Matching and Learning – An Experimental Study

Guillaume Haeringer*
Lan Nguyen†
Silvio Ravaioli‡

September 2, 2021

Abstract

We use a lab experiment to study the patterns and effects of learning in two classic centralized matching mechanisms widely used in school choice and other real-world settings. We focus on the Deferred Acceptance algorithm (DA, strategyproof but not efficient) and contrast it with the Immediate Acceptance algorithm (IA, efficient but not strategyproof). Each matching problem (round) is repeated for several periods: after being informed about the match outcome of the previous period, subjects are asked again the order in which they would apply to the same schools. Each participant experiences multiple rounds of the same mechanism. Our design allows for two types of learning: to coordinate within the same environment (within round) as well as to understand the underlying mechanisms (across rounds). We observe that subjects are more truthful under the DA mechanism and achieve higher payoffs under the IA mechanism. We provide additional evidence to previous work that the majority of the deviations from truth-telling in DA do not affect payoffs. Furthermore, by explicitly analyzing learning, we can confirm that at least some of the participants learn about the optimality of truth-telling, and their departures from it happen primarily when the same environment is repeated. Finally, we find that when learning to coordinate, agents tend to retain their previous strategy when the payoff from this strategy is high, which is suggestive of reinforcement learning.

1 Introduction

Gale and Shapley (1962)'s student-proposing Deferred Acceptance mechanism has become increasingly ubiquitous in school choice to assign students to schools. This is, in part, thanks to its strategyproofness, an attractive property, which, informally in this context, means that it is in the students' best interest to report their true preferences. However, a growing body of experimental literature has recorded large deviations from truth-telling when subjects participate in the Deferred

*Baruch College. Email: guillaume.haeringer@baruch.cuny.edu. [Link to webpage.]
†Department of Economics, Columbia University. Email: tn2304@columbia.edu. [Link to webpage.]
‡Department of Economics, Columbia University. Email: sr3300@columbia.edu. [Link to webpage.]

Acceptance mechanism in laboratory settings. At the same time, there has also been empirical evidence that some of these "mistakes" have little to no effects on payoff (Artemov, Che and He, 2020).

This paper introduces learning into a lab experiment of centralized matching mechanisms, Deferred Acceptance in particular. Our contribution is two-fold. First, we shed light on the nature of "mistakes" that occur in the lab. We find that the majority of these mistakes are payoff-irrelevant as in Artemov, Che and He (2020). We show that payoff-irrelevant mistakes are more common when subjects have had the opportunities to learn to coordinate within the same environment, and they occur despite evidence that subjects have gained an understanding of the strategyproofness of the mechanism. Second, our results provide insights into the learning process of students who receive feedback in centralized matching.

In our laboratory experiment, we adopt a 2x2 design, with two matching mechanisms (between-subjects) and two treatments (within-subjects). The two mechanisms are Gale-Shapley Deferred Acceptance (DA) and Boston Immediate Acceptance (IA) algorithms. While DA is our focus, IA serves as a control. In each experimental session, we use only one of the mechanisms. At the beginning of the session, the rules of the relevant mechanism are clearly explained, accompanied by an example. Each session is divided into two parts for the two different treatments. Part one involves a random environment without school zones, and part two, a designed environment with school zones (where students are informed of their higher chance of being admitted in two out of five schools). Within each part, there are 8 rounds, each consisting of 5 periods. In every period, each participant plays the role of a student who applies to multiple schools. There are 20 students and 5 schools. Participants are informed of their own payoffs for being admitted into the schools and asked for an order in which they would apply to those schools. Within a round, the environment (schools' priorities and students' preferences) is unchanged, and periods differ only in the rankings submitted by participants. Across the rounds, the environment changes, but the mechanism does not.

Focusing on DA, we observe two opposite patterns in truthfulness within a round and across rounds. Across periods within the same round, when the environment remains constant, truthful reporting decreases over time. However, payoffs do not change much throughout the round, and the vast majority of participants obtain the same payoff as if they were fully truthful. In contrast, truthful reporting increases over time across rounds when the environment changes but the same matching mechanism is used. We do not observe the same trend for IA, for which truth-telling is not necessarily a good strategy.

Given the patterns above, we argue that participants learn about the strategyproofness of the DA mechanism throughout the experimental session. This is supported both by a further comparison with the results in the IA sessions and by an examination of the first period of every round when a new environment is generated. Within each round, when subjects are allowed to learn to coordinate within the same environment, the best possible match each of them can get is often easy to predict in later periods. Subsequently, their untruthfulness is mostly payoff-irrelevant and relates to the concepts of robust equilibrium and stable-response strategy introduced by Artemov, Che and He (2020). We also find that a higher payoff leads to a subject being more likely to keep the same strategy in the next period, consistent with reinforcement learning.

This paper contributes to the experimental literature on matching markets, particularly school choice, by exploring the learning process of students who receive feedback in interactive environments. Hakimov and Kubler (2019) offer a systematic review of the relevant literature. Starting from the earliest lab experiments on school choice (Chen and Sonmez, 2006; Pais and Pinter, 2008; Calsamiglia, Haeringer and Klijn, 2010, 2011), the designs put significant emphasis on the comparison of different mechanisms (DA, IA, and TTC), from the perspective of efficiency and stability. The main results on strategyproofness of the mechanisms are systematically aligned with the theoretical predictions, with the lowest proportion of truthful reporting under IA, but efficiency results are mixed. Surprisingly, there is unequivocal evidence of departure from the complete truth-telling, with fractions of truthful reporting under DA ranging between 40% and 80% (on average, larger fractions were registered when the design included fewer schools to rank). Other common results include systematic biases in favor of schools with many seats and schools that belong to the same "zone" as the student (in both cases, perceived with a higher probability of admission). More recent experiments compared traditional and modified mechanisms, such as the parallel mechanism (Chen and Kesten, 2019) and the secure Boston mechanism (Dur, Hammond and Morrill, 2019).

A small number of experimental papers on school choice provide more detailed results on the learning process in this domain, either by testing it directly or by adopting designs in which participants play multiple rounds of matching with the same mechanism. In the vast majority of the studies, we observe a positive trend of reported truthful over time under DA (up to a 30% increase), but with a ceiling around 80% regardless of the characteristics of the environment. Furthermore, the starting level and slope of truthfulness trends over time are negatively correlated with the number of schools in the task, with slower increases (or even decreases) being observed when the environment contains more than four schools (Chen and Kesten, 2019; Bo and Hakimov, 2020).

In contrast with all the previously listed experiments, we focus more explicitly on learning in an interactive environment. As in the previous cases, we let the participants play under the same mechanism for several rounds, allowing them to reach a clearer understanding of how the algorithm operates. In addition, we have an explicit form of feedback, that is represented by the realized match given the (unobservable) schools' priorities, and the combination of own and others' application lists. Each participant submits a list for multiple periods within each environment (experimental round), with the possibility to revise the list at the beginning of each period. This procedure allows the participant to learn about the elements of the environment that are not directly observable. At the same time, the comparison of the effect of own change across periods is confounded by the presence of other human participants, who also can revise their own lists.

Ding and Schotter (2019) also consider learning in an interactive environment, and they address similar research questions, but with a different paradigm. In their experiments, participants repeatedly played DA or IA in a market comprised of five students and three schools (one seat each). Participants were informed about their own preferences and schools' priorities, and groups were randomly re-matched in every round. They observe a significant increase in truthful reporting under DA (from 64% to 77% over 20 rounds) and no increase in IA. Our design differs in many aspects, such as the size of the environment (20 students and 5 schools), the fact that all students can be successfully assigned to a school, the lack of explicit information about school priority, and the absence of re-matching across rounds.

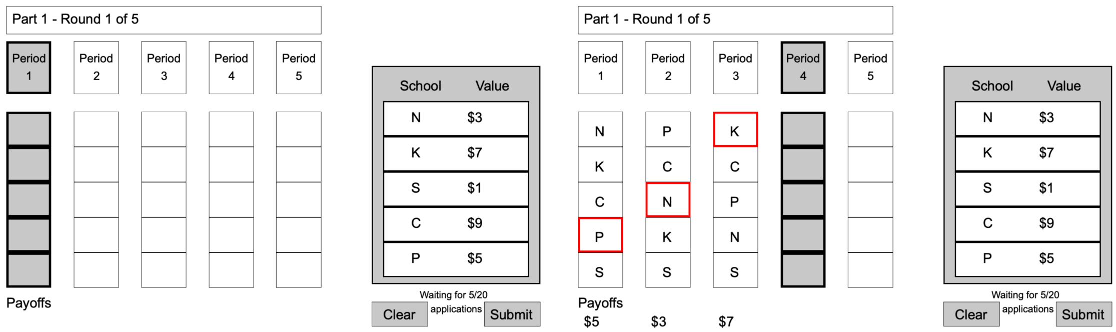
A different group of experimental papers focus on learning in the form of advice, either from a previous participant to the experiment (Ding and Schotter, 2017) or from a third party (Guillen and Hing, 2014). Most of the studies conclude that providing correct advice increases the rates of truthful reporting in strategyproof mechanisms, but the opposite effect is even stronger when wrong advice is provided, with a large number of participants manipulating their preference reports.

The rest of the paper is organized as follows. Section 2 describes the experimental design. Section 3 presents our analysis of the results, and section 4 concludes.

2 Experimental Design

We design our experiment to compare the learning behavior under different assignment mechanisms. We implement a two-by-two design: for each of the two mechanisms, Deferred Acceptance (DA) and Immediate Acceptance (IA), we examine a *random* environment (without school zones) and a *designed* environment (with school zones). The first manipulation occurs between subjects, with each participant being assigned to one mechanism only. The second manipulation occurs within subjects, with the random environment being presented first, followed by the designed environment. Despite the possible concern for the order effect, we prioritize clarity of the task for the participants and frame the second environment as a variation on the first part. Each environment condition is used for eight subsequent rounds, with each round comprised of five periods. In each period, twenty students compete for twenty seats (five schools, four seats each). We now describe the details of the experimental design, which apply to all the conditions unless otherwise specified.

Students' Preferences and Schools' Priorities. Each student receives a random ranking of the schools, with values ranging from \$1 to \$9, with \$2 increments. Subjects are informed that different participants may have different payoff values, and receive detailed information about how the values are generated, with particular emphasis on the fact that the ranking is not independent across students, and on average participants like the schools located at the top of the list more than the schools at the bottom. The ranking of each student is generated by summing, for each school, a common score and a random score. The common score (0 to 12 points, equally spaced)



(a) Beginning of period 1

(b) Beginning of period 4 (same round)

Figure 1: Experimental interface (Part 1)

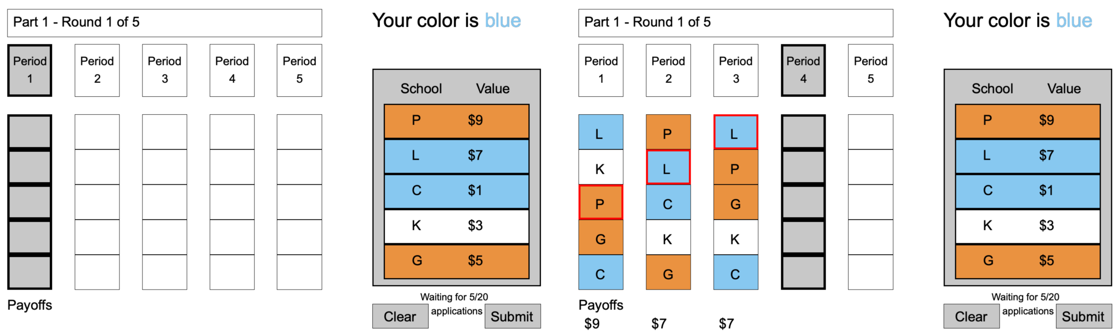
Notes: The experimental interface contains the list of the schools, each with an associated monetary value, and an empty application list that the participant can fill freely by clicking on the schools' names. The screen also contains a shrinking timer that indicates the time available to complete the choice, and two buttons, to submit the application in advance or restart the application list.

is the same for all students and determined by the rank used to display the schools on the screen. The random score (0 to 8 points) is independent across students and across schools, which allows students to have different preferences. The sum of the two scores determines the subjective ranking of the schools, and this ranking is associated with the monetary amounts displayed as values for the seat. Participants can observe their own preferences, but not other students' preferences. They also know that all preferences stay constant within each round but change across rounds. In the designed environment condition, students' zones and schools' zones have no effect on the ranking.

Schools' priorities are generated differently in the *random* environment (without school zones) and in a *designed* environment (with school zones). In both environments, students are not directly informed about schools' priorities, but they are aware that priorities astay constant within each round but change across rounds. In the random environment, each school independently generates a random ranking of the twenty students. In the designed environment, each student receives a color (orange or blue, indicating her own zone, ten students for each color), and the five schools are partitioned into two schools for each color, plus a white (neutral) school. Colors are randomly assigned at the beginning of each round and remain unchanged within the round. For each school, the probability of a student being among the top 4 positions of the priority list depends on her color. If the student's color matches [does not match] the school, she has a 30% probability [10% probability] of being among the top four positions.[1]

School Application Task. Each period is comprised of an *application* phase and a subsequent assignment phase. During the application phase, each student submits an application list, which

[1]A separate random draw determines which students are among the top 4 positions. The relative ranking of those four students is random. Similarly, the relative ranking of the students ranked 5th and below is randomly generated.



(a) Beginning of period 1

(b) Beginning of period 4 (same round)

Figure 2: Experimental interface (Part 2)

Notes: The experimental interface differs from the previous one only because of the “student’s color" (top right of the screen) and the colors associated with the various schools.

indicates in which order they want to apply to schools. Students are not forced to apply to all five schools, although not doing so means that they may end up unassigned (an outcome with value \$0, strictly lower than a seat in any school). The application phase has a limited time (60 seconds in period 1, 30 seconds in all the subsequent periods), but participants can submit their own application lists before the timeout.

In the assignment phase, each participant observes to which school she would be assigned, given other students’ current application lists as well as her own. The assignment is salient, and it is displayed both with the school’s name in the application list and its value below the application list (see Figures 1 and 2). No other information (schools’ priorities or other students’ assignments) is communicated to the participants. The application phase has a fixed time of 10 seconds, after which all the participants move to the next period.

Each round is comprised of five periods, and within a round, there is no change in students’ preferences, schools’ priorities, nor zones (in the designed environment condition). From period 2 to 5, participants can revise the application list, and submit either the previous one or a new one, with no direct cost or benefit from the revision. The task is repeated for eight different rounds for each environment (random and designed).

Procedure. The experiment was run in CELSS (Columbia Experimental Laboratory for Social Sciences, Columbia University, New York, USA) between March and April 2019. The experiment was coded in MATLAB (Release 2018b) using Psychotoolbox 3 (Psychophysics Toolbox Version 3). Seventy-six volunteers (38 in the Boston treatment, and 38 in the Gale-Shapley treatment) were recruited from the students registered in the Marked Design undergraduate course at Columbia University as optional assignment for the class’ requirements.[2] On average, the whole experiment

[2]Crucially, the experiment was run before the instructor discussed the mechanisms for school assignment in class.

took 60 minutes, including instructions. Participants did not receive any monetary incentive for their participation in the study, but they were motivated by the fact that their scores in the task would be posted on the website of the course. Anecdotal evidence suggests that participants were motivated to get a high score and put effort into the task as if they were facing a monetary incentive.

We ran four sessions in total (two sessions for each mechanism). Each session is designed for a group of twenty subjects, and the actual number of participants ranged between eighteen and twenty. When the number was not exactly twenty, the missing participants were replaced by automatic computer programs. Human participants were informed that each program was submitting a list of schools in the first period of each round, but not revising it in the following periods. Instructions were read aloud by the experimenter, and were also provided to each participant as paper printout (see Appendix 6). There are two versions of the instructions (one for each matching mechanism), containing the same information verbatim with the exception of the tailored information on the mechanism being used. The instructions included detailed information about the task, the data generating process used to generate priorities and payoffs, as well as the matching mechanism (described formally first, and illustrated with an example after). In addition to the written instructions, the main projectors of the laboratory were used to illustrate the matching mechanism in the example and to display screenshots of the experimental interface. Clarifying questions, if any, were answered publicly.

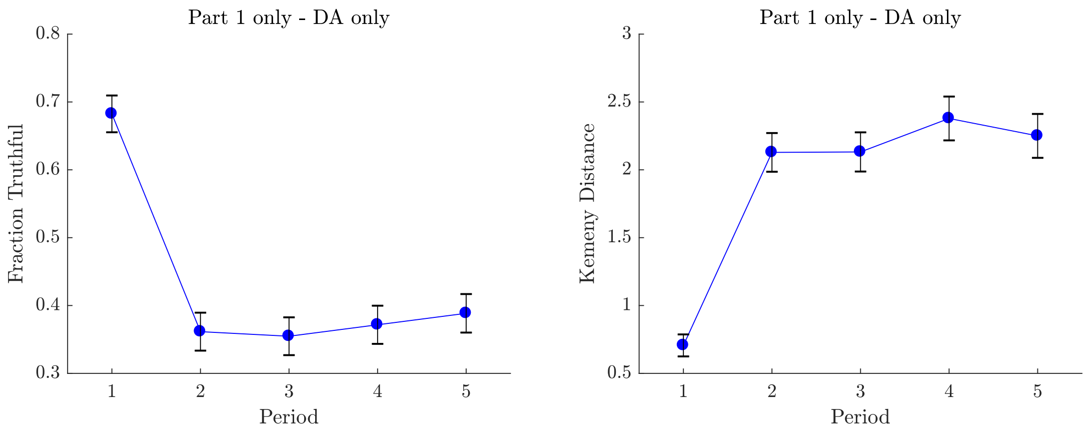
3 Results

We focus our analysis on the results of the Deferred Acceptance (DA) sessions, and contrast it with Immediate Acceptance (IA) when appropriate. We also restrict our attention to the first part of the experiment to avoid the potentially confounding effect of time and school zones in part 2.

3.1 Across Periods within A Round

We find that for all ex-ante measures of (un)truthfulness, that is, measures that are independent of the resulting match, truthful reporting decreases over time within a round. However, payoffs remain relatively stable throughout the round because the majority of untruthful participants behave in such a way that they still obtain the same payoff as under truthful reporting.

Figure 3 includes the plots for two ex-ante (un)truthfulness measures for the five periods, averaged over rounds. Both show that truthful reporting decreases further into the round. As seen in Panel (a), the fraction of participants who are fully truthful (i.e., reporting their exact true preferences) drops sharply from the first to the second period before mostly leveling out. Panel (b) shows are similar story when we move from the aforementioned binary measure to one that also takes into account the intensity of truthful reporting for each participant. The Kemeny distance from the submitted list to the true preference is the number of pairwise comparisons in the submitted list



(a) Fraction of fully truthful participants

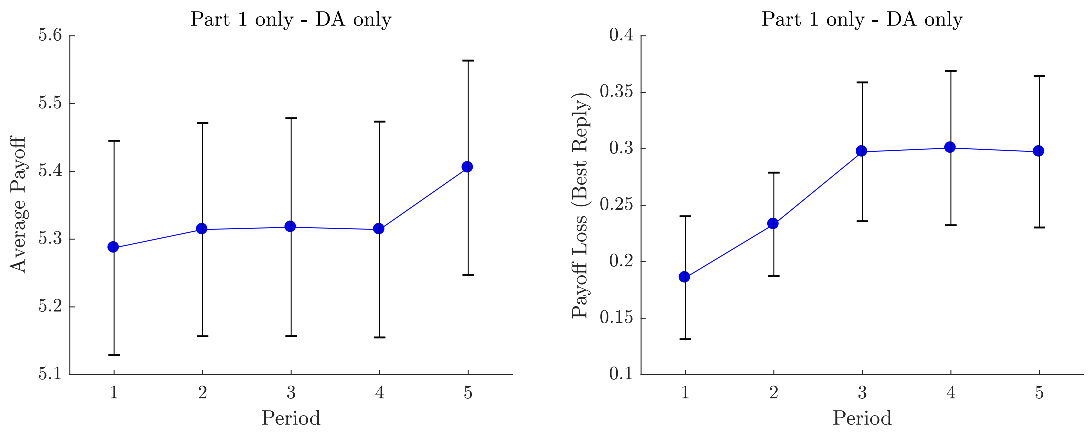
(b) Kemeny distance from submitted ranking to true preference

Figure 3: Across periods - Measures of ex-ante (un)truthfulness

that deviate from the true preference: the higher this measure is, the less truthful the participants are. It increases considerably from close to 0 in the first period to more than 2 in period 2 and then remains between 2 and 2.5 until the end of the round. With five schools, there are a total of ten pairwise comparisons, so during periods 2 to 5, an average participant is untruthful for 20-25% of the pairwise rankings.

Figure 4 plots two different measures of payoff for the five periods, averaged over rounds. Panel (a) demonstrates that average payoffs remain relatively stable across the five periods. Meanwhile, Panel (b) shows the payoff loss for each participant compared to what they would have gotten by best-replying to the actual strategy profile of their opponents. For DA, as is the case here, full truthfulness is a best-reply to any strategy profile of the opponents, so this measure can also be considered as payoff loss compared to truthful reporting. There is a slight increase in payoff loss at the beginning of the round before it levels out. The magnitude of both the average payoff loss (ranging from 0.15 to 0.3) and the increase in payoff loss are small compared to the average payoff (5.3-5.4). This suggests that each participant was very close to best-replying to their opponents' strategies even though they might not be using the dominant strategy (truth-telling).

To further investigate this phenomenon, we divide the participants into three categories within each period: truthful, untruthful with no effect on payoff (i.e. the payoff loss outlined above is zero), and untruthful with effect (i,e, the payoff loss is positive). We find that the vast majority of the 38 participants (between 90-94%) belong to the first two groups for all periods, which explains why although the fraction of full truthfulness decreases, this has little effect on payoff. In line with the pattern in truthful reporting, the most drastic change in composition of participants occur in the second period, where 49% of those who were truthful in the first period switch to untruthful with no effect on payoff (see the transition matrix on the left in Figure 5). While there exist movements



(a) Average Payoff (b) Payoff loss (compared to best reply)

Figure 4: Across periods - Measures of Payoff

Period 1 \ Period 2	Truthful (36%)	N-Truthful No effect (55%)	N-Truthful W/ effect (9%)
Truthful (68%)	45%	49%	6%
N-Truthful No effect (26%)	18%	73%	9%
N-Truthful W/ effect (5%)	18%	38%	44%

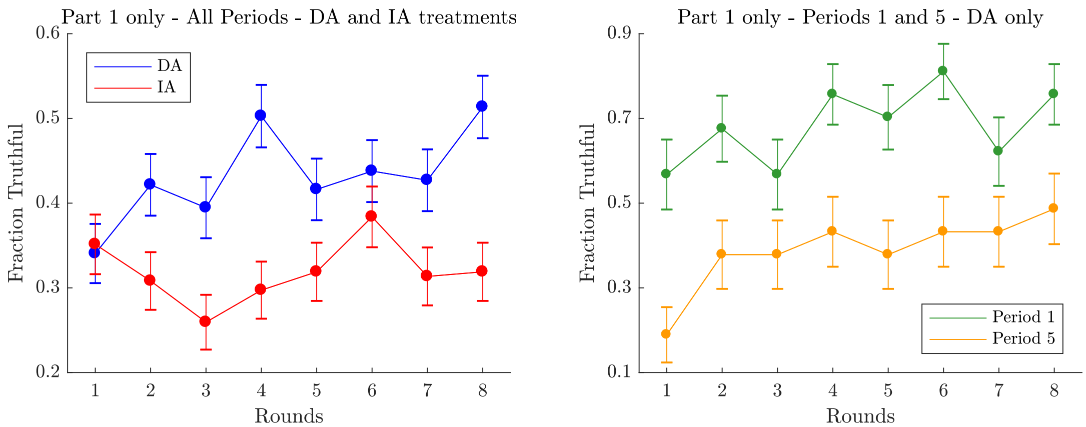
Period 4 \ Period 5	Truthful (39%)	N-Truthful No effect (53%)	N-Truthful W/ effect (8%)
Truthful (37%)	75%	24%	1%
N-Truthful No effect (54%)	16%	75%	9%
N-Truthful W/ effect (9%)	23%	46%	31%

Figure 5: Transition matrices: periods 1 to 2 (left) and periods 4 to 5 (right).

between groups for all the later periods, the composition always stay close to that attained in period 2. As seen in the transition matrix on the right in Figure 5, by the time we reach the transition from period 4 to period 5, a large fraction (75%) of those in the truthful and untruthful unaffected groups stay within the same group. The proportion of participants in the untruthful affected group rises from 5% in period 1 to 8-9% in later periods, which accounts for the slight increase in payoff loss documented above.

3.2 Across Rounds

When we make the comparison across rounds, truthful reporting increases over time during the session for DA. The same trend is not observed for IA (the control), suggesting that (some) participants eventually learn that DA is a strategy-proof mechanism while IA is not.



(a) DA vs. IA - All periods

(b) DA only - Periods 1 and 5

Figure 6: Across rounds - Fraction of truthful participants

Table 3.1: Truthfulness and Payoff - DA vs. IA

	(1) Full Truthfulness	(2) Truthful or Untruthful Unaffected	(3) Payoff	(4) Payoff Loss (vs best reply)
Constant	0.542***	0.905***	5.291***	0.284***
	(0.029)	(0.016)	(0.163)	(0.048)
IA Dummy	-0.057		0.089	1.712***
	(0.042)		(0.236)	(0.138)
Round	0.019***	0.007**	-0.003	-0.023**
	(0.006)	(0.003)	(0.031)	(0.011)
Period	-0.047***	-0.005	0.024	0.029
	(0.009)	(0.005)	(0.050)	(0.019)
IA * Round	-0.013*		0.055	0.010
	(0.008)		(0.044)	(0.027)
IA * Period	0.004		-0.040	-0.134***
	(0.013)		(0.072)	(0.043)
R-squared	0.029	0.004	0.0023	0.1562
N	2960	1480	2960	2960

Notes: Observations are at the session-round-period-participant level. For columns (1), (3), and (4), the sample includes observations from part I (rounds 1-8) of all four experimental sessions. For columns (2), the sample is restricted to the two DA sessions.
The round and period numbers are normalized to 0-7 and 0-4, respectively, so that the constant indicates the average level of truthfulness in the first period of the first round under DA.

While the fractions of truthful participants start out at comparable levels for the first round of both mechanisms, they diverge in later rounds. Panel (a) of Figure 6 shows an upward trend in the fraction of full truthfulness under the DA mechanism, from about 34% in the first round to 51% in the eighth round (within the range of the previous literature). In contrast, the equivalent fraction under the IA mechanism fluctuates between 25 and 40% throughout the session. The same pattern can be seen in the first column of Table 3.1, which presents the results for the regression of an indicator for full truthfulness on a dummy for the IA mechanism, round number (1 to 8), period number (1 to 5), and their interactions. In the first period of the first round, there is no significant difference in the level of truthfulness between the two mechanisms. As the experiment progresses, the probability of truthfulness improves by 1.9 percentage point per round on average (significant at the 1% level) under DA as opposed to the 1.3 percentage point decrease (significant at the 10% level) under IA.

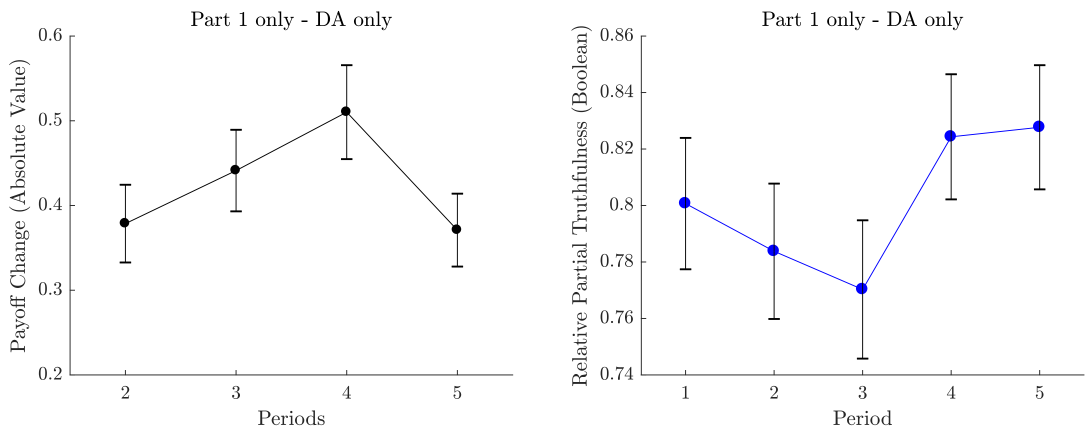
Under DA, there is no significant change in payoff, but payoff loss decreases marginally across rounds (according to columns (3) and (4) of Table 3.1). The decrease in payoff loss is attributable to the modest increase in the fraction of participants who are either truthful or untruthful but unaffected in terms of payoffs. Although both of the aforementioned changes are significant (at the 5% level), the magnitudes are negligible.

3.3 Explaining the Results

3.3.1 Learning about the Mechanism and Payoff-Irrelevant Mistakes

Although participants learn about the optimality of truth-telling under the DA mechanism, when they are allowed to learn to coordinate within the same environment, a large fraction deviate from full truthfulness in a way that is payoff-irrelevant. Two main arguments support the existence of learning about the mechanism over time: when we compare DA to the control IA and when we concentrate our attention on periods in which participants were faced with a new environment. We find that subjects were presumably able to predict their best possible match in later periods within each round, and their untruthfulness relates to the concepts of robust equilibrium and stable-response strategy introduced by Artemov, Che and He (2020).

The comparison between DA and IA points to not only the fact that subjects had some understanding of the strategyproofness of DA but also that at least some of them acquired this over time during the experimental session. As we described in the previous subsection, there is a significant increase in truthful reporting across rounds under DA that is not replicated under IA. Given the random assignment of subjects into the two mechanisms, it is likely that this result is driven by subjects' understanding of the mechanism's underlying property. In theory, this understanding can be obtained immediately after reading the description and example of the mechanism in the instructions. However, we can attribute it to learning during actual participation (with feedback) in the mechanism for at least some participants because there is no significant difference in truthfulness



(a) Across Periods - Absolute changes in payoff (b) Across Periods - Relative partial truthfulness

Figure 7: Payoff Predictability and Observed Strategy

Notes: Relative partial truthfulness is defined as reporting the truthful relative rankings among schools above or equal to the realized match.

between IA and DA in the first round (and the first period of that round, in particular).

Now, let us focus on the first period of each round, when participants had not yet learned to coordinate with each other in the new environment. Panel (a) of Figure 6 demonstrates that the increase in truth-telling across rounds, previously reported for the average over all periods, persists under this restriction. Moreover, we can see that although truthfulness is lower in period 5 than period 1 of every round, this low level in the last period of a round does not carry over the period immediately after it (i.e. period 1 of the next round) when the environment is new. This further confirms that the decrease in truthful reporting we observe across periods within a round can be ascribed to learning to coordinate and does not discount the learning-about-the-mechanism hypothesis. We also confirm that participants who became consistently truthful in the first periods towards the end of the experimental session account for a considerable part of the increase in truthful reporting across rounds. Figure 5.1 shows the fractions of participants (compared to the total) who were truthful in the first period for all 8 rounds, all of the last 7 rounds, all of the last 6 rounds, and so on. Half of of the participants who were truthful in the last round (76% of the total) had been truthful for the last 5 rounds (inclusive of round 8).

The observed pattern in the data provides suggestive evidence of the ease of predicting the best possible match during later periods of each round. We remarked in subsection 3.1 that average payoffs remain stable across periods, but this could be masking large positive and negative changes occurring at the same time for different participants. Panel (a) of Figure 7 shows that this is not the case: the absolute changes in payoff, ranging on average between 0.3 and 0.5 through out the round, are small compared to the increment of \$2 between consecutive prizes and the difference of \$1 between the unmatched outcome and the lowest prize. Note that there is a one-to-one correspondence between payoffs and schools. Therefore, given little change in payoffs and 90-94% of participants receiving as-if-truth-telling payoffs, participants could assume (often correctly) that the best match they could obtain was the same as their previous match.

To examine the subjects' ranking strategies, we define an "ex-post" notion of truthfulness that relies on knowledge of the match. From a participant's point of view when choosing their strategy, it is their best guess of the match, which is unobservable to the experimenter. Hence, we use the realized match, which as argued above was easy to predict. Relative partial truthfulness is, then, a binary measure indicating that a subject reported the truthful relative rankings among schools of higher or equal positions to the match in the submitted list. In other words, all schools ranked above the match are in the correct order and preferred to the match. Panel (b) of Figure 7 shows that the majority of paticipants (between 76 and 84%, depending on the period) satisfy relative partial truthfulness with respect to the realized match. In period 1, the large fraction of relative partial truthfulness is mostly driven by fully truthful participants (68 out of the 80 percentage points). Since this measure uses realized instead of predicted match, the decrease in periods 2 and 3 can be plausibly explained by the fact that it may take time to notice a pattern and accurately predict the best possible match, after which it rises in periods 4 and 5. Notably, even though full truthness is low in these last two periods (37-38%), relative partial truthfulness is even higher than that in period 1.

The untruthful behavior of the participants in our experiment relates to the theory in Artemov, Che and He (2020). First, they formalize the robust equilibrium concept, which allows for mistakes along as they become payoff irrelevant when the market grows arbitrarily large and the set of feasible schools for each participant becomes close to deterministic. Even though the market is small in our experiment, we show above that empirically the most preferred feasible school is stable and easy to predict in the later periods of each round. Second, the relatively partially truthful strategies employed by the majority of our participants belong to a subset of Artemov, Che and He (2020)'s stable-response strategy, which requires the most preferred feasible school to be ranked above all other feasible schools in the limit continuum economy. Note that while a stable response strategy allows for a worse but infeasible school to be above the most preferred feasible school, relative partial truthfulness rules out this behavior. It is understandable that this is the case in a small economy, where there is more uncertainty regarding the set of feasible schools.

3.3.2 Reinforcement Learning within a Round

We examine subjects' tendency to revise their strategies when going from one period to another and find that the higher the payoff the more likely a subject would keep the same strategy in the next period, which is consistent with reinforcement learning.

Table 3.2 presents logistic regressions of the probability of no revision in four different spec-

Table 3.2: Logistic Regression - Probability of No Revision

	Pr(No revision in period t)			
	(1)	(2)	(3)	(4)
Constant	-3.848***	-4.938***	-1.602***	-3.438***
	(0.354)	(0.780)	(0.475)	(1.007)
Full Truthfulness $(t-1)$	2.049***	1.631***	1.767***	1.706***
	(0.299)	(0.320)	(0.385)	(0.428)
Relative Partial Truthfulness Only $(t-1)$	0.764**	0.644*	0.925***	0.834**
	(0.311)	(0.335)	(0.354)	(0.392)
Payoff $(t-1)$	0.365***	0.481***	0.210***	0.351***
	(0.033)	(0.044)	(0.046)	(0.059)
ΔPayoff $(t-2$ to $t-1)$			0.314*	0.301
			(0.161)	(0.193)
Revision $(t-1)$			-2.326***	-1.891***
			(0.226)	(0.243)
ΔPayoff *Revision $(t-1)$			-0.080	-0.084
			(0.172)	(0.206)
Round FEs	Yes	Yes	Yes	Yes
Participant FEs	No	Yes	No	Yes
Periods	2-5	2-5	3-5	3-5
Pseudo R-squared	0.265	0.390	0.410	0.476
N	1184	1184	888	888

Notes: "Relative partial truthfulness only" refers to strategies that satisfy relative partial trutfulness without being fully truthful.

ifications, considering the previous period (specifications 1 and 2) and the previous two periods (specifications 3 and 4). In all specifications, the coefficient on lagged payoff is positive and significant at the 1% level although the magnitude are smaller in specifications 3 and 4 when we look at a sample of period 3 onwards and control for revision in the previous period (or lack thereof) and the associated payoff change. In other words, the probability of keeping the current strategy is increasing in the payoff said strategy brings, consistent with reinforcement learning. This result is not driven by individual's intrinsic propensity to revise. In fact, the estimated impact of payoff grows when we include participant fixed effects (specification 2 as opposed to 1 and 4 as opposed to 3). The change in payoff also has a positive effect on not revising, but it becomes insignificant once participant fixed effects are controlled for. Note that having adopt a relatively partially truthful strategy (whether or not it is fully truthful) in the previous period is also less likely to lead to revision.

4 Conclusion

We use a lab experiment to study learning when participants receive feedbacks in centralized matching mechanisms widely used in school choice and other real-world applications. Our results provide a deeper understanding of participants' "mistakes" in strategyproof mechanisms like Deferred Acceptance as well as their learning process.

A novel aspect of our experimental design is to include the distinction between rounds and periods, allowing for two types of learning. Within a round, the environment (schools' priorities and students' preferences) is unchanged, and periods differ only in the rankings submitted by participants. Thus, agents learn to coordinate, which in our experiment induces them to become untruthful in a way that is payoff irrelevant. Across the rounds, the environment changes, but the mechanism does not. There is suggestive evidence that agents do have a better grasp of the underlying property of the mechanism through this experience.

We plan to run more experimental sessions with an improved design, which includes randomizing the order of the environments with and without school zones and allowing for more periods within a round. Another direction of future research would be to explore sequential admission procedures, as seen in German university admission, where the feedback and learning process (although very different from our current experiment) can be implemented in a real-world setting.

5 Appdendix A: Additional Figure

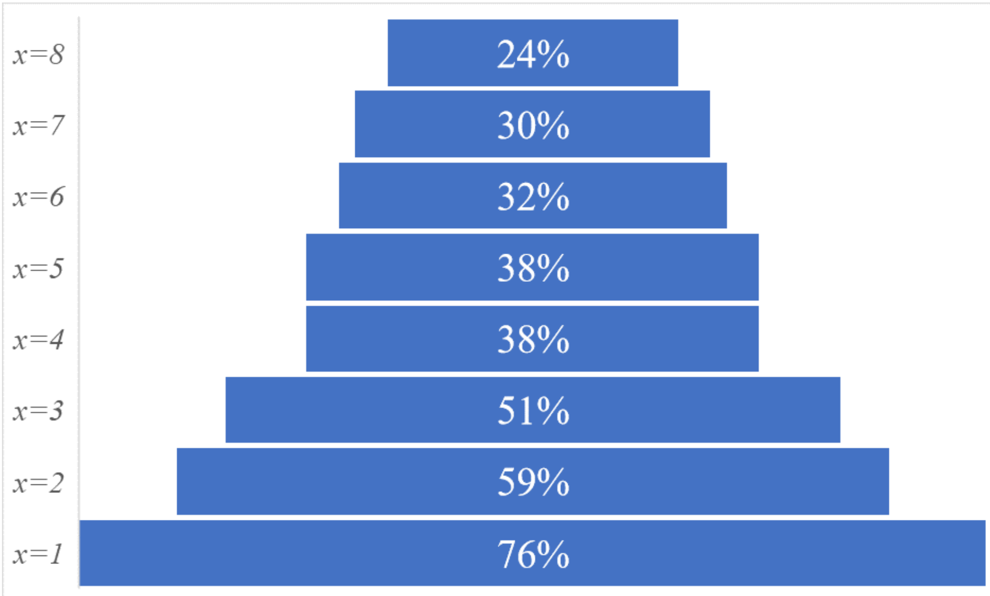


Figure 5.1: Fraction of participants who have been truthful in the first period of the last x round(s)

6 Appendix B: Instructions for the Experiment

6.1 Instructions for the Deferred Acceptance Mechanism (Mechanism G)

Thank you for participating in this experiment on decision making. From now until the end of the session any communication with other participants is forbidden. Foods and drinks are not allowed in the lab, and we ask you to **turn off your mobile phone**.

If you have any question, feel free to ask at any point of the experiment. Please do so by raising your hand and one of us will come to your desk and answer your question privately.

This session will last approximately 90 minutes and comprise of multiple rounds. Throughout the experiment the following rules apply.

You and every other participant play the role of a student who is applying to several schools sequentially. Different schools have different values for you (i.e. how much you like them), and you will choose in which order to apply to the schools. Your goal is to be accepted by a school with a high value. **Application** is done by submitting an *application list* of schools.

Each school has a **priority** list over students (i.e., a ranking of students) ,which determines in which order applicants are considered for acceptance.

Schools have different priorities over students; students have similar but not identical preferences over schools.

The computer will run a **matching** algorithm that utilizes the *application lists* submitted by the students and the schools' *priority lists* over students. This algorithm will determine who is assigned to which school.

The experiment involves N human participants and 20-N automated computer program(s), which will make decisions under the same conditions as yours.

Submitted application list

Each student (including the computer programs) submits an application list, which indicates in which order she wants to apply to schools. That is:

- The first school in the list is the first school to which the student will apply.
- The second school in the list is the the next school to which the student will apply if rejected by the first school.
- and so on...

Priority list over students

Each school has a priority list that will not be communicated to you. Schools decide to accept or reject students who applied to them using their own priority list. There are 5 schools with 4 seats at each school.

The priority list of each school is generated independently.

Student's payoff if accepted to a school

Your payoff depends on the value of the school where you are accepted. You will see on the screen a table outlining the value for each school.

Different participants may have different payoff tables.

Matching Algorithm G

The participants are assigned to schools by a multi-step algorithm.

Step 1:

- Each student applies to the school that is ranked first in her submitted application list.
- Each school tentatively accepts the applicants who have applied to it, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

Step 2, 3, ...: This step is only for the students who have been rejected in the previous step.

- Each of those students apply to the next school on their submitted application list. If a student has already applied to all the schools in her application list, then that student remains unassigned (and the algorithm ends for that student).
- Each school considers the set of students it accepted at the previous step together with the set of new applicants. From this larger set, the school tentatively accepts students, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

End: The algorithm stops when either no school receives any new applications, or no application is rejected.

Example (Mechanism G)

This is a small example that shows how the algorithm works. We consider

- 3 students: A, B and C
- 3 schools: 1, 2, and 3. Each school has 1 seat.

Schools' priority lists of students are described in Table 6.1.

Table 6.1: Schools' priority lists of students

Priority	School 1	School 2	School 3
1st	A	A	C
2nd	C	B	B
3rd	B	C	A

So, for instance, at school 3, student C is considered first, then B, then A.

The application lists submitted by the students are in Table 6.2.

Table 6.2: Students' submitted application lists

	Student A	Student B	Student C
1st application	1	1	2
2nd application	2	2	3
3rd application	3	3	1

So, for instance, student C will first apply to school 2. Then, if rejected, she will apply to school 3. Finally, if rejected again, she will apply to school 1. The other students' application lists are read similarly.

Let's run the algorithm.

Step 1 Students A and B apply to school 1, and C applies to school 2.

- School 1 starts accepting the students who applied to it following the order given by its priority. A is the top priority student, so she is accepted. School 1 cannot accept any more student (it reached its capacity of 1), so B is rejected.
- The other student, C, applies to school 2. She is the only applicant, so she is accepted.

At the end of the first step, A is tentatively accepted at school 1, C is tentatively accepted at school 2, and B is not accepted at any school.

Step 2 Student B applies to school 2.

- School 2 has now two applicants: C (from Step 1) and B. B has higher priority than C, and since school 2 has only one seat available, B is now tentatively accepted at school 2 and C is rejected.

At the end of the second step, A is tentatively accepted at school 1, B is tentatively accepted at school 2, and C is not accepted at any school.

Step 3 Student C applies to school 3.

- She is the only applicant, so she is accepted.

No student is rejected, so the algorithm stops.

The final assignment is

Student A	Student B	Student C
School 1	School 2	School 3

Part I

You can see on the main screen how the interface of the experiment looks like. The first half of the session consists of 8 Rounds, each divided into 5 Periods. We are in the first part of the session, first round, first period.

On the right-hand side, you can see the five schools (each with a different letter) and your value in case you are accepted there. Each school has 4 seats. If you are not accepted to any school, your payoff will be $0.

You create your application list by clicking on the school boxes on the right. After each click, the school is crossed out and it is added to your application list on the left, in the highest empty box still available. Keep clicking on the schools until you are ready to submit. Click the submit button to confirm your application list, or click on the clear button to reset the list and start from scratch. You can submit your list without applying to all 5 schools.

You have a limited amount of time to submit your application list. The available time is indicated by a timer below the school values. If you run out of time before clicking submit, the current application list on your screen is submitted automatically.

After a few seconds, a new period starts with the *same* school values for all students and *same* priority lists for all schools as in the previous period. You can observe the school you were accepted to in the previous period (highlighted in the red square) and the payoff associated with it (written below the application list). Now you can review your application list and submit the same or a different one. The experiment proceeds in this way for five periods.

After the end of the fifth period, a new round starts with *different* school priorities and *different* student payoffs. Remember that your objective is to get as many dollars as possible in each period.

The schools' priority lists and students' payoffs are generated as follows.

Priority list over students

Each school independently generates a random ranking of the 20 students.

Students' payoffs

Each student receives a random ranking of the schools, with values ranging from $1 to $9 with $2 increments. This ranking is not independent across students: on average students like the schools on the top more than the schools on the bottom. This means that the top school (N in the figure) is more likely to have a high payoff both for you and for the other students, and so on.

The ranking for each student is generated in this way. Schools in position 1 to 5 (ranked from top to bottom) start with a *common score* of 12, 9, 6, 3, and 0 points, that is the same for all the students. Then, each school receives a *random score*, uniformly distributed from 0 to 8. The random score is independent across students and across schools, and allows different students to have different preferences over schools. Common and random scores are summed, and the ranking is generated from the school with more points to the school with less points.

Computer program actions

At the beginning of each round, each computer program will choose an application list and submit the same list for all 5 periods. Only human participants can review their own list between periods.

If something is unclear, you can ask questions now. After answering your questions, we can proceed with the first part.

Part II

The second half of the session consists of 8 Rounds, each divided into 5 Periods. The same rules as Part I apply, with only a difference in how schools generate priority lists. You can see on the main screen the interface for the second part.

At the beginning of each round, half of the students will be randomly assigned the color orange, and the other half will be assigned the color blue. Your color is visible above the school values, and will remain the same during the round, **but your color might change between rounds**.

Schools are also randomly assigned to a color. There are three types of school colors:
2 orange schools, 2 blue schools, and 1 white school.
The color is used to fill the box with the school's name and value.

For each school, the probability of a student being among the top 4 positions of the priority list depends on her color:

- Each orange student has a 30% probability of being among the top 4 positions in the priority list of each orange school, 20% for the white school, and 10% for each blue school.
- Each blue student has a 30% probability of being among the top 4 positions in the priority list of each blue school, 20% for the white school, and 10% for each orange school.

- A separate random draw will determine which students are among the top 4 positions in each school's priority list. The relative ranking of those 4 students will be random. Similarly, the relative ranking of the students who will be ranked 5th, 6th, ... will be randomly generated.

How you value schools does not depend on your color, and values are generated as in part 1. As before, in each period you create your application list by clicking the school boxes. Each round is comprised of 5 periods.

If something is unclear, you can ask questions now. After answering your questions, we can proceed with the second part.

6.2 Instructions for the Immediate Acceptance Mechanism (Mechanism B)

All instructions, except for the description of the mechanism and the illustrating example (as seen below), are identical to the instructions for the Deferred Acceptance mechanism.

Matching Algorithm B

The participants are assigned to schools by a multi-step algorithm.

Step 1:

- Each student applies to the school that is ranked first in her submitted application list.

- Each school accepts the applicants who have applied to it, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

  The accepted applicants and their seats are removed from the system: these are their final acceptances.

Step 2, 3, ...: This step is <u>only for the students who have been rejected in the previous step</u>.

- Each of those students apply to the next school on their submitted application list. If a student has already applied to all the schools in her application list, then that student remains unassigned (and the algorithm ends for that student).

- If a school still has available seats remaining at the end of the previous step, then it accepts the applicants who have applied to it, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

The accepted applicants and their seats are removed from the system: these are their final acceptances.

End: The algorithm stops when either no school receives any applications, or no application is rejected.

Example (Mechanism B)

This is a small example that shows how the algorithm works. We consider

- 3 students: A, B and C
- 3 schools: 1, 2, and 3. Each school has 1 seat.

Schools' priority lists of students are described in Table 6.3.

Table 6.3: Schools' priority lists of students

Priority	School 1	School 2	School 3
1st	A	A	C
2nd	C	B	B
3rd	B	C	A

So, for instance, at school 3, student C is considered first, then B, then A.

The application lists submitted by the students are in Table 6.4.

Table 6.4: Students' submitted application lists

	Student A	Student B	Student C
1st application	1	1	2
2nd application	2	2	3
3rd application	3	3	1

So, for instance, student C will first apply to school 2. Then, if rejected, she will apply to school 3. Finally, if rejected again, she will apply to school 1. The other students' application lists are read similarly.

Let's run the algorithm.

Step 1 Students A and B apply to school 1, and C applies to school 2.

- School 1 starts accepting the students who applied to it following the order given by its priority. A is the top priority student, so she is accepted. School 1 cannot accept any more student (it reached its capacity of 1), so B is rejected.
- The other student, C, applies to school 2. She is the only applicant, so she is accepted.

At the end of the first step, students A and C have received their final school acceptances: A at school 1, and C at school 2. B is not accepted at any school. Both schools 1 and 2 have reached full capacity whereas school 3 still has 1 empty seat.

Step 2 Student B applies to school 2.

- School 2 is already full, so B is rejected.

At the end of the second step, B is still not accepted at any school, and school 3 still has 1 empty seat.

Step 3 Student B applies to school 3.

- She is the only applicant, so she is accepted.

No student is rejected, so the algorithm stops.

The final assignment is

Student A	Student B	Student C
School 1	School 3	School 2

References

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2020. "Stragic 'Mistakes': Implications for Market Design Research." Working Paper. 1, 3.3.1, 3.3.1

Bo, Inacio, and Rustamdjan Hakimov. 2020. "Iterative versus standard deferred acceptance: Experimental evidence." *The Economic Journal*, 130: 356–392. 1

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2010. "Constrained school choice: An experimental study." *American Economic Review*, 100: 1860–1874. 1

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2011. "A comment on âSchool choice: An experimental study." [J. Econ. Theory 127 (1) (2006) 202-231]." *Journal of Economic Theory*, 146: 392–396. 1

Chen, Yan, and Onur Kesten. 2019. "Chinese college admissions and school choice reforms: An experimental study." *Games and Economic Behavior*, 115: 83–100. 1

Chen, Yan, and Tayfun Sonmez. 2006. "School choice: An experimental study." *Journal of Economic Theory*, 127: 202–231. 1

Ding, Tingting, and Andrew Schotter. 2017. "Matching and chatting: An experimental study of the impact of network communication on school-matching mechanisms." *Games and Economic Behavior*, 103: 94–115. 1

Ding, Tingting, and Andrew Schotter. 2019. "Learning and mechanism design: An experimental test of school matching mechanisms with intergenerational advice." *The Economic Journal*, 129: 2779–2804. 1

Dur, Umut, Robert G. Hammond, and Thayer Morrill. 2019. "The secure Boston mechanism: Theory and experiments." *Experimental Economics*, 22: 918–953. 1

Gale, D., and L. S. Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69(1): 9–15. 1

Guillen, Pablo, and Alexander Hing. 2014. "Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism." *European Economic Review*, 70: 178–185. 1

Hakimov, Rustamdjan, and Dorothea Kubler. 2019. "Experiments on Matching Markets: a Survey." Working Paper. 1

Pais, Joana, and Agnes Pinter. 2008. "School choice and information: An experimental study on matching mechanisms." *Games and Economic Behavior*, 64: 303–328. 1