# Iterative Versus Standard Deferred Acceptance: Experimental Evidence*

Inácio Bó† Rustamdjan Hakimov‡

November 8, 2017

**Abstract**

We run laboratory experiments where subjects are matched to colleges, and colleges are not strategic agents. We test the Gale-Shapley Deferred Acceptance (DA) mechanism versus the Iterative Deferred Acceptance Mechanism (IDAM), a matching mechanism based on a new family of procedures being used in the field, in which students make applications one at a time. We consider two variations of IDAM: one in which students are only informed about whether they have been tentatively accepted or not (IDAM-NC) and one in which they are additionally informed at each step of the tentative cutoff values for acceptance at each school (IDAM). A significantly higher proportion of stable outcomes is reached both under IDAM and IDAM-NC than under DA. The difference can be explained by a higher proportion of subjects following an equilibrium strategy akin to truthful behavior under IDAM and IDAM-NC than the truthful behavior itself under DA. Moreover, the provision of intermediate cutoff values in IDAM leads to higher rates of equilibrium behavior than in IDAM-NC. We associate the benefits of iterative mechanisms relative to DA with the feedback of the outcome of applications provided between steps of the iterative mechanisms. This feedback allows subjects to learn that the deviating strategies from truthful do not work in an intended way. Our findings provide substantial support for the rising practice of using iterative mechanisms in centralized college admissions in practice.

*classification*: C78, C92, D63, D78, D82

*Keywords*: Market Design, Matching, Iterative Mechanisms, College Admissions, Experiments.

---

*We thank Marina Agranov, Nina Bonge, Li Chen, Yan Chen, Morimitsu Kurino, Dorothea Kübler, Jennifer Rontganger, Alexander Nesterov, Alvin Roth, Andrew Schotter, and Utku Ünver for helpful comments.

†Address: WZB Berlin Social Science Center, Reichpietschufer 50, D-10785 Berlin, Germany; website: http://www.inaciobo.com; e-mail: inacio.bo@wzb.eu; phone: +49 (0)30-25491-292.

‡Address: WZB Berlin Social Science Center, Reichpietschufer 50, D-10785 Berlin, Germany; e-mail: rhakimov@wzb.eu.

# 1 Introduction

The use of central clearinghouses to match prospective students to universities or schools[1] has been steadily increasing in recent decades. As have the theoretical and empirical studies of the corresponding procedures. The vast majority of these procedures require students to submit rank-ordered lists of universities, and universities to submit rankings (or priorities) of students (often based on the students' grades in exams). These submissions are subsequently used to allocate students to universities. One of the primary goals of the designers of these procedures is to ensure the stability of the allocation, that is, the allocation is non-wasteful and no student envies another with a lower priority at the college she was assigned. The stability is important since it constitutes fairness in this context, as priorities are typically based on ability-related measures. Given that designers also account for the incentives induced in the students, the Gale-Shapley Deferred Acceptance (DA) mechanism is often seen as the best mechanism: it is strategy-proof and (constrained) efficient from the students' perspective [Balinski and Sönmez, 1999]. One problem, however, is that growing experimental and empirical evidence suggests that participants often misreport their preferences under DA despite truth-telling being a dominant strategy under the mechanism. This may lead to allocations that are unstable, and thus unfair.

Bo and Hakimov [2016] recently proposed a mechanism for matching students to colleges, called Iterative Deferred Acceptance Mechanism (IDAM), based on a modification of a mechanism currently being used to match students to university programs in Brazil. IDAM asks students to apply to one of the universities available in each of a number of periods. Throughout the allocation process, a university can tentatively retain no more applications than its number of seats and if a university receives more applications than it can accept then it rejects the students with lower priorities and retains the remaining applications. After each step, all students are informed about whether their application was rejected or retained. Moreover, the minimum grades of the retained students are publicly announced by each university at the end of each step. If an applicant is rejected, she can apply to any other university where the cutoff grade is lower than her grade. Students are not allowed to change their choices while tentatively selected by some university. IDAM is based on a family of mechanisms that we have observed in the field over the last few years, notably in Brazil and in the province of Inner Mongolia (China). These mechanisms differ from most of those analyzed in the literature in that they do not require students to submit rank-ordered lists of colleges, but instead repeatedly ask them to apply to one college among a set of "available" ones. Another common feature of these mechanisms is that

[1] From this point on, we will refer to the institutions as universities. Unless explicitly stated, all arguments also hold for schools.

they provide feedback about tentative cutoff grades after each step. When the number of periods is large enough, if students follow the simple strategy of applying to the most preferred college among those available at each step of the IDAM mechanism (denoted by the *straightforward strategy*), the matching produced as an outcome is the Student Optimal Stable Matching. This is the matching that is the most preferred by all students among all stable matchings. While, unlike the standard DA mechanism, IDAM does not have a dominant strategy, the authors show that the Student Optimal Stable Match is an equilibrium outcome under a robust equilibrium concept – Ordinal Perfect Bayesian Equilibrium (OPBE) [Bo and Hakimov, 2016]. In a sequential game, a strategy profile is an OPBE if in every period, any deviation strategy is stochastically dominated by following the equilibrium strategy. Though in theory IDAM has arguably "worse" incentive properties for students than its direct counterpart (DA), in practice, students might benefit from the iterative nature of the game, as students interact repeatedly with the mechanism during the determination of the allocation.

This paper addresses the following questions: Are there benefits of using iterative, non-direct mechanisms compared to the theoretically superior DA? Is the equilibrium strategy in IDAM (straightforward strategy) a better predictor of behavior than the dominant strategy of truthful reporting in DA? What effect does the provision of intermediate cutoff values have on individual behavior and the stability of outcomes?

To answer these questions, we run laboratory experiments. Subjects interact with the mechanism for 20 rounds, playing new market parameters (preferences and priorities) in each round. Subjects know their preferences and the preferences of other students, they know their grades, but not the grades of other students. Additionally, they know that the grades were drawn from a uniform distribution with support [0,100]. We run three treatments between subjects: DA, IDAM, and IDAM-NC. Unlike in IDAM, in IDAM-NC students do not observe intermediate cutoff grades at every step. Thus, by comparing IDAM and IDAM-NC, we can identify the effect of the provision of cutoffs on the students' behavior and on the final allocations. We show that both IDAM and IDAM-NC dominate DA in terms of stability. As for individual strategies, there is a statistically significant and large increase in the proportion of straightforward and truthful behavior in all treatments when we compare the first 10 to the last 10 rounds of each treatment, but the proportion of truthful (or straightforward) behavior is the highest in IDAM, second highest in IDAM-NC, and the lowest in DA, with all pairwise differences being statistically significant.

What contributes to the benefits of the iterative mechanisms? First, the iterative feature of IDAM and IDAM-NC leads to the biggest advantage in terms of stability improvements and proportions of "truthful" behavior relative to DA. Second, the provision of cutoffs in IDAM leads to a higher proportion of subjects following the straightforward strategy than in IDAM-NC. This difference is driven by an almost universal adoption of the straightforward strategy by subjects who were rejected at some step of the mechanism and who thus observed intermediate cutoffs. Additionally, the provision of cutoffs allows subjects to observe the fairness of the allocation, which is especially crucial for the persistence of stability between rounds. In other words, a student can see that any envy she may have results from her not choosing a university that was more desirable, when that was still available.

Thus, we identify a disparity between the theory and subjects' behavior in the lab: equilibrium predicts subjects' choices better than the weakly dominant strategy. What drives the success of those iterative mechanisms when compared to DA? We investigate several alternative explanations, and conclude that the benefit of iterative mechanisms comes from the feedback on the outcome of the previous actions that these mechanisms provide within a round. This feedback allows subjects to learn that the deviating strategy from truthful behavior does not work as intended.

The typical deviating strategy from truthful reporting that we observe is essentially the same in all treatments, and we call it the *skipping strategy*. Subjects who play the skipping strategy avoid applying to universities where they think they will be rejected. They do it either because they want to avoid unnecessary rejections, or because they think that rejections at early steps of the mechanism might hamper their chances at subsequent steps. A skipping strategy is "partial" in the sense that it does not include all possible sequences of choices/iterations, but is based on the belief that the outcome will be one among $X$ first-choices in iterative mechanisms or top-ranked in DA. Therefore, in the iterative mechanisms there is no plan for what to do if this belief is incorrect, and in DA there is no clear reasoning behind the ranking below those $X$ top choices. In iterative mechanisms, however, after a subject receives the feedback that none of the $X$ choices were realized, she is allowed to re-strategize. She may plan for further skipping, in which case she may also have to re-strategize again later on. Every time the expectation of being accepted in a certain set of universities is not matched, a subject might update her beliefs regarding how good the skipping strategy is. In this respect there is an asymmetry between the iterative mechanism and DA: there may be multiple failures of the skipping strategy in one round in the iterative mechanism and only one in DA. Thus, the probability that subjects will be "disappointed enough" to abandon the skipping strategy is higher in the iterative mechanisms. If a subject abandons the skipping strategy, she is more likely to play the truthful strategy. While we are unable to test this explanation exactly in the data, we argue that the switch from skipping to the straightforward choices within a round might point to the subjects who realized the failure of the skipping strategy during that round. They would thus be more likely to play the truthful strategy in the next round. This prediction is supported by the data for iterative mechanisms, and not for DA. Subjects who used the skipping strategy at the beginning of a round and then rolled back to "truth-telling" in the same round are more likely to play the truthful strategy in the following round. These individuals, according to our interpretation, started the round with a skipping strategy in mind and acted upon it, potentially multiple times (re-strategized in line with skipping strategy). The fact that they were rejected from the top choices under the skipping strategy leads to an update or multiple updates of the belief about the success of the skipping strategy. Thus, they might realize the skipping strategy does not work, as the mechanism gave a direct indication that the acceptance that was expected did not materialize. Under DA, on the other hand, belief updating about the success of the skipping strategy happens between rounds, and thus the choice of lower-ranked universities does not correlate with actual disappointment in the skipping strategy.

We believe that this finding extends to other comparisons of iterative and direct mechanisms. For instance, our argument might be applied to explain the benefit of the English auction relative to the second price sealed-bid auction. The better performance of the English auction may not come from the fact that players understand it better, but from the fact that during the English auction players quickly realize that a deviating strategy, for instance overbidding, does not work, through observing the hands of other players raised once the clocks are above the true value of the object. The better performance of the (sequential) Ausubel auction relative to the multiple unit Vickrey auction and the benefit of the uniform price multiple unit English auction relative to the sealed-bid version Kagel and Levin (2001, 2009) can be explained by a similar argument. In fact, it is evident from Kagel and Levin (2001, 2009) that the advantage of the dynamic auctions comes from the information about dropouts and not from the fact that it is sequential.

Another example of an iterative mechanism that performs better in experiments is the sequential Serial Dictatorship (SD), when compared to the static SD Li [2015]. In this case the feedback that the sequential mechanism provides to subjects is the set of the available options at the moment of their choice. A subject who at the beginning of sequential SD has a strategy similar to the skipping strategy in mind and who guesses the object correctly would not learn anything, but if there is a better object available in the choice set and the moment of their decision, she would learn that the strategy is wrong and she will thus choose the better object.

The contribution of our paper, therefore, goes beyond the matching literature, as it suggests the channel through which learning is happening in the iterative mechanisms, and which could explain the observed superior performance of sequential allocation mechanisms relative to their direct counterparts. This channel is worth studying further and emphasizes the importance of behavioral considerations in market design.

## 2 Related literature

One common objective of policymakers, when designing college and school admission mechanisms, is to ensure that the outcomes produced are fair. In a fair matching, if a student is not matched to a more preferred university then every student who is matched to that university has a higher priority than her. While this objective is debatable, for university admissions it is the most natural, as the priorities of students are almost always based on grades or other performance-based measures. Balinski and Sönmez [1999] showed that DA can be characterized as the "best" fair mechanism, in that it is strategy-proof and Pareto dominates any other fair mechanism (that is, it is constrained efficient). In fact, variations of the DA mechanism are used in many real-life student matching programs around the world. College and secondary school admissions in Hungary [Biró, 2012], high school admissions in Chicago [Pathak and Sönmez, 2013] and New York City [Abdulkadiroğlu et al., 2009] as well as elementary school admissions in Boston [Abdulkadiroğlu et al., 2006] are examples of real-life use of the DA mechanism.

While strategy-proofness itself may be considered an element of fairness [Pathak and Sönmez, 2008], it also aligns the students' incentives such that the outcomes have the desirable efficiency and fairness properties. Recent empirical and experimental evidence suggests, however, that those incentives may not be fully understood by the subjects and may lead to assignments that are neither fair nor constrained efficient.

The seminal matching experiment in Chen and Sönmez [2006] suggests that truth-telling rates are higher under DA than under the Boston mechanism in the school choice problem. Later experiments, however, showed that truth-telling rates might drop if more information was made available to the participants [Pais and Pintér, 2008, Pais et al., 2011]. Lately, some experiments have been conducted to study learning in DA, in which subjects interact with the mechanism for several rounds. In the baseline treatment of Ding and Schotter [2015], subjects play DA repeatedly and only limited learning is observed after 20 periods. Similar results are presented in Zhu [2015], where the subjects increase the truth-telling rates with experience, but only in one out of the two markets analyzed. Chen and He [2016] show that subjects in the experiments are ready to invest in information about the preferences and priorities of others in DA even after gaining experience with the mechanism, thus showing that they do not fully understand its strategy-proofness. Moreover, evidence of manipulation was found in the flagship application of the student-proposing DA – the match of medical doctors to residency programs in the US (the National Resident Matching Program). Rees-Jones [2017] shows that around 17% of participants manipulate preference reports, and around 5% of them declare this manipulation as strategic.[2]

[2]Though the participants can use only constrained lists of preferences, and thus DA is not strategy-proof, the

The numbers might come across as small, but the misreporting of preferences by even a small fraction of students often translates into non-stable allocations. Similarly, Hassidim et al. [2015] show that some students manipulate submitted lists in the Israeli college admissions, Shorrer and Sóvágó [2017] in the Hungarian college admissions, and Chen and Pereyra [2015] identify manipulations for school admissions in Mexico City.

Several papers have also tried to address the question of how factors outside of the mechanism being used affect truth-telling. Guillen and Hing [2014] and Guillén and Hakimov [2015] test the effect of advice on the properties of the mechanism (in their case, the top trading cycles mechanism, which is also strategy-proof) on the truthful reporting in the lab and the field, respectively. Ding and Schotter [forthcoming] test how the ability to communicate with other participants before the submission of preferences affects truth-telling rates and obtain a negative result: communication makes more subjects change their rank-order lists relative to the no communication case, but changes go both in the direction of truth-telling and in the opposite direction. Ding and Schotter [2015] test the effect of inter-generational advice on the truth-telling rates and show that it has a detrimental effect on truth-telling under DA. Summing up, most of the experimental and empirical evidence raises concerns about participants' understanding of incentives to report truthfully under DA.

Dur et al. [forthcoming] use the fact that the school choice mechanism used in the Wake County Public School System allows for students to interact multiple times with the procedure as a method for empirically identifying strategic players. Interestingly, the dynamic nature of the procedure, and the information that is made available to the participants during the process, makes it somewhat comparable to the IDAM mechanism.

Three recent papers also evaluate experimentally indirect sequential mechanisms. Echenique et al. [2016] consider a two-sided market, with DA being implemented dynamically. The authors found that 48% of outcomes are stable and, surprisingly, the receiving side of optimal stable matching is more likely to be reached than the proposing side. Our results are not comparable to these results as we run experiments in one-sided settings, where colleges are not strategic agents. Moreover, authors do not compare the performance of the dynamic DA to the static counterpart. Gong and Liang [2016] consider, both theoretically and experimentally, the mechanism currently in use to match students to universities in the province of Inner Mongolia, in China. The authors find that, when compared to DA, the Inner Mongolia mechanism exhibits higher truth-telling rates in the environment with low preference correlation, but that this does not translate into a higher rate of stable outcomes. In the high preference correlation environment, on the other hand, there

manipulations that are identified cannot be rationalized by the constrained lists. For instance, the switch of the order of any two hospitals in the submitted list is clearly pointing to a misunderstanding of the incentives of DA.

is a higher proportion of stable outcomes under DA. Although the dynamic mechanism used in Gong and Liang [2016] has some similarities to IDAM, such as the availability of tentative cutoff grades, it is in fact a different mechanism, with different timing and incentives. Our results, therefore, are not directly comparable. Klijn et al. [2016] compare dynamic versions of both the school-proposing and student-proposing versions of DA to its static counterparts in one-sided settings of the school choice problem. The dynamic version of the student-proposing DA that they implement is equivalent to our IDAM-NC treatment.[3] In contrast to our results, however, Klijn et al. [2016] find no statistically significant difference in the proportion of stable outcomes between the standard and dynamic student-proposing DA. We attribute the distinct results to two differences in the experimental setup: the size of the market and the information environment. First, we consider the size of the market, i.e. the number of schools or universities. While the market size used in Klijn et al. [2016] is of four schools and four students, we use markets with eight universities and eight students. Experimental evidence suggests that larger markets (and thus longer rank-order lists) lead to lower truth-telling rates in direct strategy-proof mechanisms, such as DA and TTC [Chen and Kesten, 2015, Hakimov and Kesten, 2014], and thus the larger the market the higher the potential scope is for improvement over DA. Moreover, larger markets make subjects have more interactions with the iterative mechanism within one round, involving on average more choices and observations of its operation, emphasizing its contrast with the one-shot nature of DA. Considering the fact that the mechanism used, IDAM-NC, differs from DA essentially in its sequentiality, a larger market is more likely to lead to the differences between the mechanisms. Regarding the information environment, the difference in the information that subjects have about preferences and priorities might also influence the results: Klijn et al. [2016] use complete information, while in our setup there is uncertainty regarding priorities.

Finally, our experiments are related to the vast experimental auction literature that compares static and dynamic implementation of different auction formats.[4] Kagel and Levin [2001] and Kagel and Levin [2009] compared the sealed-bid multiple items Vickrey auction to the dynamic uniform price English clock auctions and Ausubel auction. Just like IDAM when compared to DA, the Ausubel auction has a weaker incentive property than the sealed-bid Vickrey auction. More specifically, truthful bidding equilibrium is obtained after the iterative elimination of dominated strategies, while in the sealed-bid Vickrey auction sincere bidding is a dominant strategy. In experiments, however, a higher proportion of sincere bidding is observed in the Ausubel auction.

[3] Despite the similarities, those experiments in Klijn et al. [2016] were performed simultaneously and independently from the ones in this paper.

[4] See an extensive survey of the experimental auction literature in Kagel and Levin [2014].

## 3 Mechanisms and hypotheses

In this section we describe the matching mechanisms that we test in the lab. They correspond to the three treatments that we run. At the end of the section we formulate the experimental hypotheses.

### The student-proposing Gale-Shapley deferred-acceptance mechanism (DA)

DA is a direct mechanism. It collects universities' capacities and students' submitted rank-order lists of universities simultaneously, which are used by the algorithm below to produce the final allocation. Universities' priorities over students are strict and are exogenously given.[5]

- Step 1
  - Every student applies to her first choice. Each university rejects the least-ranked students in excess of its capacity and temporarily holds the others.
- Step $k > 1$
  - Every student who is rejected in step $k-1$ applies to the next most preferred university according to the submitted rank-order list. Each university pools together new applicants and those who are held from step $k-1$ and rejects the lowest-ranked students in excess of its capacity. Those who are not rejected are temporarily held by the universities.

The process terminates after any step without rejections. Each university is then matched to the students it holds, and students who are not held at any university are left unmatched.

### The iterative deferred-acceptance mechanism with no cutoffs (IDAM-NC)

IDAM-NC is not a direct mechanism. The allocation procedure is implemented in an iterative way, in which a student applies to one university at a time. After each step a student receives intermediate feedback of whether she was temporarily accepted or rejected at that step. Whenever a student is asked to make a choice, the only restriction is that she cannot choose a university that has rejected her in a previous step. More specifically:

- Step 1

[5] Note that for the goal of this study we use students' grades as an instrument to impose priorities.

- Each student applies to one university.
- Each university rejects the least-ranked students, among those who applied to it, in excess of its capacity and temporarily holds the others. If no application is rejected, the procedure will stop at this step, matching the universities to the students they hold.

- Step $t > 1$
  - Each student who is not held at some university applies to any university that has not rejected her in the previous steps in case it exists.
  - Each university rejects the least-ranked students, among those held and those who applied to it, in excess of its capacity and temporarily holds the others. If no application is rejected, the procedure stops at this step, matching the universities to the students they hold and leaving students who are not held at any university unmatched.

**The iterative deferred-acceptance mechanism (IDAM)**

Just like in IDAM-NC, the allocation procedure is implemented in an iterative way. Differently from the IDAM-NC mechanism, in the IDAM students are additionally informed about the cutoff values of all universities at the end of each step. Whenever a student is asked to make a choice, she is restricted to choosing among universities with a cutoff grade in a previous step that is lower than her grade. More specifically:

- Step 1
  - Each student applies to one university.
  - Each university rejects the least-ranked students, among those who applied to it, in excess of its capacity and temporarily holds the others. If no application is rejected, the procedure stops at this step, matching the universities to the students they hold.
  - The cutoffs of each university are made public. For universities where the number of students held is smaller than their capacities, the value of the cutoff by the end of step 1 is zero. For the other universities, the value of the cutoff is the lowest grade among the grades of the students held at each university.

- Step $t > 1$

Table 1: Predictions of individual behavior by treatments

| DA | IDAM | IDAM-NC |
|---|---|---|
| Truthful reporting (weakly dominant strategy) | Straightforward strategy (OPBE) | Straightforward strategy (OPBE) |

– Each student who is not held at some university applies to any university where the cutoff value is lower than her grade in that university in case it exists.

– Each university rejects the least-ranked students, among those held and those who applied to it, in excess of its capacity and temporarily holds the others. If no application is rejected, the procedure stops at this step, matching the universities to the students they hold, and leaving students who are not held at any university unmatched.

– The cutoffs of each university are made public.

The summary of treatments and predictions of individual behavior are presented in Table 1. In DA, the truthful submission of a rank-ordered list is a weakly dominant strategy, while the straightforward strategy is an equilibrium in IDAM and IDAM-NC.[6]

Based on these observations, we form the following hypotheses:

***Hypothesis 1:*** The proportion of subjects submitting truthful rank-order lists in DA is higher than the proportion of subjects playing the straightforward strategy in IDAM and IDAM-NC.

***Hypothesis 2:*** The proportion of stable outcomes is higher in DA than in IDAM and IDAM-NC.

The second hypothesis is based on hypothesis 1. If subjects do not follow the predictions in Table 1, any matching can be produced as an outcome. Therefore, as we expect more subjects to submit truthfully in DA, on average, we should expect a higher proportion of stable outcomes in DA than in IDAM or IDAM-NC.

# 4 The Experiment

In this section, we present a series of experiments designed to test DA versus IDAM and IDAM-NC. This experimental design is not the first one we tried. We originally ran sessions under

[6]Note that Bo and Hakimov [2016] show that the straightforward strategy is OPBE and not a weakly dominant strategy for IDAM by providing an example of students conditioning their actions on the cutoffs. As for IDAM-NC, even in the absence of cutoffs, a straightforward strategy is not dominant as the strategy space allows students to condition their applications on the step at which they are rejected from a certain university.

constant market conditions (the same preference profiles and priorities) and complete information. Our findings were in line with findings of the current experiments from the perspective of stability and efficiency. However, we identified that some of the design choices made did not allow us to observe the effects of the different mechanisms on individual strategies in a robust way, and so we ran additional experiments which are reported here. As this experiment supersedes the previous one, we report the results of the first one in the online appendix.[7]

## 4.1 Experimental design

In the experiment, there were eight universities that differed in quality and specialization. Each university had only one seat. Universities admitted students based on an exam grade. There were eight students who applied for seats at universities, and each student had a grade for math and a grade for language. Universities M1, M2, M3 accepted based on the math grade only. Universities L1, L2, L3 accepted based on the language grade only. Universities H1 and H2 accepted based on the average grade between math and language. In all treatments, students received 22 euros if they were assigned to their most preferred university, 19 euros to their second most preferred university, 16 euros to their fourth most preferred university, and so on. Students received 1 euro if they were assigned to their least preferred university. Each treatment lasted for 20 rounds. At the end of the experiment, one round was randomly drawn to determine the participants' payoffs. Each round represented a new market. The preferences used in each market were generated following the designed market idea of Chen and Sönmez [2006]. For each of market, we generated the qualities of universities uniformly and randomly distributed between [0,40]. It corresponded to the utility of each university for subjects and was common for all subjects. Additionally, for each subject and university we generated a random component of utility from the interval [0,20]. Finally, each subject had an additional utility of 20 for one of the groups of universities: either math, language or hybrid. This is used to model the student-specific preference for a field of study. The group for each subject was determined randomly. The resulting utilities were transformed into ordinal preferences. The grades were independently drawn, in each round, from the uniform distribution with support [1,100] for math and language.[8] There is a unique stable matching in all markets except for those used in rounds 8, 14, and 20, where there are two stable matchings.

In the experiment, subjects could see tables with the ordinal preference of all students, as well as the distribution of exam grades, but they could only see the realization of their own exam grade. This design choice was done for the following reasons: complete information about

[7] The online appendix can be found at http://www.inaciobo.com/research.html.

[8] The details of all markets are presented in the online Appendix.

preferences make it closer to reality than no information, as by observing the preference table, subjects could have an idea about the popularity of each university, which is typically known in reality. The realizations of the grades of other students were not known, and we argue that this also approximates the informational conditions to reality. After each round a subject received feedback about the university in which she had been given a slot, and not the allocation of other students.

We use markets with eight schools, which are relatively large if compared with the experimental literature, for two reasons:

1. There is evidence that a relatively high share (from 65 to 85 percent) of all subjects converge to submitting truthfully in DA in markets with three or four schools (Chen and Kesten [2015], Zhu [2015]) in repeated experiments. However, based on cross studies observations and evidence in Chen and Kesten [2015][9] for DA and Hakimov and Kesten [2014] for TTC, we expect lower proportions of truthful submissions in DA when the rank-ordered lists are longer. Thus, we create a larger room for potential improvement for the iterative mechanisms.[10]

2. The large number of schools is crucial to test the difference between DA and IDAM, as the iterative mechanisms are less sensitive than the direct ones to the increase in the number of universities, as the decision in each step is the submission of just one university from the list.[11]

As mentioned in the previous section we use three treatments: DA, IDAM, and IDAM-NC. The comparison of iterative mechanisms with DA is the focus of the current paper. We use both IDAM and IDAM-NC to disentangle the effect that the provision of cutoff grades during the execution of the mechanism has from the simple iterative nature of the procedure, on both subjects' behavior and the allocation properties.

The experiment was run at the experimental economics lab at the Technical University of Berlin. We recruited student subjects from our pool with the help of ORSEE [Greiner et al.,

[9]The truth-telling rates are lower in six-school environment than in four-school environment (75% versus 45% in the last 10 periods).

[10]Our markets with eight universities and eight students are the largest markets considered so far in repeated matching experiments: Ding and Schotter [2015] use a market with five students but only four different profiles, Chen and Kesten [2015] with four and six students, Gong and Liang [2016] and Klijn et al. [2016] both use markets with four students and four schools, and Zhu [2015] with three students. The only exception is the large-scale experiment by Chen et al. [2015], however, the large market in the experiment was created by increasing the number of students with similar profiles, while we create a higher number of different student profiles.

[11]In the case of one school, DA and IDAM are the same, while in the case of 1,000 schools, the submission of the full list in DA is almost unfeasible in practice, while following the straightforward strategy in IDAM still constitutes a simple task at each given step.

2003]. The experiments were programmed in z-Tree [Fischbacher, 2007]. For each of the three treatments, independent sessions were carried out. Each session consisted of 24 participants that were split into three matching groups of eight for the entire session. We use fixed matching groups in order to increase the number of independent observations and allow for maximum learning. We are not concerned about repeated games caveats, due to the fact that every round represents a new environment, and incomplete information does not allow subjects to identify the strategies of the players and their identities in previous rounds.[12] In total, 12 sessions with 288 subjects were conducted. Thus, we have 96 subjects and 12 independent observations per treatment. On average, the experiment lasted 115 minutes and the average earnings per subject were 25.20 euros, including a show up fee of 5 euros.

At the beginning of the experiment, printed instructions were given to the participants (see Appendix). Participants were informed that the experiment was about the study of decision-making, and that their payoff depended on their own decisions and the decisions of the other participants. The instructions were identical for all participants of a treatment, explaining in detail the experimental setting. Questions were answered in private. After reading the instructions, the experimenter went through the solution of an example of an allocation task on the whiteboard and allowed for public questions. After that, all individuals participated in a multiple choice quiz to make sure that everybody understood the main features of the experiment.

After the quiz, and before the start of the first round, participants were asked to solve an allocation task which appeared on the screen of their computers. The solution of the task had to be typed in and, if it was correct, participants earned 2 euros. For the DA treatment, participants were shown the submitted list of virtual students on the screen and their grades and they had to determine the final allocation. For the IDAM and IDAM-NC treatments, participants saw the decision of each student in the first step and had to determine the retained and rejected students in each school. If it was done correctly they were informed about the decisions of the rejected students, and so on until a final allocation was reached. In case of a mistake at any step the task stopped and the solution was counted as incorrect. We introduced the incentivized task to be sure that every participant paid enough attention to the details of the mechanisms and to have a measure of the understanding of the mechanics of mechanisms.

[12] Additionally, there are no reputational concerns in the game that could hamper the interpretation of subjects' strategies as one-shot game strategies, given the subjects' experience.

## 4.2 Results

We first present the aggregate results on the level of allocations in order to compare the treatments. In the second step, we study individual behavior in the treatments to compare it to the equilibrium predictions and to shed light on the reasons for the aggregate findings. The significance level of all our results is 5%, unless otherwise stated. In the section we use either non-parametric tests if the data are defined on the level of independent observations or clustered regressions on the level of independent observations if the data are defined on the individual or round level. The details of the regression are presented in notes of tables with p-values. In the main text we only use p-values, without mentioning the details, in order to simplify the exposition. We use signs > in the results between treatments to communicate significantly higher. We use => to communicate higher at the 10% significance level.

### 4.2.1 Aggregate results: Stability and efficiency

In this section, we compare properties of the allocations reached in each of the treatments. We take two perspectives on each of the parameters: learning within treatments and comparison of the outcomes between treatments in the last 10 rounds of the experiment and in all 20 rounds.

**Result 1 (Stability):**

1. There is a significant increase in the proportion of stable outcomes reached in the last 10 when compared to the first 10 rounds in all treatments.

2. A comparison of average proportions of stable outcomes in the last 10 rounds leads to the following results: IDAM>DA, IDAM-NC>DA. Comparison of average proportions of stable outcomes in all rounds leads to the following results: IDAM>DA, IDAM-NC>DA.

**Support.**

Figure 1 presents the average proportions of stable outcomes by treatments grouped by five rounds.

The proportions of stable outcomes by treatments and rounds are presented in Table 2. We observe a significant increase in the proportions of stable outcomes in all treatments. The average proportion of stable outcomes is significantly higher in IDAM and IDAM-NC than in DA. The difference is significant in both the first 10 and the last 10 rounds of the experiment. Thus, we reject Hypothesis 2. As for the difference in the number of stable outcomes between IDAM and IDAM-NC, we observe no significant difference. This indicates that the difference regarding stability between DA and two treatments with iterative mechanisms is driven mainly by the

Table 2: Proportions of stable allocations by treatments:

| | DA | IDAM | IDAM-NC | DA=IDAM, p -value | DA=IDAM-NC p-value | IDAM=IDAM-NC p-value |
|---|---|---|---|---|---|---|
| Round 1-10 (1) | 26.7% | 54.2% | 52.5% | 0.00 | 0.01 | 0.85 |
| Round 11-20 (2) | 47.5% | 76.7% | 68.3% | 0.00 | 0.01 | 0.34 |
| All rounds (3) | 37.1% | 65.4% | 60.4% | 0.00 | 0.00 | 0.47 |
| p-value first 10 = last 10 (4) | 0.00 | 0.01 | 0.03 | | | |

*Notes:* All the p-values are p-values for the coefficient of the dummy in the probit regression of the dummy for the stable outcome on the dummy for the corresponding treatment (columns 5, 6, 7) or the last 10 rounds (row 4). The standard errors of the probit models are clustered on the level of the matching groups. Thus, for within-treatment regressions we have 12 clusters (row 4), and for between treatments 24 clusters (columns 5, 6, 7).

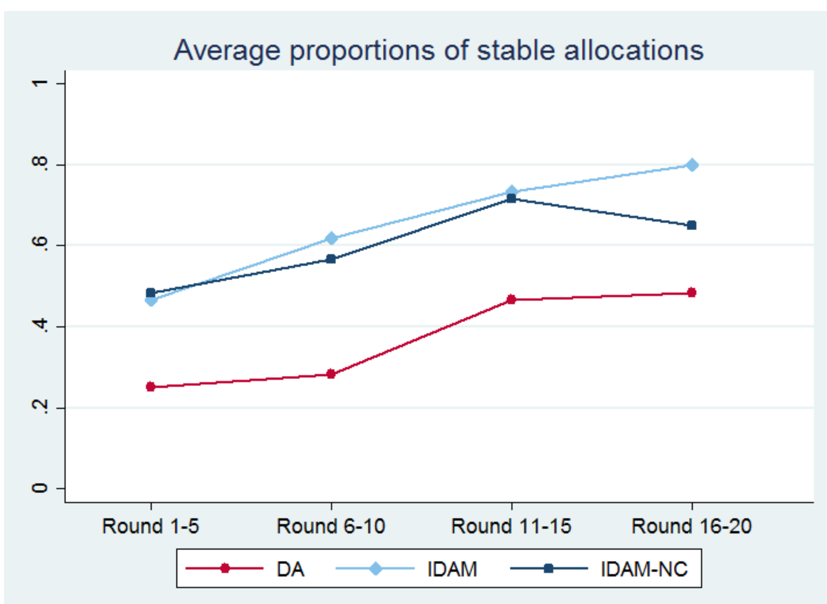


Figure 1: Proportions of stable outcomes

iterative nature of these mechanisms, and not by the provision of information about cutoffs. Note, however, that in the last five rounds of the experiment, the proportion of stable outcomes is on average 15% higher in IDAM than in IDAM-NC (and the difference is 10% significant, p=0.09).

Another way to look at the stability results is to consider only persistent stable outcomes. Sometimes a stable outcome is reached by a group for one or two rounds, but it is not reached in the following rounds. In the following, we treat these outcomes as non-persistent stable outcomes, as this pattern might be a sign that some subjects are still trying to change strategies, hoping that could lead them to obtain a better outcome. Next, we consider only persistent stable outcomes i.e., situations in which stable outcomes are reached in a certain period and then in all rounds until the last period of the experiment. For each group, we calculate the number of consecutive stable outcomes before the last round. Thus, for instance, if a group reached stable outcomes in rounds 16, 18, 19, and 20, but not in 17, we count the persistent stability of this group to be equal 3. If a group failed to reach a stable outcome in round 20 then the variable is equal to 0. The number of consecutive stable allocations can be interpreted as an indicator of how robust the groups are in reaching a stable allocation, despite the different parameters of the markets between rounds.[13]

**Result 2 (Persistent stability):** The comparison of the average number of consecutive stable allocations before the last round leads to the following results: IDAM>DA, IDAM>IDAM-NC.

**Support.** Figure 2 presents the average number of consecutive stable allocations before the last round by treatments.

The average number of consecutive stable allocations in IDAM is the highest among the three treatments and is equal to 5.7. Thus, on average, in rounds 14 or 15, stable allocations are reached in IDAM and remain so until round 20. The difference is significant in comparison to both other treatments: Wilcoxon rank-sum test (12 values versus 12 values) two-sided p-values for the comparison of IDAM with DA is 0.01, and IDAM with IDAM-NC 0.05. The difference between DA and IDAM-NC is not significant (p-value 0.38). Thus, despite the absence of difference between the average proportions of stable allocations in IDAM and IDAM-NC, once the persistent, or robust, stability is considered, IDAM performs significantly better than IDAM-NC. Thus, persistence of stability can be attributed to the effect of the information: providing intermediate feedback about the cutoff grades reveals to students the fact that the allocation reached is stable or "fair" with respect to their last choices, while in IDAM-NC, students do not

[13] For the markets in which there are more stable allocations, we consider any stable allocation. This choice is inconsequential, though. There were only four instances in which an allocation was stable but not student-optimal: two groups in DA, one in IDAM and one in IDAM-NC. All of them in the 20th round.

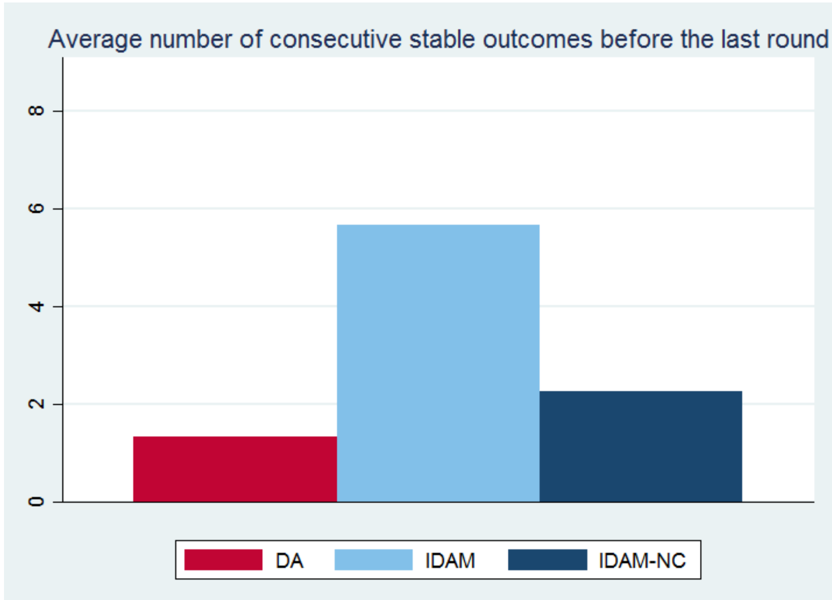


Figure 2: Number of consecutive stable outcomes

know that and it might induce them to change their strategies between rounds. We investigate this hypothesis in further detail in the next section when we analyze individual strategies.

Summing up results 1 and 2, we observe that both treatments with an iterative deferred-acceptance mechanism lead to higher proportions of stable outcomes. Interestingly, there is no significant difference in the proportion of stable outcomes for IDAM and IDAM-NC, but the availability of information about cutoffs leads to a higher persistence of stable allocations between rounds.

Though the main focus of this paper from the perspective of allocation analysis is stability, as we are motivated by college admissions where priorities have to be respected, we also compare the efficiency of the outcomes produced by the mechanisms evaluated. In order to analyze the efficiency of an allocation, we define efficiency as the sum of the payoffs of all the subjects in the allocation produced, divided by the sum of the payoff in the student-optimal stable allocation.

**Result 3 (Efficiency):** There is no significant difference in efficiency between treatments in the last 10 rounds, but if all rounds are considered, a comparison of the average efficiency leads to the following: IDAM>DA, IDAM-NC>DA.

**Support.** Table 3 presents the average payoffs of participants by treatments and rounds, grouped by five.

Row 4 in Table 3 presents the p-values for the significance of the difference in efficiency between

Table 3: Average efficiency of allocations by treatments:

| | DA | IDAM | IDAM-NC | DA=IDAM, p -value | DA=IDAM-NC p-value | IDAM=IDAM-NC p-value |
|---|---|---|---|---|---|---|
| Round 1-10 (1) | 97.4% | 99.2% | 99.4% | 0.06 | 0.03 | 0.66 |
| Round 11-20 (2) | 98.3% | 99.0% | 99.4% | 0.44 | 0.27 | 0.67 |
| All rounds (3) | 97.8% | 99.1% | 99.4% | 0.04 | 0.00 | 0.60 |
| p-value first10 = last 10 (4) | 0.55 | 0.84 | 0.97 | | | |

*Notes:* All the p-values are p-values for the coefficient of the dummy in the OLS regression of the average payoff of allocations on the dummy for the corresponding treatment (columns 5, 6, 7) or the last 10 rounds (row 4). The standard errors of the regressions are clustered on the level of the matching groups. Thus, for within-treatment regressions we have 12 clusters (row 4), and for between treatments 24 clusters (columns 5, 6, 7).

the first 10 and the last 10 rounds in each treatment. Efficiency does not differ between the first 10 and the last 10 rounds in any of the treatments. However, the level of efficiency is close to 100%. This can be explained by the fact that not every violation of stability of an allocation leads to a lower sum of payoffs, and some may even lead to a higher sum of payoffs. As for the between-treatment comparison, in IDAM and IDAM-NC treatments we observe a higher average efficiency than in DA in the first 10 rounds of the experiment. The difference is also significant if all rounds are considered (see rows 5 and 6 of Table 3).

This completes our analysis of allocations by treatments. IDAM and IDAM-NC perform significantly better than DA from the perspective of stability, and also outperform DA from the efficiency perspective.

### 4.2.2 Individual behavior

Next, we analyze individual strategies of experimental subjects in order to test hypothesis 1 and better understand the drivers of the observed differences between the proportions of stable outcomes produced between treatments. We consider the proportions of subjects following straightforward behavior in IDAM and IDAM-NC, and the proportion of subjects submitting truthful rank-ordered lists in DA. To simplify the language, we introduce the Truthful Criterion:

- **Truthful Criterion**:
  - DA: Truthful submitted list.[14]
  - IDAM: Straightforward strategy: applying to the best university from the preference

[14]Note that the only undominated strategy, given the information available for subjects at the experiments, is to submit the full truthful list. In our setting there is no "minimum guaranteed allocation," like a district school, for instance, as the relative grades of the other students are unknown.

Table 4: Proportions of behavior in line with the Truthful Criterion by treatments:

| | DA | IDAM | IDAM-NC | DA=IDAM, p -value | DA=IDAM-NC p-value | IDAM=IDAM-NC p-value |
|---|---|---|---|---|---|---|
| Round 1-10 (1) | 41% | 59% | 54% | 0.00 | 0.01 | 0.42 |
| Round 11-20 (2) | 55% | 82% | 74% | 0.00 | 0.00 | 0.04 |
| All rounds (3) | 48% | 70% | 64% | 0.00 | 0.00 | 0.15 |
| p-value first10=last10 (4) | 0.00 | 0.00 | 0.00 | | | |

*Notes:* All the p-values are p-values for the coefficient of the dummy in the probit regression of the dummy for strategy in line with the Truthful Criterion on the dummy for the corresponding treatment (columns 5, 6, 7) or the last 10 rounds (row 4). The standard errors of the probit models are clustered on the level of the matching groups. Thus, for within-treatment regressions we have 12 clusters (row 4), and for between treatments 24 clusters (columns 5, 6, 7).

list among the universities that have a cutoff value lower than the student's exam grade.

– IDAM-NC: Straightforward strategy: applying to the best university among the universities that have not rejected her.

The Truthful Criterion is based on the theoretical properties of the mechanisms. In DA it is a weakly dominant strategy, while in IDAM and IDAM-NC it is an OPBE. It leads to the student-optimal stable match in all treatments, if played by all subjects of a group.

**Result 4 (Behavior in line with the Truthful Criterion):**

1. There is a significant increase in the proportion of subjects behaving in line with the Truthful Criterion in the last 10 compared to the first 10 rounds in all treatments.

2. The comparison of average proportions of subjects behaving in line with the Truthful Criterion in the last 10 rounds leads to the following results: IDAM>DA, IDAM-NC>DA, IDAM>IDAM-NC. The comparison of the average proportions of subjects behaving in line with the Truthful Criterion in all rounds leads to the following results: IDAM>DA, IDAM-NC>DA. [15]

**Support:** Table 4 presents the proportions of behavior in line with the Truthful Criterion and the p-values for a test of equality of these proportions between treatments and between the first 10 and the last 10 rounds.

[15]Result 4 is robust to two changes of the definition of truthfulness in DA. First, if instead of requiring the full truthful list we count as truthful all truthful submissions until the student-optimal stable match for each subject, result 4 remains. Second, if instead of requiring the full truthful list in DA we count as truthful all truthful submissions until the assigned university, result 4 remains.

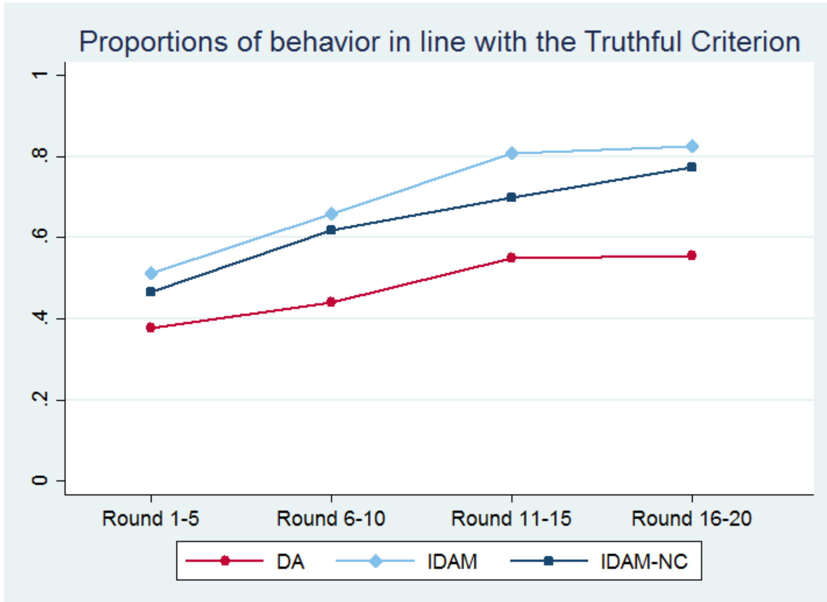


Figure 3: Behavior in line with the Truthful Criterion

Row 4 in Table 4 presents the p-values for the difference between the first 10 and the last 10 rounds of the experiment by treatments. We observe a significant increase in the Truthful Criterion behavior in all treatments, including DA (see row 4 of Table 4). Despite the relatively high increase in truthful submissions in DA, the increase of straightforward behavior in IDAM and IDAM-NC is even higher.

Figure 3 presents the average proportions of behavior in line with the Truthful Criterion by treatments and rounds. Overall, the proportion of strategies in line with the Truthful Criterion is higher in IDAM and IDAM-NC than in DA. The difference is significant for the first 10, last 10, and all rounds jointly. Thus, we can reject Hypothesis 1. Despite the fact that from a theoretical perspective we should expect higher truthful rates in DA, as it is a weakly dominant strategy, it turns out that IDAM and IDAM-NC lead to higher rates of straightforward behavior by subjects. As for the difference between IDAM and IDAM-NC, the proportions of behavior in line with the Truthful Criterion is higher in IDAM, with the difference being significant for the last 10 rounds of the experiment. Note that this dominance of IDAM relative to IDAM-NC in the last 10 rounds is translated into a higher proportion of stable outcomes, but the difference in the proportion of stable outcomes is not significant. Thus, we can conclude that the provision of intermediate feedback about cutoffs has a positive effect on the proportion of straightforward behavior.

Next, we take a closer look at the manipulations in all three treatments.

**Result 5 (Violation of the Truthful Criterion and its determinants):**

1. In all three treatments, the probability of violation of the Truthful Criterion is higher the lower the students' grades are.

2. The correct solution of the incentivized allocation task at the beginning of the experiment significantly increases the probability of submission in line with the Truthful Criterion in IDAM and IDAM-NC but not in DA.

3. If IDAM and IDAM-NC are considered: in IDAM, the most common violation of straightforward behavior is skipping the most preferred university in the first step of the mechanism. The proportion of violations of straightforward behavior (that is, not choosing the most preferred option among the available ones,) decreases dramatically after the first step and reaches almost zero after the third step of the mechanism. In IDAM-NC the proportion of violations does not change between steps. Controlling for the number of options available, the proportion of violations in IDAM-NC is significantly higher than in IDAM, starting from the second step of the procedure.

4. The violations of the Truthful criterion are costly for subjects. The highest average payoff loss from a deviation from the Truthful Criterion in DA, and equals 2.37 euros. In IDAM-NC and IDAM, the cost is significantly lower than in DA and equals 1.72 euros and 1.60 euros, respectively.

**Support:** Table 5 presents the marginal effects of probit regressions of the dummy for play in line with the Truthful Criterion by treatments. In all treatments, in line with result 4, the probability of submission in line with the Truthful Criterion increases with the experience and its marginal effect is higher in IDAM and IDAM-NC than in DA. The probability of submitting in line with the Truthful Criterion is also higher the higher the student's grades. In all three treatments, the coefficient for the average grade is positive and significant. Notably, this observation is in line with field data for college admissions in Israel, where the DA mechanism is used [Hassidim et al., 2015].

The dummy for the correct solution of the allocation task at the beginning of the experiment is positive and significant in IDAM and IDAM-NC, while not significant in DA. It is an interesting finding: if a participant of the experiment was successful in understanding the mechanics of the iterative mechanism from instructions and examples, she is 12.9% and 16.2% more likely to use a straightforward strategy in IDAM and IDAM-NC, respectively. In DA, the fact that a participant of the experiment could produce the allocation of the DA mechanism in the task does not translate into a higher probability of truthful submissions. This is potentially one of the main concerns about the DA mechanism: the fact that truthful reporting is a weakly dominant strategy is not easy to infer even with a good understanding of the mechanics of the mechanism. For IDAM and IDAM-NC, going through the steps of an allocation problem one by one, subjects are more likely to use a straightforward strategy when interacting with the mechanism.

Table 5: Marginal effects of probit model of submissions in line with the Truthful Criterion by treatments

| | (1)<br>Truthful Criterion<br>dummy<br>DA | (2)<br>Truthful Criterion<br>dummy<br>IDAM | (3)<br>Truthful Criterion<br>dummy<br>IDAM-NC |
|---|---|---|---|
| Round | .015*** | .025*** | .024*** |
| | (.002) | (.003) | (.003) |
| Average grade | .003*** | .004*** | .006*** |
| | (.001) | (.001) | (.001) |
| Correct solution of the allocation task | -.048 | .129*** | .162*** |
| | (.076) | (.031) | (.059) |
| Female dummy | -.005 | .008 | .057 |
| | (.077) | (.048) | (.064) |
| Math-related major dummy | .081 | .017 | .058 |
| | (.070) | (.037) | (.063) |
| Observations | 1920 | 1920 | 1920 |
| No. of clusters | 12 | 12 | 12 |
| log(likelihood) | -1282.22 | -1042.37 | -1110.40 |

Standard errors in parentheses, and are clustered on the level of matching groups.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Math-related major dummy is equal to 1 if a subject is studying economics, computer science or math, and 0 otherwise.

Next, we take a closer look at the differences in straightforward behavior between IDAM and IDAM-NC. Table 6 presents the marginal effects of probit regressions of straightforward submission (that is, choosing the most preferred available university in each step) depending on the round, the number of available options, and the treatment by steps of the mechanism (the sample includes only IDAM and IDAM-NC). Figure 4 presents the average proportions of straightforward behavior by groups of five rounds in IDAM and IDAM-NC, depending on the step of the mechanism. Decisions in step 1 are the first decisions, and being straightforward requires choosing the most preferred university. There is no significant difference between IDAM and IDAM-NC in terms of the proportions of truthful most preferred university submissions (see column one of Table 6). Note that one of the possible explanations for deviations from straightforward behavior is that subjects try to skip applications to universities where they think they are likely to be rejected. We call this "skipping" behavior. The fact that we observe the average grade as a

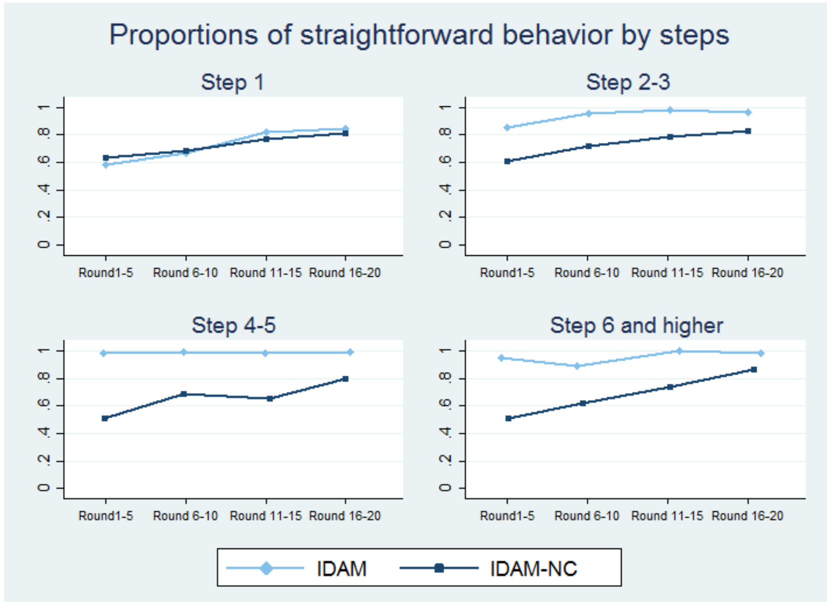


Figure 4: Behavior in line with the Truthful Criterion

predictor of straightforward behavior supports this explanation.

Decisions in steps 2 and 3 are taken after one and two rejections, respectively. Note that in IDAM straightforward behavior requires the application to the best university among those available given the published cutoffs. In IDAM-NC subjects are not aware of the cutoffs, and thus straightforward behavior requires the application to the best university among those that have not rejected the subject in the previous steps. If a subject applied to the true top choice at the first step, straightforward behavior requires the application to the second best university at the second step. If the subject did not apply to the best university at the first step, straightforward behavior requires the application to the best university at the second step. This leads to the following: due to the simple fact that the number of options available after each step in IDAM is weakly smaller than under IDAM-NC, the likelihood of a behavior being consistent with a straightforward strategy, by pure mechanics, is greater under IDAM than under IDAM-NC. To control for that fact, we added the number of options available to the student at each step to the regression.

Starting from steps 2 and 3, the probability of straightforward behavior is significantly higher in IDAM than in IDAM-NC, controlling for the number of options available (see columns 2, 3, and 4 of Table 6). The difference can be explained by the fact that in IDAM participants see the cutoffs in the feedback of the first and all consequent steps, which makes the mechanism different from IDAM-NC. Skipping behavior is most likely driven by participants' beliefs about

Table 6: Marginal effects of probit model of straightforward submission depending on the average grade by cycles of the mechanism.

| | (1) Straightforward step=1 | (2) Straightforward step=2 or step=3 | (3) Straightforward step=4 or step=5 | (4) Straightforward step>5 |
|---|---|---|---|---|
| Round | .017*** | .011*** | .008*** | .020*** |
| | (.002) | (.001) | (.002) | (.004) |
| IDAM | .007 | .186*** | .215*** | .272*** |
| | (.041) | (.033) | (.031) | (.042) |
| # of available options | | -.009 | -.074*** | -.025*** |
| | | (.006) | (.012) | (.008) |
| Observations | 3840 | 3081 | 1290 | 1269 |
| log(likelihood) | -2107.61 | -1228.01 | -578.49 | -672.46 |

Standard errors in parentheses, and are clustered on the level of matching groups.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. # of available options represents the number of universities with corresponding cutoff grades being lower than the student's at the moment of decision in IDAM, and equals the number of schools that have not rejected the student before the current step of the mechanism.

which university is available to them and a desire to avoid unnecessary choices and rejections or a belief that rejection at the first steps might decrease the chances of acceptance in later steps. We are unable to disentangle these motives for skipping. In IDAM-NC, the feedback of each step does not provide any information except that the chosen university is not available to them (the one that rejected them in the previous step). In IDAM, however, starting from the second step, participants see the cutoffs and thus some uncertainty about the availability of universities is resolved, and therefore subjects do not need to form beliefs about universities where they have a chance of being accepted.[16]

From the previous section we observe differences between treatments in terms of stability and efficiency, which implies that deviation from equilibrium strategies are consequential for the allocation. This means that at least some skipping done by subjects is irrational, and subjects are over-pessimistic about their prospects in some universities. But how costly are these mistakes for subjects from an individual perspective? Table 7 presents results of the OLS estimation of the effect of misreporting on the payoff of the subjects.

[16]Unfortunately, we cannot run simulations to have a clear understanding of whether some manipulations were the best-response to the strategies of other players. While it would be possible to verify whether students who are not submitting truthful preferences under DA are best responding to their counterpart strategies, that analysis is not possible under both the IDAM and the IDAM-NC mechanisms. The reason is that, in sequential games, strategies are themselves functions that specify different actions for each configuration an agent may face at a given step. Therefore, knowing how other agents would respond to a counterfactual strategy is something that requires more information that can be observed in the experimental setting. Thus, the rates of deviations from the straightforward behavior are just suggestive measure of the proportions of subjects who actually fail to best respond.

Table 7: OLS regression of payoff on the dummy for truthful behavior

| | (1)<br>Payoff |
|---|---|
| Non-truthful | -2.37*** |
| | (0.16) |
| Non-truthful in IDAM-NC | 0.65*** |
| | (0.13) |
| Non-truthful in IDAM | 0.77*** |
| | (0.20) |
| Dummies for each student ID in each round | (yes) |
| Observations | 5760 |
| $R^2$ | 0.79 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are clustered on the level of matching groups and are presented in in parentheses

OLS includes 160 dummies for each combination of ID and round, to account for the "role-specific" fixed effects, as the roles (combination of preferences and grades) vary the prospects of earning high payoffs. Thus, the dummies for misreporting presents the average differences between subjects who play truthfully relative to subjects who play non-truthfully, controlling for the role of the subjects. In DA, misreporting on average leads to a loss of 2.37 euros (note that maximum payoff for the allocation is 22 euros), while the deviations are 65 cents less costly in IDAM-NC and 77 cents less costly in IDAM. The difference can be explained by the fact that skipping in iterative treatments is more often non-consequential for the payoff than in DA. Summing up the results of the experiment, we found that IDAM and IDAM-NC lead to a significantly higher proportion of stable allocations than DA. The iterative structure of the game leads to a higher proportion of stable allocations, while the provision of intermediate feedback during the procedure contributes to the persistence of this stability between rounds. The aggregate findings are explained by a higher proportion of subjects playing in line with the Truthful Criterion: the iterative nature of the game leads to a large improvement in the proportion of Truthful Criterion behavior when compared to DA, but the provision of cutoffs makes the improvement significantly larger. The latter is explained by an almost universal adoption of a straightforward behavior by the subjects after the first rejection, once the cutoff grades of universities are published.

Summing up individual behavior analysis, we found that IDAM and IDAM-NC lead to a significantly higher proportion of stable allocations than DA. The iterative structure of the game leads to a higher proportion of stable allocations, while the provision of intermediate feedback during the procedure contributes to the persistence of this stability between rounds. The aggregate findings of the higher proportion of stable allocations in IDAM and IDAM-NC than DA are explained by a higher proportion of subjects playing in line with the Truthful Criterion: the iterative nature of the game leads to a large improvement in the proportion of Truthful Criterion behavior when compared to DA, but the provision of cutoffs makes the improvement significantly larger. The latter is explained by an almost universal adoption of a straightforward behavior by the subjects after the first rejection, once the cut-off grades of universities are published. The main puzzling question remains: why do subjects behave more truthfully in iterative mechanisms than in DA, despite the presence of the dominant strategy in the latter? The next subsection investigates possible explanations of this puzzle.

# 5 Reasons for the superior performance of the iterative mechanisms

In this subsection we explore possible reasons behind the benefits of the iterative mechanisms relative to DA. Why do subjects converge faster to truthful behavior in iterative mechanisms than in the direct mechanism? In this section we devote particular attention to the difference between DA and IDAM-NC, as these treatments differ only in the dimension of iterative implementation, and thus strategies are more directly comparable. As for IDAM, we list results only when the meaningful comparison is possible, but due to the restriction of options starting from the second step of the mechanism, strategies of DA and IDAM-NC are not comparable to the strategies in IDAM, and thus we do not present them. We present three possible explanations for the differences between the iterative and static mechanism, and provide data from the experiments either supporting or contradicting the explanation.

## 5.1 Is it about the difference in complexity of reasoning in DA and IDAM-NC, and limited ability for contingent thinking of subjects?

One of the potential reasons for why IDAM-NC performs better than DA could be the fact that if subjects think step by step when submitting the lists in DA, the reporting becomes more complex the further down the rank-ordered list they go. In the case of the decision about the top choice, the decision is simple, when making a decision about the second choice, in DA these subjects would try to understand the procedures inside of the mechanism that would lead to the use of the second reported choice and, based on that, would decide which university to report as the second choice. This requires contingent thinking. The third choice is even harder. If a subject has a limited capacity for contingent thinking, or a limited understanding of the mechanics of how DA works, she is likely to be able to make a correct or “truthful” report at the top of the lists before deviating from truthful behavior in the subsequent choices. In IDAM-NC, subjects always report only one university at each step at a time, and thus might make less mistakes. This implies that the percentage of truthful decisions should not depend on the step, and the differences from DA would be explained by more thought-through later choices relative to lower-ranked reports in DA. Figure 5 presents proportions of truthful reporting grouped by the length of the list, relevant for the allocation in DA and IDAM-NC. The top-left figure of Figure 5 presents the proportion of truthful top choices for the subjects who were assigned to the top submitted choice. The top-right figure of Figure 5 presents the proportion of truthful top two choices for the subjects who were assigned to the second submitted choice. If the argument above was the reason for the difference between DA and IDAM-NC, we should observe no difference between DA and IDAM-NC in the initial figures, but increasing difference for the later figures, as mistakes in DA should appear more in the bottom part of the lists. In fact, we observe the opposite. The biggest difference comes from submission of the relevant top, and relevant top two choices.

Thus we can conclude that the difference between the iterative mechanisms and DA does not come from the fact that in iterative mechanisms the choices are made only one at a time, which simplifies the thinking about each choice.

## 5.2 Experience

In this subsection we analyze whether the difference in the proportion of behavior in line with the Truthful Criterion between DA and the iterative mechanism can be explained by the experience. As subjects take more decisions within each round in iterative mechanism, it might be that this drives a learning of the equilibrium strategy. We thus conjecture here that the need to take a decision makes subjects think of the mechanism more, and they thus understand it better after each decision. Subjects take only one decision per round in DA, and up to eight decisions in the iterative mechanism. In order to address non-stationary data we define our database as panel, with rounds being the time dimension. Table 8 presents results of the random effects probit estimation of the probability of play in line with the Truthful Criterion. In iterative mechanisms (models (2) and (3) ) we control for the average number of decisions taken in all previous rounds. If our conjecture of understanding the mechanism better with every decision is true then this variable should positively correlate with the probability of playing in line with the Truthful Criterion, controlling for other factors. However, the coefficient is significant only for IDAM, and not for IDAM-NC. Thus, just the higher number of decisions taken in IDAM-NC cannot explain the higher proportion of truthful behavior in IDAM-NC. In case of IDAM, however, every decision enhances learning, as IDAM has a richer feedback on every decision,


Figure 5: Proportions of truthful behavior depending on the rank of submitted choice

including the feedback about the fairness of the intermediate allocations, and is thus different in this dimension from DA and IDAM-NC. In IDAM the more decisions a subject takes the more information about the consequences of her actions to the outcome of the market she observes through the updated cutoffs. Hence, the consequences of the actions for the allocations are more observable in IDAM than in IDAM-NC.

Table 8: Random effects probit estimation of probability of playing Truthful Criterion

| | (1) Truthful till reached DA | (2) Truthful IDAM-NC | (3) Truthful IDAM |
|---|---|---|---|
| Truthful in previous round | .37*** | .67*** | .60*** |
| | (.13) | (.15) | (.08) |
| Round | .04*** | .07*** | .08*** |
| | (.01) | (.01) | (.01) |
| Experience per round | | -.05 | .22*** |
| | | (.06) | (.08) |
| Average grade | -.02*** | -.01** | .00 |
| | (.00) | (.00) | (.00) |
| Increase of grade for the true top choice | .01*** | .01*** | .02*** |
| | (.00) | (.00) | (.00) |
| Rank of assigned college in previous round | -.45*** | -.38*** | -.10*** |
| | (.03) | (.04) | (.01) |
| Increase of number of decisions considered for CR1 | -.37*** | -.35*** | .09*** |
| | (.02) | (.04) | (.03) |
| Observations | 1824 | 1824 | 1824 |
| No. of individuals | 96 | 96 | 96 |
| log(likelihood) | -720.84 | -685.08 | -762.47 |

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are clustered on the level of matching groups and are presented in in parentheses

## 5.3 Directed learning within rounds and correction of mistakes

In this subsection we explore the difference between DA and IDAM-NC using the fact that the iterative version allows subjects to change the strategy within a round. First, in line with the evidence of positive correlation of the grade with the probability of playing truthfully, we concentrate on subjects who play non-truthfully in the systematic manner, namely skipping the choices where they expect to be rejected (with low corresponding grades or popular choices). We refer to it as the "skipping strategy." This deviation from the truthful strategy is the most robust deviation in DA both in experiments and in the field.[17] Subjects use skipping strategies because

[17]Skipping strategy is in line with the district school bias in DA experiments, for details see a recent survey by Hakimov and Kübler (2017), skipping of choices in Echenique et al. (2016), and evidence of Hassidim et al.

they want to avoid unnecessary rejections, or because they think that rejections at early steps of the mechanism might hamper their chances at subsequent steps. We are unable to identify the reasoning, but in both cases the skipping strategy implies fast acceptance: in IDAM-NC it means acceptance at early steps of the mechanism, while in DA it means acceptance to the top submitted choices. A skipping strategy is "partial" in the sense that it does not include all possible sequences of choices/iterations, but is based on the belief that the outcome will be one among $X$ first-choices in iterative mechanisms or top-ranked in DA. Therefore, in the iterative mechanisms there is no plan for what to do if this belief is incorrect, and in DA there is no clear reasoning behind the ranking below those $X$ elements. In iterative mechanisms, however, after a subject receives feedback that none of the $X$ choices were realized, she is allowed to re-strategize. She may plan for further skipping, in which case she may also have to re-strategize again later on. Every time the expectation of being accepted in a certain set of universities is not matched, a subject might update her beliefs regarding how good the skipping strategy is. In this respect there is an asymmetry between the iterative mechanisms and DA: there may be multiple failures of the skipping strategy in one round in the iterative mechanism and only one in DA. Thus, the probability that subjects will be "disappointed enough" to abandon the skipping strategy is higher in the iterative mechanisms. We think of the truthful strategy as a strategy that is either played by subjects who understand the mechanism well, or by subjects who fail to find a better strategy for how to outsmart the other players. In a sense it is a default strategy for the latter subjects. Therefore, if a subject abandons the skipping strategy, she is more likely to play the truthful strategy. While we are unable to test this explanation explicitly in the data, we argue that the switch from skipping to the straightforward choices within a round might point to the subjects who realized the failure of the skipping strategy during that round. They would thus be more likely to play the truthful strategy in the next round. In IDAM-NC, the rejection from choices in the first steps provides feedback that the skipping strategy might not work in the intended way, as there is a mismatch of the belief in being matched fast and the outcome. This mismatch is realized within the round and might make subjects reconsider the skipping strategy. The subject might therefore switch to the strategy of being straightforward within the round and apply to the skipped choices after experiencing one or several rejections. If so, this deviation from the skipping strategy within a round, should increase the probability of playing truthfully in the next round, due to the fact that the round taught the subject that the skipping strategy did not work the way he intended. Note that in DA the feedback on the failure of the skipping strategy is not realized within a round, and the return to a straightforward reporting of the colleges after

(2017), Artemov et al. [2017].

some skips in the rank order list does not indicate the disappointment in the skipping strategy. Potentially, from the feedback of being allocated in the choice listed below the skipped one in DA subjects might restore the same feedback as in the iterative mechanism. However it is very demanding and require the backwards counterfactual thinking, while in sequential mechanism it is immediate, at the step of turning to straightforward strategy.

We are unable to test this hypothesis in the data, as in the data we cannot precisely identify the subjects who used the skipping strategy and then disappointed in it. We argue that the switch from skipping to the straightforward choices within a round might point to the subjects who realized the failure of the skipping strategy during that round. If the argument above is correct, these subjects should be more likely to play truthfully in the following rounds relative to those who never abandoned the skipping and was assigned in one of the expected choices. Thus, we construct a variable "the proportion of consecutive straightforward decisions before last within a round." It is equal to the number of consecutive straightforward choices till the last by a subject in a round, divided by the total number of choices made by a subject within this round. For IDAM-NC, straightforward choice would mean that the choice of college at a particular step was the best college, among the colleges the subjects had not applied to yet in the previous steps of that round. At the first step, in order to be straightforward, the choice must be the true best college. If the choice at the first step was not straightforward, then all subsequent choices will not be straightforward until the subject applies to the true best college. For the rounds of truthful play, this variable is always equal to 100%. If subjects used the skipping strategy and it was successful, this variable is equal to 0%. Intermediate values of this variable point to subjects who started the round by using the skipping strategy, and went back to the skipped choices later on. Moreover these subjects were truthful in all subsequent choices until they were accepted. Thus, the higher the proportion of the straightforward decisions is, the more likely the subject realized the failure of the skipping strategy early in the round.[18] This variable was constructed for DA, taking the similar "iterative reasoning" for the reporting of the top, second, third choice and so on one after another. In order to be straightforward, the top choice must be truthful, the second choice must report the best college among all, except of reported as a top choice. The x-th reported choice must report the best college among the subset of those not reported in the first till the (x-1)th choices. For IDAM, straightforward choice means something slightly different, as subjects can only choose from a subset of colleges that have a tentative cutoff lower

[18]Note that earlier switches (after only a few rejections) are more likely to signal the disappointment is the skipping strategy than later switches, where the return to the skipped choices might be just mechanical, due to the rejections from other choices. That is why we use a continuous measure of proportion, and not binary measure of presence of the straightforward choice in a round.

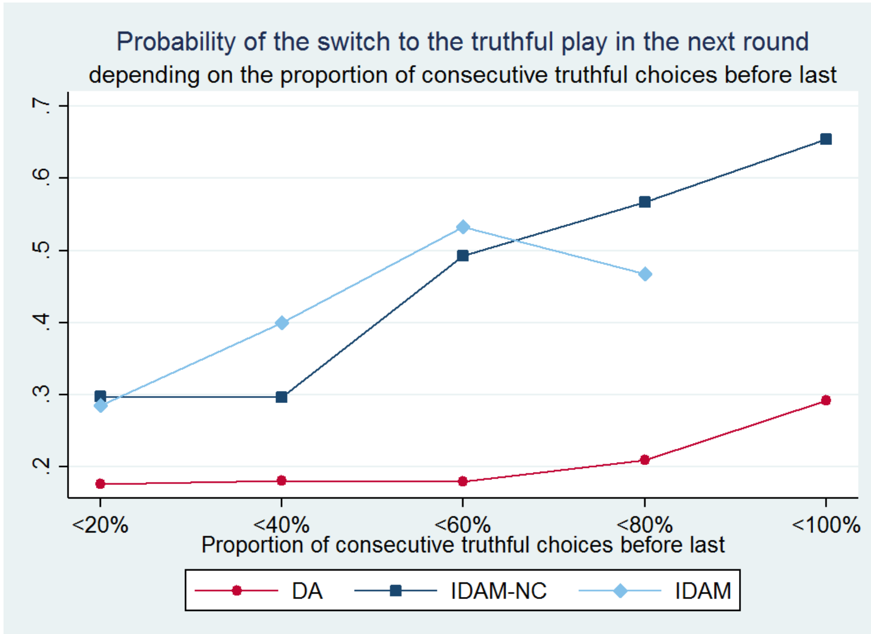


Figure 6: Proportion of switches of strategies from non -truthful till to truthful depending on proportion of consecutive straightforward choices before last in a previous round

than the subject's corresponding score. Thus, the comparison of IDAM with other treatments should be made with caution. We restrict our attention to the rounds in which a subject did not play truthfully. We further restrict our sample to the case when a subject was not assigned to the reported top choice, as in these rounds the proportion of straightforward choices is always zero, and no learning of the failure of the skipping strategy is possible.

Figure 6 presents the probability of the subject switching from non-truthful to the truthful strategy for the restricted sample (non-truthful behavior in the round of interest and assignment in lower than the top reported choice).

The probability of switching to the truthful strategy in the next round increases with the proportion of consecutive truthful choices before the last in the previous round in IDAM and IDAM-NC. This observation is in line with the argument that the within-round "return" of subjects to the truthful (previously skipped) choices can indicate the disappointment in the skipping strategy. For the statistical analysis, consider results of the random effect probit regression of the probability of switching to the truthful strategy by treatments. In line with graphical evidence from Figure 6, the proportion of truthful choices is significant in IDAM and IDAM-NC, and not DA, controlling for other variables.

We therefore cannot reject the hypothesis that the benefit of iterative mechanisms comes from the feedback these mechanisms provide within a round: information on the outcome of

Table 9: Proportion of switches to CR1 strategy depending on the proportion of straightforward choices in the previous round

| | (1) Switch in DA | (2) Switch in IDAM-NC | (3) Switch in IDAM |
|---|---|---|---|
| Rank of assigned college in previous round | -.002 | -.04*** | -.05** |
| | (.01) | (.01) | (.01) |
| Prop.of straightf. consec. ch. before last in prev. round | .10 | .47*** | .29** |
| | (.05) | (.09) | (.14) |
| Round | .004 | .02*** | .02*** |
| | (.003) | (.004) | (.005) |
| Increase of grade for the true top choice | .003*** | .005*** | .01*** |
| | (.001) | (.001) | (.001) |
| Observations | 488 | 479 | 349 |
| log(likelihood) | -245.05 | -199.36 | -194.28 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors in parentheses. Sample restriction: only if non truthful in previous round, and not assigned to the top submitted choice, as then variable of interest is always 0.

the previous actions, own or of the other players, in iterative mechanisms teaches subjects that the deviation of truthful behavior does not work in an intended way. We believe the argument extends to the other comparisons of iterative and direct mechanisms.

For instance, our argument might be applied to explain the benefit of the English auction relative to the second price sealed-bid auction. The better performance of the English auction may not come from the fact that players understand it better, but from the fact that during the English auction players quickly realize that a deviating strategy, for instance overbidding, does not work. Imagine, for example, a player who decides to bid $1,100 given her true valuation of the objects is $1,000. She does so as she expects that this would lead to a higher probability of winning while the price is still below $1,000. Under the second price sealed-bid auction she would send a bid of $1,100. Under the English auction she starts the auction with the intention of keeping the hand raised until the price reaches $1,100. If after some time she observes other players holding their hands up at a price higher than $1000, however, it becomes obvious to her that the strategy did not work as intended, since the price of the object became higher than she expected. Once this is realized, she would then immediately put her hand down. Thus, we can expect almost immediate learning that overbidding does not work.

The better performance of the (sequential) Ausubel auction relative to the multiple unit Vickrey auction and the benefit of the uniform price multiple unit English auction relative to the sealed-bid version can be explained by a similar argument. In fact it is evident from Kagel and Levin (2001, 2009) that the advantage of the dynamic auctions comes from the information about dropouts, and not from the fact that it is sequential. At first sight this might come as a contradiction to our results, but in fact it supports our explanation: in order to outperform the direct counterpart, the sequential mechanism must provide intermediate feedback on the actions of other players. The Ausubel and multiple unit uniform price English auctions without dropout information as tested by the authors is not truly sequential in this sense: it is just a different framing for bidding in the sealed-bid auction. The analogue of it in our case would be an implementation of DA in which subjects are sequentially asked for the top, second, and other choices. These sequentially elicited choices are then transformed into rank-ordered lists to be used in DA. That is, at no point would the participants receive any feedback on the feasibility of their choices. We conjecture that this mechanism would not improve upon DA.

Another example of an iterative mechanism that performs better in experiments is the sequential Serial Dictatorship (SD), when compared to the static SD [Li, 2015]. In fact, SD is a simple, direct mechanism, and deviations from truthful reporting are not observed very often. The typical manipulation is done by subjects who have a lower priority. These subjects manipulate submitted rank-ordered lists of objects, putting at the top of the list the objects that they guess would be the best among those available at the point of their turn. It is closely related to the skipping strategy in DA. In the sequential SD subjects take turns and just choose the object to keep, once their turn is realized, without submitting any rankings beforehand. A subject who at the beginning of Sequential SD has a similar strategy to that of skipping in mind and who guesses the object correctly would not learn anything, but if there is a better object available in the choice set and the moment of their decision, she would learn that the strategy is wrong and will thus choose the better object.

# 6 Discussion

Our findings demonstrate the benefits of the IDAM mechanism, especially with the provision of intermediate feedback about the cutoff grades, when compared to the standard student-proposing DA. This is another evidence of better performance of the sequential mechanisms relatively to the direct counterpart, despite theoretical predictions. The comparison between the IDAM and DA mechanisms is somewhat similar to previous comparisons between dynamic and sealed-bid auctions. The disparity in behavior in strategically equivalent second-price sealed-bid auctions and English auctions is well documented in the experimental literature (see for instance Kagel et al. [1987]). Li [2015] presented one possible explanation for the difference in behavior observed in the lab: While the English auction is *obviously strategy-proof*, the second price sealed-bid auction is not. That, however, cannot be used to explain our results between IDAM and DA (neither can the results of Kagel and Levin [2001] and Kagel and Levin [2009]). We explain the superior performance of the sequential mechanism by the feedback these mechanisms provide on a strategy within a round that teaches subjects that the deviating strategies from the truthful equilibrium do not work as they are intended to. Thus the presence of intermediate feedback is crucial for the superior performance of the sequential mechanisms relative to the direct counterparts. This intuition is in line with the results of Kagel and Levin [2009], where only the Ausubel auction with the provision of drop out prices outperforms the multiple-item Vickrey auction. The Ausubel auction without the provision of drop out prices contains no feedback on the strategies of players, and thus differs from the sealed-bid auction only in a frame of the strategy.

Thus, our paper shows that the static versus dynamic disparity between theory and behavior, which is present in the auction literature, is also present in matching and therefore opens a new platform for the development and research of iterative matching mechanisms, despite their inferior incentive properties. Given the popularity of DA, these findings might make policymakers reconsider DA as the optimal mechanism for implementing matching procedures, where stability is put as the main criterion of interest. On one hand, DA is easier to implement, as it requires only a one-shot submission of preferences, while the IDAM mechanism requires a more intensive use of technology, such as access to the internet and perhaps commitment on the part of participants to stay online for some time. Given the rapid development of technologies, in most countries access to the internet, especially among those who aim to enter a university, is no longer a binding factor, and we thus argue that the potential benefits of the iterative version are higher than the costs. The mere fact that iterative mechanisms are being used to match millions of students to universities in Brazil and China indicates that these constraints are not extremely limiting.

We expect especially high advantages of using mechanisms such as the IDAM in large markets with a high number of universities (this argument is in line with the observation that our surveyed attempts of using the iterative mechanisms are observed in massive markets like those in Brazil and Inner Mongolia). In these markets, where there is a high number of universities or schools, real-life use of DA would require a dramatic constraint of the length of the list when compared to the total number of universities and would thus lose its main benefit, i.e., strategy-proofness. In the IDAM mechanism, the constraint comes as the number of steps.[19] However, in reality, we argue that the mechanism should be implemented with as many steps as possible.[20]Although one

[19]In the online Appendix we describe an additional series of experiments in which constrained versions of IDAM and DA (restricted number of periods in IDAM and restricted length of the ranking in DA) are compared. There we show that the benefits of the IDAM mechanism over DA also takes place when using the restricted versions.

[20]Note that the mechanism implemented in Inner Mongolia is run in almost continuous time, and thus the number of steps could be counted as unconstrained.

could argue that requiring multiple interactions with the mechanism may pose a significant cost over the students, we believe that modern technologies can provide an interactive feedback, which would allow participants to follow the straightforward strategy by spending a very short period of time in each step and thus simplifying the implementation of the almost unconstrained IDAM. In Brazil, for example students are currently able to update their choices via a free cellphone "app."

Moreover, in markets with a large number of universities, the feedback of the iterative mechanism is more valuable than in small markets. This is due to the fact that in big markets participants must be aware of what universities are within their reach, and providing real-time feedback with the updated information about the cutoff grades serves as a very valuable tool.

Lastly, we would like to mention that one important factor which may influence the success of iterative mechanisms such as IDAM, would be the exact way in which the interaction and design of the feedback are implemented. Although these features are beyond this study, we argue that they are of significant importance, and should thus be tested before the implementation with real participant by methods such as focus groups and field testing.

# References


Atila Abdulkadiroğlu, Parag A. Pathak, and Alvin E. Roth. Strategy-proofness versus efficiency in matching with indifferences: redesigning the nyc high school match. *American Economic Review*, 99(5):1954–78, 2009. 2

Atila Abdulkadiroğlu, Parag Pathak, Alvin E Roth, and Tayfun Sönmez. Changing the boston school choice mechanism. NBER Working Paper, 2006. 2

G. Artemov, Y-K. Che, and Y. He. Strategic "mistakes": Implications for market design research. Technical report, 2017. 17

Michel Balinski and Tayfun Sönmez. A tale of two mechanisms: student placement. *Journal of Economic theory*, 84(1):73–94, 1999. 1, 2

Péter Biró. University admission practices – hungary. MiP Country Profile 5., 2012. 2

Inacio Bo and Rustamdjan Hakimov. The iterative deferred acceptance mechanism. Working Paper, 2016. 1, 6

Li Chen and Juan Sebastián Pereyra. Self-selection in school choice. ECARES Working Papers, 2015. 2

Yan Chen and Yinghua He. Information acquisition and provision in school choice. Working paper, 2016. 2

Yan Chen and Onur Kesten. Chinese college admissions and school choice reforms: an experimental study. Working Paper, 2015. 2, 1, 10

Yan Chen and Tayfun Sönmez. School choice: an experimental study. *Journal of Economic theory*, 127(1):202–231, 2006. 2, 4.1

Yan Chen, Ming Jiang, Onur Kesten, Stéphane Robin, and Min Zhu. Matching in the large: an experimental study. Working paper, 2015. 10

Tingting Ding and Andrew Schotter. Learning and mechanism design: an experimental test of school matching mechanisms with intergenerational advice. Working Paper, 2015. 2, 10

Tingting Ding and Andrew Schotter. Matching and chatting: An experimental study of the impact of network. *Games and Economic Behavior*, forthcoming. 2

UM Dur, Robert G Hammond, and Thayer Morrill. Identifying the harm of manipulable school-choice mechanisms. *American Economic Journal: Economic Policy*, forthcoming. 2

Federico Echenique, Alistair J Wilson, and Leeat Yariv. Clearinghouses for two-sided matching: an experimental study. *Quantitative Economics*, 7(2):449–492, 2016. 2

Urs Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2):171–178, 2007. 4.1

Binglin Gong and Yingzhi Liang. A dynamic college admission mechanism in inner mongolia: Theory and experiment. Working paper, 2016. 2, 10

Ben Greiner et al. *The online recruitment system ORSEE: a guide for the organization of experiments in economics*. Max-Planck-Inst. for Research into Economic Systems, Strategic Interaction Group, 2003. 4.1

Pablo Guillén and Rustamdjan Hakimov. How to get truthful reporting in matching markets: A field experiment. WZB Discussion Paper, 2015. 2

Pablo Guillen and Alexander Hing. Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism. *European Economic Review*, 70:178–185, 2014. 2

Rustamdjan Hakimov and Onur Kesten. The equitable top trading cycles mechanism for school choice. WZB discussion paper, 2014. 2, 1

Avinatan Hassidim, Deborah Marciano-Romm, Assaf Romm, and Ran I Shorrer. Strategic behavior in a strategy-proof environment. Working paper, 2015. 2, 4.2.2

John Kagel and Dan Levin. Auctions: A survey of experimental research. 2014. 4

John H Kagel and Dan Levin. Behavior in multi-unit demand auctions: Experiments with uniform price and dynamic vickrey auctions. *Econometrica*, 69(2):413–454, 2001. 2, 6

John H Kagel and Dan Levin. Implementing efficient multi-object auction institutions: An experimental study of the performance of boundedly rational agents. *Games and Economic Behavior*, 66(1):221–237, 2009. 2, 6

John H Kagel, Ronald M Harstad, and Dan Levin. Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica: Journal of the Econometric Society*, pages 1275–1304, 1987. 6

Flip Klijn, Joana Pais, and Marc Vorsatz. Static versus dynamic deferred acceptance in school choice: A laboratory experiment. Working paper, 2016. 2, 3, 10

Shengwu Li. Obviously strategy-proof mechanisms. *Working paper, Stanford University*, 2015. 1, 5.3, 6

Joana Pais and Ágnes Pintér. School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior*, 64(1):303–328, 2008. 2

Joana Pais, Ágnes Pintér, and Róbert F Veszteg. College admissions and the role of information: an expermimental study. *International Economic Review*, 52(3):713–737, 2011. 2

Parag A Pathak and Tayfun Sönmez. Leveling the playing field: Sincere and sophisticated players in the boston mechanism. *American Economic Review*, 98(4):1636–1652, 2008. 2

Parag A Pathak and Tayfun Sönmez. School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103 (1):80–106, 2013. 2

Alex Rees-Jones. Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behavior*, 2017. 2

Ran I Shorrer and Sándor Sóvágó. Obvious mistakes in a strategically simple college-admissions environment. Technical report, Working paper, 2017. 2

Min Zhu. Experience transmission: Truth-telling adoption in matching. Working paper, 2015. 2, 1, 10

# Appendix

## Instructions of the experiment

Welcome! This is an experiment about decision-making. You and the other participants in the experiment will participate in a situation where you have to make a number of choices. In this situation, you can earn money that will be paid out to you in cash at the end of the experiment. How much you earn depends on the decisions that you and the other participants in the experiment make.

These instructions describe the situation in which you have to make a decision. The instructions are identical for all participants in the experiment. It is important that you read the instructions carefully so that you understand the decision-making problem well. If something is unclear to you while reading, or if you have other questions, please let us know by raising your hand. We will then answer your questions individually.

Please do not, under any circumstances, ask your question(s) aloud. You are not permitted to give information of any kind to the other participants. You are also not permitted to speak to other participants at any time throughout the experiment. Whenever you have a question, please raise your hand and we will come to you and answer it. If you break these rules, we may have to terminate the experiment.

Once everyone has read the instructions and there are no further questions, we will conduct a short quiz where each of you will complete some tasks on your own. We will walk around, look over your answers, and solve any remaining comprehension problems. The only purpose of the quiz is to ensure that you thoroughly understand the crucial details of the decision-making problem.

Your anonymity and the anonymity of the other participants is guaranteed throughout the entire experiment. You will neither learn about the identity of the other participants, nor will they learn your identity.

**General description**

This experiment is about students who are trying to enter a university. The 24 participants in the room are grouped into three groups of eight persons each. These eight participants represent students competing for university seats. You will compete with the same seven participants for the whole experiment. The experiment consists of 20 independent decisions, which represent the students' admission processes. At the end of each round every student will receive at most one seat in one of the universities or will remain unassigned. In the end of the experiment one round will be randomly drawn to determine your payment.

There are eight universities that differ in quality and specialization. Each university has only one seat. Universities admit students based on their final exam grade. Each student has a grade for math and a grade for language. Universities M1, M2, M3 accept based on the math grade only. Universities L1, L2, L3 accept based on the language grade only. Universities H1 and H2 accept based on the average grade of math and language.

## Instructions for DA

Your payoff depends on a seat of the university you will be assigned to. In order to get a seat at any university you will have to participate in the centralized allocation mechanism. At the beginning of each round you will submit the list of your preferences to the centralized allocation mechanism, and it will use the lists of all the participants in order to determine allocation. All the steps of the allocation described below take place without any further interactions from the students. Thus, the only thing required for the allocation from the students' side is to submit their lists of preferences.

The allocation procedure is implemented in the following way:

1. The mechanism sends applications from all students to the university of their top choice (the one which is stated first in the submitted list sent to the allocation mechanism).

2. Throughout the allocation process, a university can hold no more applications than its number of seats. If a university receives more applications than its capacity then it rejects the students with the lowest relevant score (math grade for M1, M2, M3; language grade for L1, L2, L3, and average grade for H1,H2). The remaining applications are retained.

3. Whenever an applicant is rejected at a university, her application is sent to the next highest university on her submitted list.

4. Whenever a university receives new applications, these applications are considered together with the retained applications for that university. Among the retained and new applications,

those with the lowest relevant grades in excess of the number of the slots are rejected, while the remaining applications are retained.

5. The allocation is finalized when no more applications can be rejected. Each participant is assigned a slot at the university that holds his/her application at the end of the process.

**Example for DA**

Example: In order to understand the mechanism better, let us go through an example together. If you have any questions about any step of the allocation procedure please feel free to ask at any point. There are six students (ID numbers from 1 to 6) on the market, and three universities (University M1, University L1, and University H1) with two seats in each university. Students have the following grades in their exams:

| | Student1 | Student2 | Student3 | Student4 | Student5 | Student6 |
|---|---|---|---|---|---|---|
| Math | 80 | 90 | 60 | 90 | 70 | 40 |
| Language | 50 | 20 | 80 | 30 | 76 | 82 |
| Average | 65 | 55 | 70 | 60 | 73 | 61 |

University M1 ranks students based on the math grade only, University L1 grades students based on the language grade only, and university H1 ranks students based on the average of the two grades. Students submitted the following school rankings in their decision sheets:

| Student ID | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Top choice | L1 | H1 | M1 | H1 | H1 | M1 |
| Middle choice | H1 | M1 | H1 | L1 | M1 | H1 |
| Last choice | M1 | L1 | L1 | M1 | L1 | L1 |

This allocation method consists of the following steps:

Step 1.

Students 3 and 6 apply for a seat at M1. University M1 has two seats available for allocation and two applicants, thus students 3 and 6 are retained at University M1.

Student 1 applies to University L1. University L1 has two seats and only one applicant, thus student 1 is retained in University L1.

Students 2, 4, and 5 apply for University H1, but it has only two seats available for allocation, thus one of the applicants must be rejected. University H1 ranks students based on the average grade for math and language: student 2 has average grade of 55, student 4 has 60 and student 5 has 73. Among the applicants, student 2 has the lowest average grade, thus student 2 is rejected, and students 4 and 5 are retained at University H1.

| | Retained students in the beginning of the round | Applications of the step | Rejected students |
|---|---|---|---|
| University M1 | - | 3,6 | - |
| University L1 | - | 1 | - |
| University H1 | - | 2,4,5 | 2 |

Step 2.

Student 2 is the only student who was rejected in the previous step. She applies to her second choice – University M1. Now University M1 considers student 2 together with the retained students who applied to University M1 in the previous step – students 3 and 6. So the University has three applications for two seats, thus one of the applicants must be rejected. University M1 ranks students based on the math grade: student 2 has a math grade of 90, student 3 has 60 and student 6 has 40. Student 6 has the lowest math grade among the applicants, thus student 6 is rejected from University M1, while students 2 and 3 are retained.

| | Retained students in the beginning of the round | Applications of the step | Rejected students |
|---|---|---|---|
| University M1 | 3,6 | 2 | 6 |
| University L1 | 1 | - | - |
| University H1 | 4,5 | - | - |

Step 3.

Student 6 applies to University H1. So the University has three applications for two seats, thus one of the applicants must be rejected. University H1 ranks students based on the average grade: student 4 has the average grade of 60, student 5 has 73 and student 6 has 61. Student 4 has the lowest average grade among applicants, thus he is rejected from University H1.

| | Retained students in the beginning of the round | Applications of the step | Rejected students |
|---|---|---|---|
| University M1 | 2,3 | - | - |
| University L1 | 1 | - | - |
| University H1 | 4,5 | 6 | 4 |

Step 4.

Student 4 applies for University L1. Thus, there are two applications for two seats at University L1. No one is rejected. All current retained allocations are finalized.

| | Retained students in the beginning of the round | Applications of the step | Rejected students |
|---|---|---|---|
| University M1 | 2,3 | - | - |
| University L1 | 1 | 4 | - |
| University H1 | 5,6 | - | - |

Thus, the final allocation looks as follows: University M1 – students 2, 3; University L1 – students 1, 4; University H1– students 5, 6.

## Instructions for IDAM

Your payoff depends on a seat of the university you will be assigned to. In order to get a seat at any university you will have to participate in the centralized allocation mechanism.The allocation procedure is implemented in the following way:

1. All students apply to one of the universities.

a. Throughout the allocation process, a university can hold no more applications than its number of seats. If a university receives more applications than its capacity then it rejects the students with the lowest relevant score (math grade for M1, M2, M3; language grade for L1, L2, L3, and average grade for H1,H2). The remaining applications are retained.

2. Each student is informed about whether her application was rejected or retained. Moreover, the minimum corresponding grades of retained students of all universities are publicly announced. If an applicant is rejected at a university, she can send the application to any other university. If an applicant is retained at a university, she is not active at this step.

3. Whenever a university receives new applications, these applications are considered together with the retained applications for that university. Among the retained and new applications, those with the lowest relevant grades in excess of the number of the slots are rejected, while the remaining applications are retained. All students see the result of the step. Each university publishes the minimum corresponding grade of the retained students. If an applicant is rejected at a university, she can send the application to any other university. If an applicant is retained at any university, she is not active at this step.

4. Step 3 is repeated until allocation is finalized. The allocation is finalized when no more applications can be rejected. Each participant is assigned a slot at the university that holds his/her application at the end of the process.

**Example for IDAM**

In order to understand the mechanism better, let us go through an example together. If you have any questions about any step of the allocation procedure please feel free to ask at any point.

There are 6 students (ID numbers from 1 to 6) on the market, and three universities (University M1, University L1, and University H1) with two seats in each university. Students have the following grades in their exams:

| | Student1 | Student2 | Student3 | Student4 | Student5 | Student6 |
|---|---|---|---|---|---|---|
| Math | 80 | 90 | 60 | 90 | 70 | 40 |
| Language | 50 | 20 | 80 | 30 | 76 | 82 |
| Average | 65 | 55 | 70 | 60 | 73 | 61 |

University M1 ranks students based on the math grade only, University L1 grades students based on the language grade only and university H1 ranks students based on the average of the two grades.

This allocation method consists of the following steps:

Step 1.

Students took the following decisions about their application: Students 3 and 6 apply to M1, student 1 applies to L1, and students 2, 4 and 5 apply to H1.

Students 3 and 6 apply for a seat at M1. University M1 has two seats available for allocation and twp applicants, thus students 3 and 6 are retained at University M1.

Student 1 applies to University L1. University L1 has two seats and only one applicant, thus student 1 is retained in University L1.

Students 2, 4 and 5 apply for University H1, but it has only two seats available for allocation, thus one of the applicants must be rejected. University H1 ranks students based on the average grade for math and language: student 2 has average grade of 55, student 4 has 60 and student 5 has 73. Among the applicants, student 2 has the lowest average grade, thus student 2 is rejected, and students 4 and 5 are retained at University H1.

| | Retained students in the beginning of the step | Applications of the step | Rejected applications | Minimum accepted grade |
|---|---|---|---|---|
| University M1 | - | 3,6 | - | 40 |
| University L1 | - | 1 | - | 0 |
| University H1 | - | 2,4,5 | 2 | 60 |

Note, that if a university has a free seat the minimum accepted cutoff grade is zero.

Step 2.

Student 2 is the only student who was rejected in the previous step, thus she is the only one who is active at this step.

She decides to apply to University M1.

Now University M1 considers student 2 together with the retained students who applied to University M1 in the previous step – students 3 and 6. So the University has three applications for two slots, thus one of the applicants must be rejected. University M1 ranks students based on the math grade: student 2 has a math grade of 90, student 3 has 60 and student 6 has 40. Student 6 has the lowest grade among the applicants, thus student 6 is rejected from University M1, while students 3 and 2 are retained.

| | Retained students in the beginning of the step | Applications of the step | Rejected applications | Minimum accepted grade |
|---|---|---|---|---|
| University M1 | 3,6 | 2 | 6 | 60 |
| University L1 | 1 | - | - | 0 |
| University H1 | 4,5 | - | - | 60 |

Step 3.

Student 6 is the only student who was rejected in the previous step, thus she is the only one who is active at this step.

Student 6 decides to apply to university H1.

University H1 considers student 6 together with the retained students – students 4 and 5. So the university has three applications for two seats, thus one of the applicants must be rejected. University H1 ranks students based on the average grade: student 4 has an average grade of 60, student 5 has 73 and student 6 has 61. Student 4 has the lowest average grade among the applicants, thus he is rejected from University H1.

| | Retained students in the beginning of the step | Applications of the step | Rejected applications | Minimum accepted grade |
|---|---|---|---|---|
| University M1 | 2,3 | - | - | 60 |
| University L1 | 1 | - | - | 0 |
| University H1 | 4,5 | 6 | 4 | 61 |

Step 4.

Student 4 is the only student who was rejected in the previous step, thus she is the only one who is active at this step.

Student 4 decides to apply for University L1.

University H1 considers student 4 together with the retained students – student1. Thus, there are two applications for two seats at University L1. No one is rejected. All current retained allocations are finalized.

| | Retained students in the beginning of the step | Applications of the step | Rejected applications | Minimum accepted grade |
|---|---|---|---|---|
| University M1 | 2,3 | - | - | 60 |
| University L1 | 1 | 4 | - | 30 |
| University H1 | 4,5 | 6 | 4 | 61 |

Thus the final allocation looks as follows: University M1 – students 2, 3; University L1 – students 1, 4; University H1 – students 5, 6.

## Instructions for IDAM-NC

Your payoff depends on a seat of the university you will be assigned to. In order to get a seat at any university you will have to participate in the centralized allocation mechanism.The allocation procedure is implemented in the following way:

1. All students apply to one of the universities.

a. Throughout the allocation process, a university can hold no more applications than its number of seats. If a university receives more applications than its capacity then it rejects the students with the lowest relevant score (math grade for M1, M2, M3; language grade for L1, L2, L3, and average grade for H1,H2). The remaining applications are retained.

2. Each student is informed about whether her application was rejected or retained. If an applicant is rejected at a university, she can send the application to any other university. If an applicant is retained at a university, she is not active at this step.

3. Whenever a university receives new applications, these applications are considered together with the retained applications for that university. Among the retained and new applications, those with the lowest relevant grades in excess of the number of the slots are rejected, while the remaining applications are retained. All students see the result of the step. If an applicant is rejected at a university, she can send the application to any other university. If an applicant is retained at any university, she is not active at this step.

4. Step 3 is repeated until allocation is finalized. The allocation is finalized when no more applications can be rejected. Each participant is assigned a slot at the university that holds his/her application at the end of the process.

### Example for IDAM-NC

In order to understand the mechanism better, let us go through an example together. If you have any questions about any step of the allocation procedure please feel free to ask at any point.

There are 6 students (ID numbers from 1 to 6) on the market, and three universities (University M1, University L1, and University H1) with two seats in each university. Students have the following grades in their exams:

| | Student1 | Student2 | Student3 | Student4 | Student5 | Student6 |
|---|---|---|---|---|---|---|
| Math | 80 | 90 | 60 | 90 | 70 | 40 |
| Language | 50 | 20 | 80 | 30 | 76 | 82 |
| Average | 65 | 55 | 70 | 60 | 73 | 61 |

University M1 ranks students based on the math grade only, University L1 grades students based on the language grade only and university H1 ranks students based on the average of the two grades.

This allocation method consists of the following steps:

Step 1.

Students took the following decisions about their application: Students 3 and 6 apply to M1, student 1 applies to L1, and students 2, 4 and 5 apply to H1.

Students 3 and 6 apply for a seat at M1. University M1 has two seats available for allocation and twp applicants, thus students 3 and 6 are retained at University M1.

Student 1 applies to University L1. University L1 has two seats and only one applicant, thus student 1 is retained in University L1.

Students 2, 4 and 5 apply for University H1, but it has only two seats available for allocation, thus one of the applicants must be rejected. University H1 ranks students based on the average grade for math and language: student 2 has average grade of 55, student 4 has 60 and student 5 has 73. Among the applicants, student 2 has the lowest average grade, thus student 2 is rejected, and students 4 and 5 are retained at University H1.

| | Retained students in the beginning of the step | Applications of the step | Rejected applications |
|---|---|---|---|
| University M1 | - | 3,6 | - |
| University L1 | - | 1 | - |
| University H1 | - | 2,4,5 | 2 |

Step 2.

Student 2 is the only student who was rejected in the previous step, thus she is the only one who is active at this step.

She decides to apply to University M1.

Now University M1 considers student 2 together with the retained students who applied to University M1 in the previous step – students 3 and 6. So the University has three applications for two slots, thus one of the applicants must be rejected. University M1 ranks students based on the math grade: student 2 has a math grade of 90, student 3 has 60 and student 6 has 40. Student 6 has the lowest grade among the applicants, thus student 6 is rejected from University M1, while students 3 and 2 are retained.

| | Retained students in the beginning of the round | Applications of the step | Rejected applications |
|---|---|---|---|
| University M1 | 3,6 | 2 | 6 |
| University L1 | 1 | - | - |
| University H1 | 4,5 | - | - |

Step 3.

Student 6 is the only student who was rejected in the previous step, thus she is the only one who is active at this step.

Student 6 decides to apply to university H1.

University H1 considers student 6 together with the retained students – students 4 and 5. So the university has three applications for two seats, thus one of the applicants must be rejected. University H1 ranks students based on the average grade: student 4 has an average grade of 60, student 5 has 73 and student 6 has 61. Student 4 has the lowest average grade among the applicants, thus he is rejected from University H1.

| | Retained students in the beginning of the round | Applications of the step | Rejected applications |
|---|---|---|---|
| University M1 | 2,3 | - | - |
| University L1 | 1 | - | - |
| University H1 | 4,5 | 6 | 4 |

Step 4.

Student 4 is the only student who was rejected in the previous step, thus she is the only one who is active at this step.

Student 4 decides to apply for University L1.

University H1 considers student 4 together with the retained students – student1. Thus, there are two applications for two seats at University L1. No one is rejected. All current retained allocations are finalized.

| | Retained students in the beginning of the round | Applications of the step | Rejected applications |
|---|---|---|---|
| University M1 | 2,3 | - | - |
| University L1 | 1 | 4 | - |
| University H1 | 4,5 | 6 | 4 |

Thus the final allocation looks as follows: University M1 – students 2, 3; University L1 – students 1, 4; University H1 – students 5, 6.