSIM and in the second stage of MSEQ. Together with the instructions, we provided participants with an explanation of this incentive property on a separate piece of paper. This information was given in the form of an informal proof to provide the information as objectively as possible. We chose this neutral framing because of the concern that the specific phrasing of the advice might be driving the results (e.g. through experimenter demand effects), instead of the mechanism itself.

Giving explicit advice to experimental subjects might seem unusual.[17] But "strategic coaching" is an integral part of many real-life centralized matching mechanisms. For instance, the German central clearinghouse advises applicants on its homepage to *"[t]hink twice about whether you are willing to accept a university below your first preference rank* [in the first part of the procedure]. [...] *If you want to maintain your chance of being admitted in the second part of the procedure, you should only list your favorite universities* [...]. [...] *However, keep in mind that there is also a possibility of rejection* [in the second part of the procedure], *since there can be no guarantee for acceptance."*[18] Similarly, real-life implementations of strategy-proof mechanisms are usually accompanied by the explicit advice that truth-telling will always be optimal. For example, Boston public schools (BPS), which have recently adopted the strategy-proof student-proposing deferred acceptance algorithm, advise students that *"the new [assignment] formula enables parents to list their true choices of schools, in true order of preference, without having to 'strategize' about the rank order."*[19]

We also ran a small control experiment without advice.[20] The control experiment was identical to the main experiment except that we did not provide participants with advice about the strategic issues (see the Online Appendix, part E, for details). The comparison of the main and the control experiment allows us to test the effectiveness of giving advice in strategy-proof and non strategy-proof matching mechanisms.

Another potential control would have been to run a separate treatment for each experimental market. This would have allowed us to give participants in MSEQ treatments explicit advice on *how* to manipulate their preferences for the given market. However, such advice would necessarily have to depend on extremely detailed information about market characteristics that is unlikely to be available in practice. In contrast, the type of advice that we give to participants in our experiment relies only on structural features of the two mechanisms we

[17]Note, however, that it has been employed in previous matching experiments, see e.g. the study of unraveling by Kagel and Roth (2000).

[18]See http://hochschulstart.de/index.php?id=683 (accessed on September 6, 2013, translation by the authors).

[19]See http://www.bostonpublicschools.org/frequently-asked-questions, accessed on September 6, 2013).

[20]We employed two sessions per mechanism in the control experiment. As in the main experiment, each session hosted three groups of eight participants. Therefore, 48 (= $2 \times 3 \times 8$) subjects per mechanism participated in the control experiment (or 96 in total).

study.

3.2 Differences to the German admissions procedure

In this section, we discuss how our experimental setting differs from the actual assignment procedure for German universities and argue that these differences do not bias the performance of top-grade students in favor of MSIM.

First, our experimental markets are of much smaller size than the real markets we are interested in. Yet, also in larger markets will the number of top-grade students equal the number of seats in the top-grade quota. Given the high degree of correlation in students' preferences that is characteristic of the German market (Braun et al. 2010),[21] there is no reason to expect less competition in large markets. Rather, students whose average grades were good enough to make them eligible for the top-grade quota, but not good enough to guarantee an assignment to one of their top choices in the first part of the procedure, can be expected to face similar trade-offs in real markets as subjects faced in our controlled laboratory environment.

Second, we implemented a setting of complete information among students. In practice, the German central clearinghouse does indeed provide detailed information on past runs of the assignment procedure and grade distributions are very stable over time. Assuming stationary distributions of students' and universities' preferences, this allows applicants to form a good estimate of other applicants' preferences. Hence, it would not have been realistic to provide participants with no information at all about the environment. However, with only partial information about the environment finding an optimal application strategy in MSEQ becomes very difficult. Since information about others' preferences plays no role in MSIM, due to its strategy-proofness, a setting with partial information could have biased the results in favor of MSIM. Complete information makes it relatively straightforward for top-grade students to identify profitable manipulation strategies for MSEQ in our experimental markets.

Third, in the experiment MSEQ is sequential only in the sense that seats in the top-grade

[21] Individual preferences are also likely to be correlated in other environments, for example when preferences are determined by institutional quality and proximity, see Chen and Sönmez (2006).

quota are allocated before seats in the regular quota are assigned. In the actual assignment procedure, assignments in the regular quota are determined about one month after seats in the top-grade quota are assigned.[22] Thus, in reality applicants can be expected to have a strict preference for being matched as early as possible (in the sense of Definition 1), given that an earlier match means more time to search for an apartment, prepare to move, etc. This difference to the real world setting should again work to the benefit of MSEQ in the experiment, since it makes it less risky for a top-grade student to wait for the second part of the procedure in the lab (and may thus make it more likely for students to submit a truncated preference list).

Finally, in the experiment the student-proposing deferred acceptance algorithm is applied in the second part of MSEQ. In the actual German assignment procedure, in contrast, the university-proposing deferred acceptance algorithm is used. We chose to apply the student-proposing deferred acceptance algorithm in the second stage of MSEQ to make the two treatments, MSIM and MSEQ, as symmetric as possible in their allocation of regular seats. Furthermore, our focus is on the incentives to truncate the rank-order list in the first stage of MSEQ, not on the strategic incentives in the university-proposing deferred acceptance algorithm.

4 Results

This section contains our experimental results. First, we describe the application strategies of participants. Second, we compare the pay-offs that students realize under each mechanism and across the different markets. We then turn to the analysis of learning effects. Finally, we compare the two mechanisms from the point of view of the universities.

4.1 Application strategies

Table 2 presents the application strategies chosen by the participants in our experiment with advice (upper panel), and contrasts them with the strategies played in the control experiment

[22] The reason for this is to give universities enough time to evaluate those students who remain unassigned after the first part of the procedure.

without advice (lower panel). We distinguish between four classes of application strategies: (1) reporting a preference ordering that corresponds exactly to the ranking induced by monetary payoffs (we will refer to this as *truthful preference revelation*); (2) submitting an ordering that contains less than four universities (we will refer to this as *truncating*); (3) ranking a university first that is not the true first choice (we will refer to this as *over-reporting*); and (4) other strategies.

Consider first the participants who receive advice regarding optimal strategies (upper panel of Table 2). In MSIM, 81.02% of all reports are truthful. For the second part of MSEQ, where truthful revelation is also weakly dominant, this share drops to 75.35%. The share of applicants playing truthfully is much lower for the first stage of MSEQ, where top-grade students can often benefit from misrepresenting their preferences. Here, only 13.68% of top-grade students' reports are entirely truthful. This share is significantly lower than in the Boston mechanism with full information studied by Pais and Pintér (2008), where 46.7% of all applicants reveal their preferences truthfully. The lower share in our experiment could be attributed (1) to our advice about the potential value of misrepresenting the preferences, and (2) to the fact that applicants in MSEQ have a second chance of obtaining a place *after* the termination of the Boston mechanism.

Table 2 also reveals significant differences between the mechanisms in the way applicants misrepresent their preferences. Many top-grade students (52.08%) truncate their preference list in the first stage of MSEQ, presumably because they are afraid of being matched "too early" to a lower ranked university. A significant fraction of top-grade students also over-reports (10.49%) or over-reports and truncates (22.08%) their preference list, presumably to increase their chances of being matched to a relatively high ranked university already in the first part of the procedure. However, only about a quarter of top-grade students submits a complete preference list, which would guarantee a match in the first part. In contrast, for MSIM and the second stage of MSEQ, more than 90% of all applicants submit a full preference list containing four universities.[23]

[23] Comparing the second stage of MSEQ to MSIM we find that the share of truncated preference lists is considerably larger in the second stage of MSEQ (7.78%=5.63%+2.15%) than in MSIM (2.98%=2.38%+0.60%).

If we do not provide participants with explicit advice, the share of participants who report their preferences truthfully when they should do so is considerably smaller. The share of truthful reports drops from 81.02% with advice to 57.99% without advice in MSIM and from 75.35% to 49.83% in the second stage of MSEQ. Interestingly, the rate of truthful preference revelation is significantly higher in MSIM than in the second stage of MSEQ both with and without advice. This suggests that when subjects are exposed to a combination of manipulable and non-manipulable mechanisms, they are less likely to report their preferences truthfully in the non-manipulable part of the mechanism. Our results also show that the share of top-grade students who submit truthful reports in the first stage of MSEQ, where manipulation can be beneficial, increases significantly if we do not advise them on their optimal strategies. Even without advice, however, the large majority of top grade students (78.12%) does not play truthfully in the first stage of MSEQ.

Overall, the results suggest that strategic coaching has a significant effect on the choices of participants. In what follows, we will concentrate on the results of the experiment with advice, as we want to compare the performance of matching mechanisms under real-life conditions where advice typically plays an important role. The Online Appendix (part G) reports the corresponding results for the experiment without advice. Our main qualitative findings from the experiment with advice carry over to the experiment without advice.[24]

Being exposed in the first stage of MSEQ to a mechanism in which truncation may pay off thus seems to encourage some applicants to also truncate their preference list in the second stage (where truncation does not pay off).

[24]In particular, we find that the share of rounds resulting in the equilibrium matching is significantly lower in MSEQ than in MSIM. Also, all top-grade students are better off in terms of preference ranks obtained in MSIM than in MSEQ (significant for students 2 and 3). The results of the experiment with and without advice differ only in two respects: First, in the experiment without advice, we do not find a statistically significant difference between MSIM and MSEQ in the average aggregate performance measure. While such a difference does exist in the experiment with advice, it is small (see Result 1 in Section 4.2 and the associated discussion). Second, the difference in the preference ranks obtained by top-grade students in MSIM and MSEQ does not change monotonically in the course of the experiment without advice. In fact, differences first increase and then decrease again. This is in contrast to the experiment with advice, in which differences between MSIM and MSEQ become continuously smaller over time.

Table 2: Proportion of truthful preference revelation and misrepresentation, by mechanism

Mechanism	Truth-telling		Misrepresentation of preferences			
	All pref.	1st pref	Truncation[2]	Over-reporting	Over-reporting & truncation	Other
			Main experiment with advice			
MSIM	81.02%	87.82%	2.38%	11.55%	0.60%	4.45%
MSEQ, first stage[1]	13.68%	60.83%	52.08%	10.49%	22.08%	1.67%
MSEQ, second stage	75.35%	85.42%	5.63%	11.67%	2.15%	5.21%
			Control experiment without advice			
MSIM	57.99%	66.49%	3.99%	31.94%	1.56%	4.51%
MSEQ, first stage[1]	21.88%	66.67%	47.92%	15.28%	11.81%	3.13%
MSEQ, second stage	49.83%	66.49%	12.50%	27.08%	3.13%	7.47%

Notes: [1]In the first stage of MSEQ, we only consider the choices of students s_1 to s_4 who are eligible under the quota for top-grade students. [2]Entries refer to individuals who are exclusively truncating (over-reporting). Individuals who do both are considered in column 6.

4.2 Performance of Mechanisms: Student Perspective

We now analyze the performance of the two mechanisms with respect to the students' preferences. To do so, we compare the actual performance of the mechanisms relative to each other as well as relative to the theoretical benchmark of MSIM under truth-telling (we will refer to the latter outcome as "the equilibrium outcome" in the following).[25]

Equilibrium outcomes and aggregate performance. Table 3 reports on how often the theoretical benchmark is reached as a fraction of the total number of rounds. It shows that the equilibrium matching is much more often realized in MSIM than in MSEQ. Across all four markets, the equilibrium matching is reached in 77.31% and 22.78% of all rounds in MSIM and MSEQ, respectively. The difference is highly statistically significant both overall and for each of the four markets individually. While the equilibrium outcomes of MSIM and MSEQ coincide in theory, the equilibrium outcome is thus reached much less frequently in treatment MSEQ than in treatment MSIM.[26]

[25]In all four experimental markets, all equilibria of the game induced by MSEQ yield the outcome of MSIM under truth-telling if we restrict attention to strategies that are not weakly dominated (see Appendix A.1).

[26]We can also study how far the actual matching outcome deviates from the equilibrium matching outcome by looking at the average share of students who realize their equilibrium outcome by mechanism and market (irrespective of whether the equilibrium matching is reached for all eight students in a round). The results (reported in the Online Appendix, part F) indicate that a considerable share of students receive their equilibrium outcome in MSEQ although the equilibrium outcome is only rarely reached for all eight students at once. However, the share of students who receive their equilibrium outcome under MSIM is significantly higher than

Table 3: Share of rounds in which the realized matching coincides with the equilibrium matching, by mechanism and market

	MSIM	MSEQ	MSIM − MSEQ
Market 1: Fully aligned	0.9111	0.2778	0.6333***
	(0.2862)	(0.4504)	[0.0563]
Market 2: Student aligned	0.7701	0.3333	0.4368***
	(0.4232)	(0.4740)	[0.0676]
Market 3: University aligned	0.8333	0.1667	0.6667***
	(0.3748)	(0.3748)	[0.0559]
Market 4: Split aligned	0.5778	0.1333	0.4444***
	(0.4967)	(0.3418)	[0.0636]
Markets 1–4	0.7731	0.2278	0.5453***
	(0.4194)	(0.4200)	[0.0313]

Notes: *** denotes statistical significance at the 1%-level. Entries in columns two and three are the cell-specific mean values of an 'equilibrium dummy' that takes on a value of one if the realized matching in a round coincides with the equilibrium matching. Entries in column four are the (market-specific) coefficient estimates of an indicator variable for MSIM from an OLS regression of the equilibrium dummy on a constant and the MSIM indicator itself. Standard deviations are in round and standard errors in squared brackets. The unit of observation is a round.

Next, we analyze how deviations from the equilibrium matching are reflected in the aggregate performance of the two mechanisms over all eight students. As a measure of the aggregate performance, we use the average difference between equilibrium assignments and assignments realized in our experiment (where differences are measured in rank points).[27]

We group observations according to the rounds in which they were observed in the experiment. For mechanism $M \in \{\text{MSEQ}, \text{MSIM}\}$, let y_{ij}^M be the preference rank that the participant in the role of student type $i \in \{1, \ldots, 8\}$ was assigned to in the jth round of the experiment.[28] Let $k^M(j) \in \{1, \ldots, 4\}$ denote the market that was played in the jth round of mechanism M. Finally, let $y_{ik^M(j)}^e$ denote the preference rank that student type i obtains in equilibrium of market $k^M(j)$. The aggregate performance measure of M in round j is then defined as

the corresponding share under MSEQ.

[27]This aggregate performance measure puts equal weight on each student type. Aggregate performance measures that use different weighting schemes can be calculated from the type-specific individual performance measures reported in Table 5.

[28]For each mechanism we had 30 groups of participants. Each group played for 12 rounds, so that $j \in \{1, \ldots, 360\}$.

$$M_j^{\text{agg}} = \frac{\sum_{i=1}^{8}(y^e_{ik^M(j)} - y^M_{ij})}{8}. \tag{1}$$

This performance measure takes on positive values if the realized preference ranks under mechanism M are on average lower than those in the outcome of MSIM under truth-telling, i.e., if M outperforms the theoretical equilibrium in our experiment. Negative values, in contrast, mean that M underperforms relative to the theoretical equilibrium outcome. In the following, we study the mean of M_j^{agg}, denoted by $\overline{M}^{\text{agg}}$, across all experimental rounds. If, say, $\overline{M}^{\text{agg}} = -0.2$, the realized matching is on average 0.2 rank points higher than the equilibrium matching. This means that, on average, in every fifth observation on mechanism M a student then obtains an assignment that is one preference rank higher/worse than in equilibrium.

Table 4 shows how the aggregate performance measure differs across the two mechanisms and across the different markets. Differences between realized and theoretical outcomes are statistically significant but small (i.e., $\overline{M}^{\text{agg}}$ is close to zero) in MSIM, and realized rank points are slightly higher in MSIM than in MSEQ. Across all markets, the aggregate performance measure is 0.0411 rank points higher in MSIM than in MSEQ and the difference is statistically significant at the one percent level.

The finding of relatively small differences in *aggregate* performance between the two mechanism is not surprising as the preferences of students are strongly and in two markets even perfectly correlated. Gains for one applicant thus often come at the expense of another applicant.[29] Table 4 further shows that MSEQ outperforms MSIM in market 3 in particular. In this market, in which only the preferences of universities but not those of students are perfectly correlated, the performance measure is 0.1319 rank points higher in MSIM than in MSEQ.

We summarize the above findings in:

Result 1: Equilibrium outcomes and aggregate performance. *The equilibrium*

[29]With perfectly correlated preferences, differences in average rank points between the two mechanisms can only occur if an applicant is not matched in one of the two mechanisms.

Table 4: Aggregate performance measure, mean value by mechanism and market

	MSIM	MSEQ	MSIM − MSEQ
Market 1: Fully aligned	0.0000	-0.0111	0.0111***
	(0.0000)	(0.0358)	[0.0038]
Market 2: Student aligned	-0.0029	-0.0125	0.0096*
	(0.0188)	(0.0421)	[0.0049]
Market 3: University aligned	-0.0319	-0.1639	0.1319***
	(0.0984)	(0.1662)	[0.0204]
Market 4: Split aligned	-0.0639	-0.0764	0.0125
	(0.1203)	(0.1460)	[0.0199]
Markets 1–4	-0.0249	-0.0660	0.0411***
	(0.0824)	(0.1296)	[0.0081]

Notes: ***,* denotes statistical significance at the 1%- and 10%-level, respectively. Entries in columns two and three are the cell-specific mean values (over all rounds) of the aggregate performance measure in MSIM and MSEQ, respectively. Entries in column four are the (market-specific) coefficient estimates of an indicator variable for MSIM from an OLS regression of the aggregate performance measure on a constant and the MSIM indicator itself. The aggregate performance measure is defined in equation (1). Standard deviations are in round and standard errors in squared brackets. The unit of observation is a round.

matching is significantly more often realized in MSIM (77.31% of all rounds) than in MSEQ (22.78%). Realized ranks are, on average, close to the equilibrium outcomes in MSIM. Realized rank points are statistically significantly higher in MSIM than in MSEQ. Overall, the average difference is 0.0411 rank points per student.

Individual performance. The key motivation for the introduction of a quota for top-grade students was to provide these students with relative freedom to choose their most preferred universities (see Footnote 6 above). Hence, a key performance criterion is how well top-grade students do in terms of realized assignments. To investigate this, we now turn to differences between the two mechanisms in the matching outcomes for individual student types. Analogous to the aggregate performance measure, we define the performance of $M \in \{\text{MSEQ},\text{MSIM}\}$ for student type $i \in \{1, \ldots, 8\}$ in round j as

$$M_{ij} = y^e_{ik^M(j)} - y^M_{ij}\,. \tag{2}$$

As for the aggregate performance measure, this performance measure takes on positive (negative) values if the outcome for student type i under M in our experiment is on average better (worse) than in the theoretical equilibrium. Notice that the individual performance measure is necessarily non-positive (non-negative) for students who receive their best (worst) option in equilibrium.[30] As in the case of aggregate performance, we will concentrate on the mean of the individual performance measure, denoted by $\overline{M}_i$, across all experimental rounds. If, say, $\overline{M}_i = -0.2$, in every fifth observation on student type i under mechanism M, i obtains an assignment that is one preference rank higher/worse than in equilibrium.

Table 5 provides the average individual performance measure by student type and mechanism. It documents that the relatively small differences between the two mechanisms that we observed at the aggregate level hide considerable differences at the individual level. In MSIM, matching outcomes are generally very close to equilibrium outcomes (i.e., the performance measure is close to zero). In MSEQ, in contrast, matching outcomes for most student types differ considerably from equilibrium outcomes.

As shown in the last column of Table 5, MSIM generally benefits top-grade students and harms regular students relative to MSEQ. For students s_2 and s_3, for instance, the actual matching outcomes are 0.3278 and 0.3833 rank points below the equilibrium in MSEQ but only 0.0701 and 0.0644 rank points below the equilibrium in MSIM. Both students have to manipulate their rank-order list to obtain their equilibrium profits in MSEQ. In particular, student s_2 must manipulate her list in markets 1 to 3, student s_3 in markets 1, 3, and 4 (see Table 1). As both students often fail to optimally misrepresent their preferences, they gain from a replacement of MSEQ by MSIM. In contrast, the actual matching outcomes of students s_5 and s_7 in MSEQ are, on average, 0.3141 and 0.1662 rank points higher than in equilibrium.

Only for top-grade student s_1 and regular student s_8 do realized and equilibrium matching outcomes largely coincide in MSIM and in MSEQ. For s_1, the strategic decision problem in MSEQ is rather simple as she just needs to reveal her preferences truthfully in order to obtain

[30] In principle, students who already receive their least best option in equilibrium may do even worse in the realized matching if they fail to secure a seat at all. This can only happen if they do not submit a full preference list.

Table 5: Individual performance measure, mean value by mechanism and student

	MSIM	MSEQ	MSIM − MSEQ
Student 1	-0.0084	-0.0639	0.0555***
	(0.0126)	(0.0125)	[0.0178]
Student 2	-0.0701	-0.3278	0.2577***
	(0.0302)	(0.0301)	[0.0426]
Student 3	-0.0644	-0.3833	0.3189***
	(0.0324)	(0.0323)	[0.0458]
Student 4	-0.0812	-0.2694	0.1882***
	(0.0384)	(0.0382)	[0.0541]
Student 5	-0.0308	0.2833	-0.3141***
	(0.0283)	(0.0281)	[0.0399]
Student 6	-0.0112	0.0306	-0.0418
	(0.0278)	(0.0276)	[0.0392]
Student 7	0.0588	0.2250	-0.1662***
	(0.0311)	(0.0309)	[0.0439]
Student 8	0.0084	-0.0222	0.0306**
	(0.0099)	(0.0098)	[0.0139]

Notes: ***,** denotes statistical significance at the 1%- and 5%-level, respectively. Entries in columns two and three are the cell-specific mean values (over all rounds) of the individual performance measure in MSIM and MSEQ, respectively. Entries in column four are the (student-specific) coefficient estimates of an indicator variable for MSIM from an OLS regression of the individual performance measure on a constant and the MSIM indicator itself. The individual performance measure is defined in equation (2). Standard deviations are in round and standard errors in squared brackets.

her first choice. Student s_8, in turn, has little to gain from the mistakes of top-grade students in MSEQ, as she is consistently ranked at the bottom of the universities' preference lists.

We summarize the above in:

Result 2. Individual performance by student type. *The two mechanisms differ significantly in the actual matching outcome for the different types of students. In general, top-grade students benefit from replacing MSEQ by MSIM, while regular students are worse off under MSIM than under MSEQ.*

The individual benefit or loss from introducing MSIM does not only differ across students but also depends on the market characteristics. Table 6 provides for each student type and each market the difference in the individual performance measure between MSEQ and MSIM. The results show that student s_2, for instance, benefits significantly from the introduction of

Table 6: Difference in individual performance measure between MSIM and MSEQ, by market and student

	Market 1: Fully aligned	Market 2: Student aligned	Market 3: University aligned	Market 4: Split aligned	Markets 1-4
Student 1	0.0778^{**}	0.1000^{**}	0.0444	0.0000	0.0555^{***}
	[0.0361]	[0.0456]	[0.0270]	[0.0312]	[0.0178]
Student 2	0.4444^{***}	0.0663	0.3889^{***}	0.1333	0.2577^{***}
	[0.0736]	[0.0815]	[0.0917]	[0.0857]	[0.0426]
Student 3	0.0222	0.7084^{***}	0.1111	0.4333^{***}	0.3189^{***}
	[0.0773]	[0.1122]	[0.0696]	[0.0810]	[0.0458]
Student 4	0.3667^{***}	-0.4870^{***}	0.4778^{***}	0.4000^{***}	0.1882^{***}
	[0.0867]	[0.1023]	[0.0870]	[0.1033]	[0.0541]
Student 5	-0.6667^{***}	-0.1126^{**}	-0.1778^{**}	-0.3000^{***}	-0.3141^{***}
	[0.0793]	[0.0526]	[0.0731]	[0.0900]	[0.0399]
Student 6	-0.0667	-0.0778	0.3556^{***}	-0.3778^{***}	-0.0418
	[0.0412]	[0.0606]	[0.0802]	[0.0945]	[0.0392]
Student 7	-0.1111^{**}	-0.1314^{*}	-0.2000^{*}	-0.2222^{**}	-0.1662^{***}
	[0.0508]	[0.0705]	[0.1035]	[0.1111]	[0.0439]
Student 8	0.0222	0.0111	0.0556	0.0333^{*}	0.0306^{**}
	[0.0273]	[0.0254]	[0.0368]	[0.0190]	[0.0139]

Notes: ***,**,* denotes statistical significance at the 1%-, 5%- and 10%-level, respectively. Each entry is the (cell-specific) coefficient estimate of an indicator variable for MSIM from an OLS regression of the individual performance measure on a constant and the MSIM indicator itself. The performance measure is defined in equation (2). Standard errors are in squared brackets.

MSIM in markets 1 and 3, but her gains are not statistically significant in markets 2 and 4. Student s_4, in contrast, benefits significantly from switching to MSIM in markets 1, 3, and 4, but loses from such a switch in market 2.

Choice of weakly dominant strategies. Depending on the preferences of students and universities, it can be more or less difficult to reach the student optimal stable matching under MSEQ. In general, for all top-grade students but s_1 (who is always guaranteed her reported top choice) the optimal application strategy for MSEQ will depend on the application strategies of others. However, in two out of four experimental markets there are top-grade students who have weakly dominant strategies that are not truthful. In the following, we provide a detailed discussion of market 1 where all top-grade students have a weakly dominant strategy. The Online Appendix contains the same analysis for market 3 where top-grade students s_2 and s_4 have non-truthful weakly dominant strategies. Before proceeding, we should emphasize that the notion of "weak dominance" used in the following refers to a fixed game of complete information. In order to infer that these strategies are always optimal, students need detailed information about universities' preferences. This should be contrasted with the weak dominance of truth-telling for MSIM and the second part of MSEQ, which does not require any information about universities' or other students' preferences.

In market 1, it is a weakly dominant strategy for top-grade students s_2 to s_4 to rank only their (truly) most preferred university for the first and submit preferences truthfully for the second part of MSEQ. The reason is that all universities rank applicants exclusively on the basis of their average grades/indices. With the just mentioned strategies, (i.) s_2 can guarantee herself a place at her most preferred university, *irrespective* of the behavior of s_1, and (ii.) s_3 and s_4 can guarantee themselves a place at their second most preferred university, while maintaining an option of being matched to their first choice if s_1 and/or s_2 make a mistake.

Table 7 provides the shares of students s_2 to s_4 in market 1 who are matched to their first, second, third, and fourth preference in MSEQ. (The Online Appendix, part F.2, provides a detailed overview of the preferences received by all students in markets 1 to 4.) These shares are calculated separately for students who play their weakly dominant strategy (upper panel) and for those who do not (lower panel). Cells shaded in gray indicate equilibrium outcomes. Participants who play the weakly dominant strategy are a minority among students s_2, s_3, and s_4, and their share decreases with grade rank. In only 34 out of 90 observations (37.78%) do participants in the role of s_2 play their weakly dominant strategy (see last column in the upper panel of Table 7). This fraction shrinks to 10/90 (11.11%) and 3/90 (3.33%) for students s_3 and s_4, respectively. The lower ranked among the top-grade students thus seem to be less inclined to truncate their preferences.[31]

The failure of students to play their weakly dominant strategy leads to a significant re-

[31] This out-of-equilibrium phenomenon suggests that top-grade students may often be inclined to "lock up" positions in the first part of MSEQ due to risk aversion. While the cardinal representation of payoffs does not matter for the structure of complete information equilibria or of the set of stable outcomes, it could well have an effect on the top-grade applicants' tendency to give in to the "lock up" incentive. A satisfactory analysis of this relationship, although potentially very interesting, would have required a significant number of additional treatments. For budgetary reasons we have therefore decided against looking at this issue in more detail. (See Echenique et al. (2009) for some preliminary experimental evidence on the effects of the cardinal representation of preferences in complete information matching markets.) To the best of our knowledge the only paper that studies the effects of changes in risk aversion/the cardinal representation of payoffs in the context of matching markets under incomplete information is Coles and Shorrer (2013). We thank an anonymous referee for making us aware of the just mentioned paper and the issues discussed in this footnote.

duction in their realized payoffs: For instance, only 39.29% of s_2 students who do not play their weakly dominant strategy receive their top choice (the equilibrium outcome) compared to 100% of those who choose the weakly dominant strategy. Similar results also hold for students s_3 and s_4.

For market 3, we find similar results for student 2 and student 4 who choose their weakly dominant strategy in 40 out of 90 (44.4%) and 20 out of 90 (22.2%) cases (see the Online Appendix, part F.3, for more details). We can summarize these findings in:

Result 3. Weakly dominant strategies. *The majority of top-grade students fails to choose the weakly dominant truncation strategy when it is available.*

Of course most real markets are more complex than our experimental market 1 and applicants often do not have a weakly dominant strategy. Thus, successful preference manipulations in MSEQ are likely to be more difficult in reality. The improvement in performance due to mechanism MSIM compared to MSEQ found in the experiment should therefore provide a lower bound for the possible improvement in real markets.

Table 7: Preference received in market 1 in MSEQ, by student type and strategy

	Preference 1	Preference 2	Preference 3	Preference 4	N
	Weakly dominant strategy played				
Student 2	**100.00%**	0.00%	0.00%	0.00%	34
Student 3	60.00%	**40.00%**	0.00%	0.00%	10
Student 4	100.00%	**0.00%**	0.00%	0.00%	3
	Weakly dominant strategy not played				
Student 2	**39.29%**	44.64%	16.07%	0.00%	56
Student 3	17.50%	**58.75%**	21.25%	2.50%	80
Student 4	10.34%	**50.57%**	28.74%	10.34%	87

Notes: Entries are the share of students matched to the corresponding (induced) preference in each cell. Cells shaded in gray indicate equilibrium outcomes.

Finally, Table 7 illustrates for market 1 that students may not only end up with below- but also with above-equilibrium payoffs in MSEQ. Consider, for instance, student s_3. If student s_2 does not play her weakly dominant strategy, s_3 can secure herself a seat at her most preferred university by playing her weakly dominant strategy – and can thus realize above-equilibrium payoffs. In fact, 60% of s_3 students who play their weakly dominant strategy are matched to their first preference and are thus better off than in equilibrium. If, in contrast, s_3 fails to play her weakly dominant strategy, she might end up with below-equilibrium payoffs. Thus, MSIM eliminates both the up- and the downside risks of MSEQ.

4.3 Learning

Our experimental setting allows participants to learn over time. Although in real life the majority of students apply only once to university, we employ a multi-rounds setup to account for the fact that students might have many opportunities to learn about the mechanism in real life before making their choice. Also, allowing for some learning in the experiment seems necessary in order not to bias our results against the more complicated sequential procedure MSEQ. Each market was played three times (once in rounds 1–4, once in rounds 5–8, and once in rounds 9–12) and participants were informed about the actual matching of all players in previous rounds. Thus, participants had the opportunity to learn about the strategic properties of each market. It can be expected that the difference between the two mechanisms diminishes over time, as this difference is due to the failure of top-grade students to misrepresent their preferences optimally. This is what we test in this section.

Table 8 shows, by student type, the difference between MSIM and MSEQ in the individual performance measure of top-grade students separately for rounds 1–4, 5–8, and 9–12. We find that the difference between the two mechanisms decreases significantly in later rounds of the experiment. The individual performance measure of student s_3, for instance, is 0.5167 rank points higher in MSIM than in MSEQ in the first four rounds. This difference shrinks to just 0.2316 rank points in the last four rounds. Likewise, the difference for student s_2 declines from 0.4417 to just 0.0720 rank points. Thus, players learn over time. Nevertheless, significant differences between the two mechanisms persist in later rounds, with an average rank difference of 0.1231 for rounds 9-12. Detailed results for the learning behavior of top-grade students by market can be found in Appendix A.2.[32]

[32] The Online Appendix, part F.4, provides additional results on the learning behavior of regular students, on how often the equilibrium is reached in earlier and later rounds of the experiment, and on how differences between the two mechanisms in the aggregate performance measure evolve over time.

Table 8: Difference in individual performance measure between MSIM and MSEQ, by student and round

Student	Rounds 1–4	Rounds 5–8	Rounds 9–12	All rounds
Student 1	0.1167^{**}	0.0250	0.0250	0.0555^{***}
	[0.0457]	[0.0184]	[0.0188]	[0.0178]
Student 2	0.4417^{***}	0.2583^{***}	0.0720	0.2577^{***}
	[0.0826]	[0.0655]	[0.0693]	[0.0426]
Student 3	0.5167^{***}	0.2083^{***}	0.2316^{***}	0.3189^{***}
	[0.0893]	[0.0707]	[0.0742]	[0.0458]
Student 4	0.2167^{**}	0.1833^{**}	0.1639^{*}	0.1882^{***}
	[0.1071]	[0.0888]	[0.0837]	[0.0541]
Students 1–4	0.3229^{***}	0.1688^{***}	0.1231^{***}	0.2051^{***}
	[0.0369]	[0.0251]	[0.0290]	[0.0183]

Notes: ***,**,* denotes statistical significance at the 1%-, 5%-, and 10%-level, respectively. Each entry is the (cell-specific) coefficient estimate of an indicator variable for MSIM from an OLS regression of the individual performance measure on a constant and the MSIM indicator itself. The performance measure is defined in equation (2). Standard errors are in squared brackets. Standard errors in the last row (Students 1–4) are clustered at the group level of each session-round (as observations from students that participated in the same group of a session are not independent).

We therefore conclude:

Result 4: Learning. *There is some learning of top-grade students over time, as reflected in smaller differences between matching outcomes in MSIM and MSEQ in later rounds of the experiment. However, even in the last four rounds of the experiment (9-12), top-grade students receive significantly better matches in MSIM than in MSEQ.*

Participants may not only learn by playing each market three times. They may also learn by repeatedly playing the role of a specific student. If so, one may expect the difference between MSIM and MSEQ to not only diminish in later rounds of the experiment but also for participants who have already played the same student type in previous rounds. The Online Appendix, part F.5, shows that there is only limited evidence for this type of learning behavior. However, participants who play the exact same type of top-grade student in two consecutive rounds indeed improve their performance in MSEQ (but still do better in MSIM).

4.4 Performance of Mechanisms: University Perspective

So far, our analysis has exclusively focused on the preferences of students. But as the evaluation of students by universities is relevant for more than half of the seats allocated, we also compare the two mechanisms from the point of view of the universities. This is also important because from a social planner's perspective it can be desirable that the best universities get the best students.

Recall that in our experiment, universities were played by a computer which truthfully revealed their preferences. To analyze how universities fare in the two mechanisms, we will employ an approach that is analogous to the one we used to compare outcomes from the student perspective. Here, we directly evaluate the performance of the two mechanisms from the perspective of the individual universities. The reason is that our analysis from the perspective of students shows that there is little difference in the aggregate performance of the two mechanisms. For mechanism $M \in \{\text{MSEQ},\text{MSIM}\}$, let y_{uj}^M be the *sum* of the positions of the two students who were assigned to university $u \in \{W, X, Y, Z\}$ in round j in university u's preferences, i.e., the *aggregate student quality* assigned to u in round j under M. Each place at u that is left unassigned is counted as being assigned a student of position 9, so that $y_{uj}^M \in \{3, \ldots, 18\}$. As above, $k^M(j) \in \{1, 2, 3, 4\}$ denote the market that was played in the jth round of mechanism M. Finally, let $y_{uk^M(j)}^e$ denote the aggregate student quality of u in the outcome of MSIM under truth-telling in market $k^M(j)$. The performance of M for university u in round j is then defined as

$$M_{uj}^U = \frac{y_{uk^M(j)}^e - y_{uj}^M}{2}. \tag{3}$$

As for the students, these performance measures take on positive (negative) values if outcomes under M in our experiment are better (worse) than in equilibrium. We will again look at the average of this measure, denoted by $\overline{M}_{uj}^U$, across all experimental rounds. Note that our measure does not condition on how many students receive their assignment through the regular quota in equilibrium and the experiment. Rather, we always compare the average quality of the student(s) assigned to a university with the equilibrium quality.

Table 9: Mean of university performance measure, by mechanism and university

	MSIM	MSEQ	MSIM−MSEQ
University W	-0.0602	-0.5292	0.4689^{***}
	(0.2949)	(0.6981)	[0.0401]
University X	-0.2731	-0.6125	0.3394^{***}
	(0.6892)	(0.8866)	[0.0593]
University Y	0.1120	0.6236	-0.5116^{***}
	(0.4858)	(0.6415)	[0.0425]
University Z	0.0420	0.1833	-0.1413^{***}
	(0.3039)	(0.7138)	[0.0410]

Notes: *** denotes statistical significance at the 1%-level. Entries in columns two and three are the cell-specific mean values (over all rounds) of the university performance measure in MSIM and MSEQ, respectively. Entries in column four are the (university-specific) coefficient estimates of an indicator variable for MSIM from an OLS regression of the university performance measure on a constant and the MSIM indicator itself. The university performance measure is defined in equation (3). Standard deviations are in round and standard errors in squared brackets.

Table 9 shows the university performance by university type and mechanism. Differences between realized and theoretical outcomes are on average smaller (i.e., $\overline{M}_{uj}^{U}$ is closer to zero) in MSIM than in MSEQ for all four universities. This is not surprising and mirrors our previous findings that the equilibrium matching is significantly more often realized in MSIM than in MSEQ (see Result 1). More interesting are the substantial differences in the performance measure between the two mechanisms for the four universities. The performance measure for university W, for instance, is 0.4689 rank points higher in MSIM than in MSEQ. University W thus prefers, on average, the students that it admits in MSIM over those that it admits in MSEQ. The same applies to university X. In contrast, the two universities Y and Z, which are generally less preferred by students, on average prefer the matching under MSEQ over the matching under MSIM. We thus find

Result 5: University performance. *On average, universities W and X admit more preferred and universities Y and Z less preferred students in MSIM relative to MSEQ. The most popular universities thus fare better in MSIM than in MSEQ, while the two other universities are better off in MSEQ than in MSIM.*

This result mirrors our previous findings for top-grade and regular students. As top-grade students often fail to optimally manipulate their preference lists in MSEQ, they are frequently matched to lower ranked universities. This does not only harm the top-grade students themselves, but also the most popular universities that usually prefer top-grade over regular students. Lower ranked universities, in contrast, can benefit from the mistakes made by the top-grade students, as they might be able to admit top-grade instead of regular students.

5 Conclusion

Quotas can be implemented in centralized matching procedures in a number of ways. In a laboratory experiment we have tested two alternative mechanisms that give priority to certain groups of applicants. The first mechanism is sequential and fills the quota for students with priority first and then the remaining seats. This procedure mimics the mechanism in Germany for university admissions in medicine and related subjects, where 20% of available university seats are reserved for top-grade students (top-grade quota). The other mechanism is a modification of Gale and Shapley (1962)'s student-proposing deferred acceptance algorithm that was recently proposed by Westkamp (2013). It simultaneously fills the seats reserved for students with priority and all other students and redistributes free capacity in each round.

In theory, both mechanisms lead to the same matching outcome when restricting attention to equilibria in strategies that are not weakly dominated. However, the experimental results show that the equilibrium matching is significantly more often realized in the simultaneous than in the sequential mechanism. The simultaneous mechanism significantly improves the matching outcome for top-grade students relative to the current sequential mechanism. The current mechanism harms top-grade students, as they often fail to grasp the strategic issues involved. We therefore conclude that quotas for top-grade students should be implemented in a simultaneous manner.[33]

[33] A potentially promising alternative to such a change in mechanism would be to keep MSEQ and amend it by a mandatory training system in which applicants play several rounds against pseudodata. Given that the difference between MSIM and MSEQ becomes less pronounced in later rounds of our experiment, see Section 4.3, this also has the potential to improve the performance of the admission market. How to design

The experiment allows us to identify the reasons for why the outcome of the simultaneous mechanism under truth-telling is not reached in the sequential mechanism although it is an equilibrium of the revelation game. We find that participants fail to use truncation strategies optimally. Our finding is supported by previous results of Braun et al. (2010). Their analysis of the actual admission data of the clearinghouse shows that only about one quarter of top-grade students truncate their rank-order list submitted in the first part of the procedure. Yet, the choices of applicants observed in the admission data can always be rationalized by unobserved preferences and are thus difficult to evaluate. In the laboratory, we know the induced preferences of applicants and can thus unambiguously identify certain choices as violations of weak dominance.

There is also a more general lesson to be learnt from our experiment: If two mechanisms are available that lead to the same allocation in theory, the mechanism with a simpler optimal strategy is likely to perform better in practice. This is true even if subjects are given advice regarding their optimal strategies.

A Appendix

A.1 Experimental markets

A.1.1 Experimental market 1: Fully aligned

Students' and universities' preferences are as follows:

$$\begin{aligned} P_{s_i}: &\quad W \succ X \succ Y \succ Z, && \forall i \in \{1, 2, ..., 8\}, \\ P_u: &\quad s_1 \succ s_2 \succ s_3 \succ s_4 \succ s_5 \succ s_6 \succ s_7 \succ s_8, && \forall u \in \{W, X, Y, Z\}. \end{aligned}$$

The outcome of MSIM under truth-telling is

and implement such training systems is an interesting practical question. We thank an anonymous referee for this suggestion.

$$\mu^1 = \begin{pmatrix} & W & X & Y & Z \\ \text{Top Grade} & s_1 & s_3 & \emptyset & \emptyset \\ \text{Regular} & s_2 & s_4 & s_5, s_6 & s_7, s_8 \end{pmatrix}.$$

For this market, all equilibrium outcomes of the game induced by MSEQ have the same structure:

- s_1, s_2 are matched to W,
- s_3, s_4 are matched to X,
- s_5, s_6 are matched to Y,
- s_7, s_8 are matched to Z.

This can be shown as follows: first, given the preferences of universities it is easy to see that s_1 and s_2 must be matched to W in *any* equilibrium. Given this, s_3 and s_4 must both end up matched to X. But then, the best university that s_5 and s_6 can obtain in equilibrium is Y. By the previous arguments, they are guaranteed a place at Y (in equilibrium) as long as they rank it higher than Z. Finally, for s_7 and s_8 the only possible equilibrium allocation is to receive a place at Z. Given that for both of them this is better than remaining unmatched, they must end up matched to Z in any equilibrium.

One equilibrium of MSEQ that yields the outcome of MSIM under truth-telling is the following: let all top-grade students rank only their most preferred university for the first part of the procedure. For the second part, let all students submit their true preferences.

The only arbitrariness in equilibrium outcomes of MSEQ lies in exactly which *type* of place students get at their assigned universities. For example, there exists an equilibrium outcome in which s_2 gets the top-grade place at W, while s_1 gets the regular place at W. In our experiment, however, students were indifferent as to which type of place they received. In particular, all equilibrium outcomes were equivalent from students' perspectives. Similar comments apply to the other experimental markets below.

A.1.2 Experimental market 2: Student aligned

Students' and universities' preferences are as follows:

$$
\begin{aligned}
P_{s_i}: &\quad W \succ X \succ Y \succ Z && \forall i \in \{1, 2, ..., 8\} \\
P_W: &\quad s_1 \succ s_3 \succ s_2 \succ s_4 \succ s_5 \succ s_6 \succ s_7 \succ s_8, && \\
P_X: &\quad s_1 \succ s_5 \succ s_2 \succ s_3 \succ s_4 \succ s_6 \succ s_7 \succ s_8, && \\
P_u: &\quad s_1 \succ s_2 \succ s_3 \succ s_4 \succ s_5 \succ s_6 \succ s_7 \succ s_8, && \forall u \in \{Y, Z\}
\end{aligned}
$$

The outcome of MSIM under truth-telling is

$$
\mu^2 = \begin{pmatrix} & W & X & Y & Z \\ \text{Top Grade} & s_1 & s_2 & s_4 & \emptyset \\ \text{Regular} & s_3 & s_5 & s_6 & s_7, s_8 \end{pmatrix}.
$$

There are two types of equilibria of the game induced by MSEQ:

(Type 1) s_1 matched to W in the top-grade quota

In this case, s_3 must be matched to W in the regular quota, s_2, s_5 must be matched to X, s_4, s_6 to Y, and s_7, s_8 to Z.

(Type 2) s_1 matched to W in the regular quota

In this case, s_2 must be matched to W in the top-grade quota, s_3, s_5 must be matched to X, s_4, s_6 to Y, and s_7, s_8 to Z.

Note that equilibria of the second type involve s_1 playing the weakly dominated strategy of ranking no university for the first part.

One equilibrium that implements the outcome of MSIM under truth-telling is the one where s_3 ranks only W for the first part, s_2 ranks X first, and all other submitted rankings correspond to true preferences.

A.1.3 Experimental market 3: University aligned

Students' and universities' preferences are as follows:

$$\begin{aligned}
P_{s_i}: &\quad W \succ Y \succ X \succ Z && \forall i \in \{1,2,5,6\},\\
P_{s_i}: &\quad X \succ Y \succ W \succ Z && \forall i \in \{3,4,7,8\},\\
P_u: &\quad s_1 \succ s_2 \succ s_3 \succ s_4 \succ s_5 \succ s_6 \succ s_7 \succ s_8, && \forall u \in \{W,X,Y,Z\}.
\end{aligned}$$

The outcome of MSIM under truth-telling is

$$\mu^3 = \begin{pmatrix} & W & X & Y & Z \\ \text{Top Grade} & s_1 & s_3 & \emptyset & \emptyset \\ \text{Regular} & s_2 & s_4 & s_5, s_6 & s_7, s_8 \end{pmatrix}.$$

It is straightforward to show that all equilibrium outcomes of the game induced by MSEQ must yield the same matching of students to universities. One equilibrium which implements the outcome of MSIM under truth-telling is obtained if all top-grade students list only their true first choice for the first part of the procedure and all students submit their true preferences for the second part of the procedure.

A.1.4 Experimental market 4: Split aligned

Students' and universities' preferences are as follows:

$$
\begin{aligned}
P_{s_i}:&\quad W \succ Y \succ X \succ Z && \forall i \in \{1,3,5,7\},\\
P_{s_i}:&\quad X \succ Y \succ W \succ Z && \forall i \in \{2,4,6,8\},\\
P_X:&\quad s_1 \succ s_5 \succ s_2 \succ s_3 \succ s_4 \succ s_6 \succ s_7 \succ s_8, &&\\
P_u:&\quad s_1 \succ s_2 \succ s_3 \succ s_4 \succ s_5 \succ s_6 \succ s_7 \succ s_8, && \forall u \in \{W,Y,Z\}.
\end{aligned}
$$

The outcome of MSIM under truth-telling is

$$
\mu^4 = \begin{pmatrix} & W & X & Y & Z \\ \text{Top Grade} & s_1 & s_2 & \emptyset & \emptyset \\ \text{Regular} & s_3 & s_4 & s_5, s_6 & s_7, s_8 \end{pmatrix}.
$$

It is again straightforward to show that all equilibrium outcomes of the game induced by MSEQ must yield the same matching of students to universities. As in Markets 1 and 3, one equilibrium which implements the outcome of MSIM under truth-telling is obtained if all top-grade students list only their true first choice for the first part of the procedure and all students submit their true preferences for the second part of the procedure.

A.2 Learning Behavior by Market

Table A1: Difference in individual performance measure between MSIM and MSEQ, by market, student and round

Student	Rounds 1-4	Rounds 5-8	Rounds 9-12	All rounds
		Market 1		
Student 1	0.2000^{*}	0.0333	0.0000^{***}	0.0778^{**}
	[0.1006]	[0.0333]	[0.0000]	[0.0361]
Student 2	0.6667^{***}	0.4667^{***}	0.2000^{**}	0.4444^{***}
	[0.1345]	[0.1376]	[0.0884]	[0.0736]
Student 3	0.2333	-0.1667	0.0000	0.0222
	[0.1774]	[0.1195]	[0.0830]	[0.0773]
Student 4	0.6667^{***}	0.3000^{**}	0.1333	0.3667^{***}
	[0.1938]	[0.1215]	[0.1107]	[0.0867]
		Market 2		
Student 1	0.2000	0.0667	0.0333	0.1000^{**}
	[0.1213]	[0.0463]	[0.0352]	[0.0456]
Student 2	0.2000	0.0667	-0.0741	0.0663
	[0.1319]	[0.1511]	[0.1424]	[0.0815]
Student 3	0.8333^{***}	0.5333^{***}	0.7630^{***}	0.7084^{***}
	[0.1993]	[0.1733]	[0.2139]	[0.1122]
Student 4	-0.5667^{***}	-0.4333^{**}	-0.4667^{***}	-0.4870^{***}
	[0.1982]	[0.1820]	[0.1495]	[0.1023]
		Market 3		
Student 1	0.1000	0.0333	0.0000	0.0444
	[0.0735]	[0.0333]	[0.0000]	[0.0270]
Student 2	0.5333^{***}	0.3667^{***}	0.2667	0.3889^{***}
	[0.1733]	[0.1363]	[0.1645]	[0.0917]
Student 3	0.3000	0.1000	-0.0667	0.1111
	[0.1854]	[0.0714]	[0.0463]	[0.0696]
Student 4	0.4333^{**}	0.5333^{***}	0.4667^{***}	0.4778^{***}
	[0.1919]	[0.1333]	[0.1178]	[0.0870]
		Market 4		
Student 1	-0.0333	-0.0333	0.0667	0.0000
	[0.0571]	[0.0333]	[0.0667]	[0.0312]
Student 2	0.3667^{*}	0.1333^{*}	-0.1000	0.1333
	[0.1982]	[0.0768]	[0.1391]	[0.0857]
Student 3	0.7000^{***}	0.3667^{**}	0.2333	0.4333^{***}
	[0.1233]	[0.1407]	[0.1492]	[0.0810]
Student 4	0.3333^{*}	0.3333^{*}	0.5333^{***}	0.4000^{***}
	[0.1817]	[0.1764]	[0.1815]	[0.1033]

Notes: ***,**,* denotes statistical significance at the 1%-, 5%-, and 10%-level, respectively. Each entry is the (cell-specific) coefficient estimate of an indicator variable for MSIM from an OLS regression of the individual performance measure on a constant and the MSIM indicator itself. The performance measure is defined in equation (2).

References

Abdulkadiroglu, A.: 2010, Controlled School Choice, *mimeo*, Duke University.

Abdulkdadiroglu, A. and T. Sönmez: 2003, School Choice: A Mechanism Design Approach, *American Economic Review* **93**, 729–747.

Abdulkdadiroglu, A., Pathak, P. and A. Roth: 2005, The New York City High School Match, *American Economic Review, Papers & Proceedings* **95**, 364–367.

Abdulkadiroglu, A., Pathak, P., Roth, A. and T. Sönmez: 2006, Changing the Boston School Choice Mechanism: Strategy-proofness as equal access, *mimeo*, Harvard University

Azevedo, E. and J. Leshno: 2011, The College Admissions Problem With a Continuum of Students, *mimeo*, Harvard University.

Braun, S., Dwenger, N., and D. Kübler: 2010, Telling the Truth May Not Pay Off: An Empirical Study of Centralized University Admissions in Germany, *The B.E. Journal of Economic Analysis & Policy (Advances)* **10**, Article 22.

Cabrales, A., Charness, G. and L. Corchon: 2003, An Experiment on Nash Implementation, *Journal of Economic Behavior and Organization* **51**, 161–193.

Calsamigla, C., Haeringer, G., and F. Klijn: 2010, Constrained School Choice: An Experimental Study, *American Economic Review* **100**, 1860–1874.

Cason, T. , Saijo, T., Sjostrom, T., and T. Yamato: 2010, Secure implementation experiments: Do strategy-proof mechanisms really work?, *Games and Economic Behavior* **57**, 206–235.

Chen, Y. and O. Kesten: 2011, From Boston to Shanghai to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms, *mimeo.*

Chen, Y. and T. Sönmez: 2006, School choice: An Experimental Study, *Journal of Economic Theory* **127**, 202–231.

Coles, P. and R. Shorrer: 2013, Optimal Truncation in Matching Markets, *mimeo*

Dubins, L. and D. Freedman: 1981, Machiavelli and the Gale-Shapley Algorithm, *American Mathematical Monthly* **88**, 485–494.

Dur, U. and O. Kesten: 2012, Sequential Assignment Problems and Two Applications, *mimeo*

Echenique, F., Wilson, A. J., and L. Yariv: 2009, Clearinghouses for Two-Sided Matching: An Experimental Study, *mimeo.*

Echenique, F., and B. Yenmez: 2013, How to control Controlled School Choice, *mimeo*

Ehlers, L., Hafalir, I., Yenmez, B., and M. Yildirim: 2012, School Choice with controlled choice constraints: Hard bounds versus soft bounds, *mimeo.*

Ergin, H.: 2002, Efficient resource allocation on the basis of priorities, *Econometrica* **70**, 2489–2497

Ergin, H. and T. Sönmez: 2006, Games of School Choice under the Boston mechanism, *Journal of Public Economics* **90**, 215–237

Featherstone, C. and M. Niederle: 2011, School Choice Mechanisms under Incomplete Information: An Experimental Investigation, *mimeo.*

Fischbacher, U.: 2007, z-Tree: Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics* **10**, 171–178.

Gale, D. and L. Shapley: 1962, College Admissions and the Stability of Marriage, *American Mathematical Monthly* **69**, 9–15.

Guillen, P. and O. Kesten: 2012, Matching Markets with Mixed Ownership: The Case for a Real-life Mechanism, *International Economic Review*, **53**, 1027–1046

Hafalir, I., Yenmez, B., and M. Yildirim: 2013, Effective affirmative action in school choice, *Theoretical Economics* **8**, 325–363

Hatfield, J. and P. Milgrom: 2005, Matching with Contracts, *American Economic Review* **95(4)**, 913–935.

Kagel, J. and A. Roth: 2000, The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment, *Quarterly Journal of Economics* **65**, 201– 235.

Kamada, Y. and F. Kojima: 2010, Improving Efficiency in Matching Markets with Regional Caps: The Case of the Japan Residency Matching Program, *mimeo.*

Katok, E., Sefton, M., and A. Yavas: 2001, Implementation by iterative dominance and backward induction: an experimental comparison, *Journal of Economic Theory* **104**, 89–103.

Kojima, F.: 2012, School choice: Impossibilities for affirmative action, *Games and Economic Behavior*, **75**, 685-693

Kojima, F. and P. Pathak: 2009, Incentives and Stability in Large Two-sided Matching Markets, *American Economic Review* **99**, 608–627.

Kominers, S. and T. Sönmez: 2013, Designing for diversity in matching, *mimeo.*

Pais, J. and A. Pintér: 2008, School Choice and Information: An Experimental Study on Matching Mechanisms, *Games and Economic Behavior* **64**, 303–328.

Pais, J., Pintér, A., and R. Veszteg: 2011, College Admissions and the Role of Information: An Experimental Study, *International Economic Review* **52**, 713-737.

Pathak2011a Pathak, P.: 2011, The mechanism design approach to student assignment, *Annual Reviews of Economics* **3**, 513–536

Pathak, P. and T. Sönmez: 2013, School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation, *American Economic Review* **103**, 80-106.

Roth, A.: 1982, The Economics of Matching: Stability and Incentives, *Mathematics of Operations Research* **7**, 617–628.

Roth, A.: 1984, The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory, *Journal of Political Economy* **92(6)**, 991–1016.

Roth, A.: 1985, The College Admissions Problem is not Equivalent to the Marriage Problem, *Journal of Economic Theory* **36(2)**, 277–288.

Sefton, M. and A. Yavas: 1996, Abreu-Matsushima mechanisms: experimental evidence, *Games and Economic Behavior* **16**, 280–302.

Shapley, L. and H. Scarf: 1974, On Cores and Indivisibility, *Journal of Mathematical Economics*, **1**, 23–37.

Sönmez, T. and U. Ünver: 2011, Matching, Allocation, and Exchange of Discrete Resources, in *Oxford Handbook of Market Design*, ed. Z. Neeman et al.

Westkamp, A.: 2013, An Analysis of the German University Admissions System, *Economic Theory* **53**, 561–589.