Games and Economic Behavior 135 (2022) 411–433

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

# The iterative deferred acceptance mechanism ☆

Inácio Bó [a,*], Rustamdjan Hakimov [b]

[a] Southwestern University of Finance and Economics, 555, Liutai Avenue, Wenjiang District, Chengdu, China
[b] University of Lausanne and WZB Berlin Social Science Center, Internef 536, Quartier de Chamberonne, CH-1015, Lausanne, Switzerland

ARTICLE INFO

*Article history:*
Received 11 December 2020
Available online 18 July 2022

*JEL classification:*
C78
C73
D78
D82

*Keywords:*
Market design
Matching
Sequential mechanisms
College admissions

ABSTRACT

Lately, there has been an increase in the use of sequential mechanisms, instead of the traditional direct counterparts, in college admissions in many countries, including Germany, Brazil, and China. We describe these mechanisms and identify their shortcomings in terms of incentives and outcome properties. We introduce a new mechanism, which improves upon these shortcomings. Unlike direct mechanisms, which ask students for a full preference ranking over colleges, our mechanism asks students to sequentially make choices or submit partial rankings from sets of colleges. These are used to produce a tentative allocation at each step. If at some point it is determined that a student can no longer be accepted into previous choice, then she is asked to make another choice among colleges that would tentatively accept her. Participants following the simple strategy of choosing the most-preferred college in each step is an ex-post equilibrium that yields the Student-Optimal Stable Matching.



## 1. Introduction

The field of market design has developed rapidly in recent years, in terms of the range of objectives that are studied—different notions of efficiency, stability, fairness, etc.—and also in the number of applications and their evaluations, in both the field and the lab. In the typical framework, the attainability of a given objective is evaluated in terms of mechanisms that require the relevant agents to submit preferences over sets of outcomes before a clearinghouse combines them using some predetermined criteria to produce an allocation. This induces a game in which the action space of the participants consists of rankings over their outcomes. By studying the incentive properties of these games, one can then see how equilibrium outcomes relate to the objectives of the market designer. These mechanisms, therefore, have two common properties: they are direct (in the sense that the participants are asked for their relevant types, in this case their preferences) and they induce a simultaneous move game: all agents simultaneously interact only once.

There are many theoretical and practical reasons for focusing on direct mechanisms. First, the revelation principle guarantees that nothing is lost by using direct mechanisms instead of alternative action spaces. Second, in the induced games, the participants have a simple strategy space, whereas strategies in sequential games may consist of large sets of contin-

☆ We thank Samson Alva, Nick Arnosti, Itai Ashlagi, Orhan Aygün, Nina Bonge, Gabriel Carroll, Li Chen, Albin Erlanson, Sombuddho Ghosh, Maria Godinho, Robert Hammond, PJ Healy, C.-Philipp Heller, Bettina Klaus, Dorothea Kübler, Alexander Nesterov, Alvin Roth, Jennifer Rontganger, Alex Teytelboym, Bertan Turhan, Utku Ünver, Yosuke Yasuda, and Alexander Westkamp for helpful comments. This research was supported by the German Science Foundation (DFG) through CRC TRR 190 and the Swiss National Science Foundation project #100018_189152. This research is partially supported by a National Natural Science Foundation of China key project (no. 72033006).

* Corresponding author.
*E-mail addresses:* mail@inaciobo.com (I. Bó), rustamdjan.hakimov@unil.ch (R. Hakimov).

https://doi.org/10.1016/j.geb.2022.07.001
0899-8256/

gency plans over information structures. Finally, if the mechanism is strategy-proof, it is very simple for a participant, with truth-telling being the expected behavior.

In this paper, we consider centralized college admissions, where the designer is interested in implementing stable outcomes. The Gale-Shapley student-proposing deferred acceptance procedure (DA) is often deemed the ideal theoretical solution for this problem due to its incentive and fairness properties (Balinski and Sönmez, 1999). However, despite the availability of DA, the last few years have seen the emergence of "iterative" (or sequential) mechanisms that match students to schools and colleges, sometimes on a very large scale. Prominent examples include the college admission mechanism for the Chinese province of Inner Mongolia (Chen and Pereyra, 2016; Gong and Liang, 2016), used for matching more than 200,000 students to universities per year, the mechanism used in Brazil to determine matching for more than two million prospective public university students per year, and the mechanism currently being used in German university admissions (Grenet et al., 2019).[1] In these iterative mechanisms, the designer and the participants potentially interact multiple times, and between these interactions, some information about intermediate outcomes is communicated to the students. In the first part of the paper, we analyze the iterative mechanisms used in Brazil, Germany, and the province of Inner Mongolia in China. We show that these mechanisms have undesirable properties, such as the inability to provide reliable information about where students could be accepted and susceptibility to new types of manipulations, in which individuals or groups of students strategically manipulate their applications in order to obtain better outcomes, at the cost of other students. In the first one, which we denote by *manipulation via postponing rejection cycles*, students strategically delay their application to a college, and by doing so prevent a consequent rejection of their application from taking place before the procedure ends. In the second one, denoted by *manipulation via cutoffs*, groups of students with high exam grades temporarily inflate the cutoff grades at some colleges and change their options in the last step, with the objective of reducing the competition faced by specific low-grade students. We show that, due to the specific characteristics of college admissions in these countries, manipulations via cutoffs are feasible both in Brazil and Inner Mongolia, and we provide anecdotal evidence that they take place in real life in the latter case. Importantly, these undesirable incentive properties distort the stability of the final outcome.

An obvious solution to all these problems would be a switch to DA. However, designers might be opposed to a sharp change in the system or may value the multiple interactions with students and the possibility of intermediate communication. Finally, the designer might prefer the iterative feature of the current mechanism. In this paper, we try to answer the following question: Can a designer implement the DA allocation, maintaining the iterative nature of the system with the possibility of intermediate communication with students?

In the second part of the paper, we propose an iterative mechanism for implementing stable outcomes in many-to-one matching markets. In the mechanism we propose, instead of requesting a full preference over all private outcomes from each participant, they are instead repeatedly asked to choose one option or to submit a (potentially partial) ranking over the options available. These choices are used to produce tentative allocations. Information about these allocations is given back to the participants before asking them for further choices and/or rankings, until a final allocation is produced. This mechanism, therefore, resembles a sequential and flexible implementation of DA, in which instead of asking for preference rankings and using them in an algorithm, the agents themselves are repeatedly asked to make choices from sets of available options. This mechanism is denoted the Iterative Deferred Acceptance Mechanism (IDAM). We also consider a prominent special case of IDAM for exam-based college admissions, where the students choose one university at a time.

The IDAM mechanism shares some properties with the standard direct implementation of DA, but not all. While DA is strategy-proof, the IDAM mechanism may lack even a weakly dominant strategy. On the other hand, all students following the simple strategy of choosing the most preferred option at each step (denoted by *straightforward strategy*) constitute an ex-post equilibrium: students would never deviate from that strategy, regardless of the profile of preferences of other players (Theorem 1). This result comes from the fact that deviating strategies, from an observer's perspective, are indistinguishable from "truthful" behavior for certain preferences. The outcome produced in that equilibrium is the same as the one produced by DA: the student-optimal stable matching (Proposition 1). Moreover, this equilibrium is essentially unique when IDAM restricts students to choosing only one option at a time (Theorem 2). Thus, the designer who has a preference for iterative allocation procedures can reach a desirable outcome. However, instead of implementing these in dominant strategies, she can use an ex-post implementation in straightforward strategies.

When comparing the IDAM with the Brazilian and Inner Mongolia mechanisms, we see that switching to our proposal provides substantial improvements. First, while IDAM implements the DA outcome through the simple and intuitive straightforward strategies, this profile might not be an equilibrium in the Brazilian mechanism, since the presence of manipulations via postponing rejection cycles might create a chain reaction of players postponing their applications, unraveling what could have been a straightforward equilibrium. Moreover, we show that the set of equilibrium outcomes of both the Brazilian and Inner Mongolia mechanisms is equivalent to the entire set of stable outcomes (Proposition 3), turning the game into a complex coordination process. Finally, we show that the IDAM is immune not only to the two types of manipulations we identified for the mechanisms currently used, but also to any other individual manipulation (Proposition 4).

In addition to the improvements that the IDAM presents when compared to the real-life instances that we evaluated, there is growing evidence that these procedures might have benefits not captured by the standard school choice model.

[1] Another example of the current use of iterative mechanisms is the school district in Wake County, NC (Dur et al., 2018).

Recent experimental evidence shows that stable allocations are reached significantly more often under iterative versions of DA than under direct DA (Klijn et al., 2019; Bó and Hakimov, 2020). This result is driven by higher rates of truthful behavior under the iterative mechanism than under DA.[2]

Another reason why iterative stable mechanisms may be desirable is that the steps involved in the production of the final allocation are more transparent. In standard DA, the rankings students submit are used in a sequential computer-operated process, which may not be entirely understood by the participants. Making the agents themselves make choices and see their effects, each step of the iterative mechanisms is simpler and better understood. In France, for instance, DA was abandoned in favor of a sequential procedure in 2018, following complaints about the lack of transparency of DA.[3]

Finally, by construction, the iterative mechanisms allow for the designer to reveal intermediate information to the students between the steps of an iterative procedure. In the mechanisms used in Inner Mongolia and Brazil, the designer places an emphasis on publishing intermediate cutoff grades which should help students update their expectations about their chances of acceptance at a university.[4]

Proofs absent from the main text and additional details can be found in the Appendix.

*Related literature*

This paper mainly relates to two lines of research in market design. One is the family of works that evaluate, from both a positive and a normative perspective, mechanisms used in the field in college admissions and school choice. While evaluating the college admission process in Turkey, Balinski and Sönmez (1999) showed that the Gale-Shapley student-proposing deferred acceptance procedure (DA) (Gale and Shapley, 1962) is 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. Examples include college and secondary school admissions in Hungary (Biró, 2011), high school admissions in Chicago (Pathak and Sönmez, 2013) and New York City (Abdulkadiroğlu et al., 2009), and elementary schools in Boston (Abdulkadiroglu et al., 2006). Other mechanisms, such as the college-proposing DA, the so-called "Boston mechanism," and the "Parallel mechanism" are used to match millions of students to schools and colleges worldwide (Chen and Kesten, 2017; Abdulkadiroğlu and Sönmez, 2003; Balinski and Sönmez, 1999). Gong and Liang (2016) apply the mechanism currently in use to match students to universities in Inner Mongolia. Grenet et al. (2019) analyze the system used for college admissions in Germany, which combines a sequential phase and a direct revelation phase. In section 4, we show that one instance of the IDAM mechanism also presents this combination and has good practical properties.

This paper is also related to the study of sequential mechanisms. Kagel et al. (1987) show that, although the second-price auction is isomorphic to an English auction, experiments show that behavior is significantly different when comparing both, with truthful behavior more prevalent in the latter. Ausubel (2004) and Ausubel (2006) propose sequential auction mechanisms for multiple (homogeneous and heterogeneous, respectively) objects. While there are direct mechanisms that implement the same outcomes in dominant strategies, the author argues that the proposed sequential mechanisms are simpler and preserve the participants' privacy.

In a recent paper, Li (2017) provides a theoretical justification for why some sequential mechanisms perform better than their direct counterparts. That justification is based on a refinement of strategy-proofness, denoted obvious strategy-proofness (OSP), in which the realization that a certain strategy is dominant does not rely on contingent reasoning. The author shows that a family of mechanisms, which includes the English auction, is OSP, therefore providing a theoretical explanation for the results in Kagel et al. (1987).[5] When it comes to stable mechanisms, however, Ashlagi and Gonczarowski (2018) show that no OSP mechanism yields stable matchings.[6]

Experimentally comparing the behavior under DA and the IDAM mechanism in the school choice setup, Bó and Hakimov (2020) show that the truthful equilibrium in IDAM, which produces the student-optimal stable matching, predicts behavior better than the dominant strategy in DA, also leading to a larger proportion of stable outcomes. Klijn et al. (2019) also present evidence in that direction. Echenique et al. (2016) analyze the behavior under an iterative DA mechanism in the two-sided matching setup and show that most realized stable outcomes were receiver-optimal. Similarly, Kagel and Levin (2009) show experimental evidence that subjects are more likely to behave in line with the equilibrium prediction in the

2 This evidence is closely related to growing experimental and empirical evidence that the strategic simplicity of DA may not be matched by an understanding, on the part of the agents, of its incentives (Chen and Sönmez, 2006; Pais and Pintér, 2008; Ding and Schotter, 2019; Rees-Jones, 2018; Hassidim et al., 2020; Chen and Pereyra, 2019). For an extended survey of the experimental literature of school choice see Hakimov and Kübler (2021).

3 Hakimov and Raghavan (2020) formalize transparency in centralized allocations, and indeed iterative mechanisms that accept the submission of one object at a time are more transparent than the direct mechanism according to their notion.

4 Grenet et al. (2019); Hakimov et al. (2020) emphasize the usefulness of this communication if students have to acquire costly information about their preferences over universities, and this acquisition is possible between the steps of the procedure.

5 Baccara et al. (2012) evaluate the use of an obviously strategy-proof mechanism — Sequential serial dictatorship — in the assignment of offices for the faculty in a new building, and analyze the effect of faculty network on the choices in the mechanism. The coordination of choices would be harder to execute using a direct mechanism. Thus, one potential benefit of using iterative mechanisms might come from the possibility of accommodating non-standard preferences.

6 The authors show, however, that there is an OSP stable mechanism when the preferences on one side of the market satisfy an acyclicity condition. This is a very restrictive condition which is not satisfied, for example, by the college admission process in Brazil.

sequential mechanisms in Ausubel (2004) than with the dominant strategy direct counterpart. These results indicate that the behavior more consistent with the equilibrium prediction in these sequential implementations is not entirely captured by the refinement proposed in Li (2017).

Other papers have evaluated non-direct iterative mechanisms for matching students to colleges or schools. In a closely related paper, Mackenzie and Zhou (2020) evaluate the family of "menu mechanisms," in which participants choose alternatives from menus, and show that straightforward strategies constitute an ex-post equilibrium when implementing strategy-proof rules. Our Theorem 1 extends that result, within our setting, to more general lengths of rankings being submitted. Dur et al. (2018) 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 made available to the participants during the process, make it somewhat comparable to the IDAM mechanism. Dur and Kesten (2019) and Haeringer and Iehlé (2021) evaluate the sequential use of direct mechanisms for a single allocation problem. In Dur and Kesten (2019), assignments made at each step are final, and in Haeringer and Iehlé (2021) students might reject their matchings and submit a potentially updated preference in the following round.

A rich series of papers also consider sequential mechanisms that implement stable matchings in equilibrium, including Alcalde and Romero-Medina (2000), Alcalde and Romero-Medina (2005), Romero-Medina and Triossi (2014), and Klaus and Klijn (2017). While many of these mechanisms implement stable allocations in equilibrium, the determination of equilibrium strategies depends on coordination between students in a way that is significantly more demanding than the equilibrium strategy that IDAM has, which depends solely on (partial) information about the student's own preferences over colleges.

## 2. Iterative mechanisms in the field

In this section, we provide a brief description of three iterative mechanisms that are currently in use for college admissions. Even though these procedures have many differences, for the most part, they share the same basic setup, which we will use for their descriptions.

There is a set of students $S = \{s_1, \dots, s_n\}$, and of colleges $C = \{c_1, \dots, c_m\}$ with fixed capacities (a maximum number of students who can be matched to them) $\left(q_{c_1}, \dots, q_{c_m}\right)$. Colleges rank all students based on some score (which in Brazil and Inner Mongolia come from a national exam, and in Germany from a combination of different criteria). If using a national exam, different colleges[7] may use different weights for the various parts of the exam. For example, economics programs could give a higher weight to the exam's math section, while medical programs could give a higher weight to the biology section. Denote by $z_c(s)$ student $s$'s resulting exam grade in college $c$. Colleges may also have a minimum acceptance grade, representing the minimum value of $z_c(s)$ a student $s$ must have to be acceptable at $c$, denoted by $\underline{z}_c$.

Given a set of students applying to a certain college, one commonly used piece of information is the *cutoff* grade for that college. A cutoff grade represents the lowest grade necessary to be accepted at a college, given the set of students applying to it. When looking at all colleges' cutoff values, a student can thus infer which ones would accept her if all other students' choices remain constant. Before the centralized iterative mechanisms were introduced in Brazil and China, it was common for students to see the historical values of the cutoffs for the different colleges as an indication of where they should apply, given their information about their own exam grades or ability. One of the advantages of the new procedures would be to allow the students to make that assessment in "real time" instead of based only on historical data.

### 2.1. The German mechanism

The procedure that we denote by German Mechanism is originally called DoSV, *Dialogorientiertes Serviceverfahren*. Our description of the procedure and its characteristics is drawn from Grenet et al. (2019). The German Mechanism is used for admission to some competitive university programs across the country and was introduced in 2012. In the winter term of 2015/16, more than 180,000 students applied to 465 programs in 89 universities.

The German Mechanism operates in three stages[8]:

- **Stage 1**: During this period, students submit a ranked ordered list with at most 12 colleges to the central clearinghouse.
- **Stage 2** (32 days):
  - During this period, each college $c$ submits the students' scores $z_c$. When the scores from a college $c$ are received, the clearinghouse automatically sends emails with offers to the $q_c$ highest-scoring students with respect to $z_c$, who had $c$ among the colleges in their list submitted during Stage 1.
  - A student with one or more offers can choose to accept one of them and leave the procedure, or can choose to hold some or all of these offers.

[7] In the countries considered, as in many others, students apply directly to specific programs in the colleges or universities. For simplicity, though, we refer only to "colleges" whenever the distinction is not necessary.

[8] While the description we provide omits some details of the procedure, it provides the key elements necessary for our analysis and for identifying shortcomings that are also present in the actual procedure. A description of the other details can be found in Grenet et al. (2019).

  - When some student rejects an offer made by some college $c$, the clearinghouse sends another offer to the student with the highest score in $z_c$, who had $c$ among the colleges in their list submitted during Stage 1 but had not yet received an offer from $c$.
- **Stage 3**: The seats that were not taken by the students who left in Stage 2—including the offers that were held but not accepted— and the students who did not leave during Stage 2—including those who held offers but did not accept any—are matched using the Gale-Shapley college-proposing deferred acceptance mechanism (Gale and Shapley, 1962), using the ranked ordered list submitted by students in Stage 1 as their preferences and the scores submitted by the colleges in Stage 2 to rank the students. Students who rejected an offer from a college in Stage 2 are removed from that college's ranking, however.

Therefore, the German Mechanism matches students to colleges first through a dynamic offers procedure (Stage 2) and then uses a standard (constrained list) college-proposing deferred acceptance for the remaining seats.

### 2.2. The Brazilian mechanism

In the period between 2010 and 2016, the precise rules which define the Brazilian Mechanism were changed multiple times. The version that we describe, due to its simplicity, is the one used in the year 2010. Although later versions have different modifications, all the problems we identify are also present in the later versions.

The mechanism runs for four days.

- On each day $t = \{1, 2, 3, 4\}$, students may each choose a college to apply to. If a student makes no choice, her last choice is used again, if any. At the end of each of the first three days, the following is executed for each college $c$:
  - If the number of students who chose $c$ and have an exam grade at that college higher than $\underline{z}_c$ is smaller than $q_c$, the cutoff grade $\zeta_c^t$ is set to $\underline{z}_c$.
  - Otherwise, the cutoff grade $\zeta_c^t$ is set to be the $q_c^{th}$ highest grade at that college among those who chose it on that day.
- The values of $\zeta_{c_1}^t, \ldots, \zeta_{c_m}^t$ are made public.
- At the end of the fourth day, a student-college matching is produced, as follows:
  - For each college $c$, the top $q_c$ students who have an exam grade higher than $\underline{z}_c$ and chose $c$ on the last day are matched to it.
  - All students who were not among the ones above remain unmatched.
  - Final cutoffs, calculated in the same way, are made public.

### 2.3. The Inner Mongolia mechanism

The Inner Mongolia Mechanism is used to match residents of the province of Inner Mongolia in China to seats reserved for those students in universities across the country. Our description of the procedure is drawn from Gong and Liang (2016). In the Inner Mongolia mechanism, all colleges use the same grades in a national exam, that is, for all $c, c'$, $z_c = z_{c'} = z$. The students are partitioned into $k$ tiers $S = S_1 \cup S_2 \cup \cdots \cup S_k$, where the grades of all those in $S_1$ are greater than those in $S_2$, and so on. That is, if $s \in S_i$, $s' \in S_j$, $i < j$ implies that $z(s) > z(s')$. The number of tiers and the number of students in each tier varies from year to year, but the former ranges from eight to 11.

When the procedure starts, as in the Brazilian Mechanism, students can choose to apply to one of the colleges available. While students make these choices, each college's cutoff grades are calculated and made public in real time. As in the Brazilian Mechanism, cutoff grades represent the lowest grade in $z$ necessary to be accepted into each college, given the choices that all students made. Students can revise their choice as many times as they want, and the cutoff values are updated continuously accordingly.

After a pre-determined $T$ number of minutes,[9] the matchings of the students in $S_1$ are finalized. If by the end of this time a student chose a certain college and the cutoff grade at that college is lower than her exam score, then the student is accepted into the college she chose. If, on the other hand, her last choice after $T$ minutes is a college with a cutoff grade above her score, she will be left unmatched.[10]

After the matchings for students in $S_1$ terminate, the remaining students again have $T$ minutes to make choices, observe cutoffs, and revise their choices, involving the seats not taken by the students in $S_1$. As for the previous tier of students, after $T$ minutes, the matches of the students in $S_2$ are finalized. Some will be matched to the last college they chose, and some will be left unmatched. This procedure keeps on going until all $k$ tiers of students are finalized.

[9] The value of $T$ is either 60 or 180 minutes.

[10] Students who are left unmatched during the Inner Mongolia Mechanism participate in a scramble procedure, which allocates leftover seats. We do not model this phase.

## 3. Shortcomings of the current mechanisms

In this section, we present some shortcomings that we identified in the mechanisms described in the previous section. Since the Brazilian and the Inner Mongolia mechanisms share more similarities with themselves than with the German Mechanism, the nature of the shortcomings also shares that relation.

To explain the results, we need to make an assumption and provide a few definitions. The assumption is that each student $s$ has preferences $P_s$ over the set of colleges and remaining unmatched that can be represented by strict linear orders. That is, they are not indifferent between any two colleges and may prefer to be left unmatched than to be matched to some colleges. Next, we say that a student $s$ **justifiably envies** a student $s'$ if $s$ is matched to college $c$, $s'$ is matched to $c'$, $s$ prefers $c'$ to $c$, and $z_{c'}(s) > z_{c'}(s')$. Finally, a matching is **wasteful** if there is a college $c$ which is left with an empty seat and a student who would rather be matched to $c$ than the match she is left with.[11]

We start with the German Mechanism. One problem of the German Mechanism is that accepting offers in Stage 2 may hurt students.

**Remark 1.** In the German Mechanism, if a student accepts an offer during Stage 2, she may justifiably envy another student at the final allocation, or the final allocation may be wasteful.

The reason for that is clear. If a student $s$ accepts an offer during Stage 2, it is possible that another offer from a more-preferred college $c^*$ could arrive later in that stage or during the algorithmic offers process that takes place in Stage 3. In this case, $c^*$ would go down their ranking of students and either make an offer to another student $s'$, in which case $s$ would justifiably envy $s'$, or $c^*$ could end up leaving an empty seat, in which case the matching would be wasteful. This is not a purely theoretical possibility. In fact, Grenet et al. (2019) show empirically that the simple fact that an offer is sent to a student during Stage 2 significantly increases the likelihood that it will be accepted, despite it not being made by more desirable universities.[12]

The second problem that we identify with the German Mechanism comes from the fact that it consists of a sequentialized version of the *college-proposing* deferred acceptance mechanism (Gale and Shapley, 1962), and as a result has some of the incentive shortcomings of the standard implementation of that mechanism (Roth, 1982), as shown below.

**Remark 2.** In the German Mechanism, students may obtain better outcomes by strategically rejecting offers during Stage 2, and/or submitting ranked ordered lists that do not represent their true preferences.

One problem with the shortcoming above is that, by improving the outcomes of students who "strategically manipulate" their choices and reports, the German Mechanism may induce an advantage to students who engage in these manipulations and/or have access to information that could assist these manipulations.

Next, let us consider the Brazilian and Inner Mongolia mechanisms. A common characteristic between both mechanisms is that the choices made by the students before these are used to produce the allocation (i.e., those made during the first three days in the Brazilian Mechanism, and those made between the finalization of the matches of each tier) have no direct effect on the outcomes. That is, as long as the final choices, used when the matchings are produced, remain the same, the outcome will also not be changed.

Since cutoffs are calculated and made public during these "practice runs" that are present in both mechanisms, they could inform students about whether an application to a certain college is likely or not to be accepted in the end. If they do, then even though they do not have a direct effect on the final outcomes, they could provide information that guides students to better outcomes. However, the problem is that since students are not restricted by which college they can apply to and when, these cutoff values can freely fluctuate up and down, without necessarily providing any reliable information about which colleges would or would not accept a student at the end of the process.

**Remark 3.** During the "practice runs" of the Brazilian and Inner Mongolia mechanisms, each college's cutoff values may go up or down from one day to the next.

Since the cutoff values may fluctuate, for a given college, a student who has a grade higher than the cutoff cannot be sure that she will be accepted into that college if she chooses it, and a student who has a grade lower than that cutoff cannot be sure that she would not be accepted.[13]

In the next subsection, we discuss other shortcomings of the Brazilian and Inner Mongolia mechanisms, which are new to the literature, and provide some indications they might be affecting real-life outcomes.

[11] Also, if that college has a minimum grade, that student will have a grade above it.

[12] The authors of that paper interpret these acceptances as resulting from an endogenous formation of preferences with regret avoidance.

[13] A previous working version of this paper showed, using data from Brazil in 2016, that cutoff values for many programs did fluctuate substantially, both up and down, from one day to the next.

### 3.1. Dynamic harmful manipulations

Given the sequential nature of the Brazilian and Inner Mongolia mechanisms, and the fact that students can react to changes in the cutoff values that result from other students' choices, the space of strategies and behaviors that they can exhibit is very large. In what follows, we will restrict our analysis to the question of whether strategic agents can obtain better outcomes by following certain strategic manipulations that are inherently related to the functioning of these mechanisms.

We will say that a student $s$'s behavior, in both the Brazilian and Inner Mongolia mechanisms, is **straightforward** if at any moment during the execution of the mechanism, she chooses $c^* \in C \cup \{\emptyset\}$, where $c^*$ is the most preferred element of $C \cup \{\emptyset\}$, with respect to $P_s$, for which the cutoff grade in that period is below $s$'s grade in that college.

Let $\mu$ be the matching of students to colleges that would result from all the students following the straightforward behavior, in either the Brazilian or Inner Mongolia mechanism. We will say that a coalition of students $S^C \subseteq S$[14] has a **harmful manipulation** if there is some behavior, different from straightforward, for at least one student in $S^C$ such that when every student in $S \backslash S^C$ follows the straightforward behavior, the new matching that will be produced is weakly preferred by every student in $S^C$, strictly preferred by at least one of them, and strictly worse for at least one student in $S \backslash S^C$. If $|S^C| = 1$, we denote it as an **individual harmful manipulation**.[15]

Next, we will introduce two types of harmful manipulations that might be performed in dynamic mechanisms: *manipulations via postponing rejection cycles* and *manipulations via cutoffs*.

#### 3.1.1. Manipulations via postponing rejection cycles

Manipulations via postponing rejection cycles are individual harmful manipulations in which a student postpones the action of applying to a certain college, with the objective of preventing a rejection cycle from coming back to her before the last period.

**Example 1.** Consider the set of students $S = \{s_1, s_2, s_3\}$ and of colleges $C = \{c_1, c_2, c_3\}$, each with capacity $q_i = 1$ and no minimum scores in the Brazilian mechanism. Students' preferences and exam grades are as follows:

$$
\begin{array}{lccc}
P_{s_1} & : & c_1 & c_3 & \sim \\
P_{s_2} & : & c_1 & c_2 & \sim \\
P_{s_3} & : & c_2 & c_1 & \sim
\end{array}
$$

| | $z_{c_1}(s)$ | $z_{c_2}(s)$ | $z_{c_3}(s)$ |
|---|---|---|---|
| $s_1$ | 200 | 100 | 100 |
| $s_2$ | 100 | 300 | 200 |
| $s_3$ | 300 | 200 | 300 |

When every student follows the straightforward behavior, when $t = 1$, $s_1$ and $s_2$ apply to $c_1$, and $s_3$ applies to $c_2$. At $t = 2$, since the cutoff at $c_1$ is higher than her grade, $s_2$ applies to $c_2$. At $t = 3$, since the cutoff at $c_2$ is higher than her grade, $s_3$ applies to $c_1$. Finally, at $t = 4$, since the cutoff at $c_1$ is higher than her grade, $s_1$ applies to $c_3$. The final matching assigns $s_1$ to $c_3$, $s_2$ to $c_2$, and $s_3$ to $c_1$.

Suppose that instead of following the straightforward behavior and applying to $c_1$ at $t = 1$, student $s_1$ waits until $t = 3$ to do that.[16] Then, assuming that the other students followed the straightforward behavior, when $t = 1$, $s_2$ applies to $c_1$ and $s_3$ applies to $c_2$. At $t = 2$ no changes take place. At $t = 3$, $s_1$ applies to $c_1$. At $t = 4$, since the cutoff at $c_1$ is higher than her grade, $s_2$ applies to $c_2$. The final matching assigns $s_1$ to $c_1$, $s_2$ to $c_2$, and $s_3$ *is left unmatched*. Notice that this was a *harmful manipulation* by $s_1$: she obtained a better matching, but that came at the cost of $s_3$, who obtained a worse one.

We can define a **rejection chain**[17] as a sequence of applications and rejections $s^1 \to c^1 \to s^2 \to \cdots \to s^k \to c^1$ that can take place when students follow the straightforward behavior, where $i \neq j$ implies $s^i \neq s^j$ and $c^i \neq c^j$. The sequence represents the situation in which initially student $s^2$ applies to $c^1$, $s^3$ to $c^2$, etc, and all of them have grades higher than the cutoff value at the colleges that they are applying to. Then, first student $s^1$ applies to college $c^1$. This leads the cutoff value at $c^1$ to increase and go above $z_{c^1}(s^2)$. As a result, $s^2$ applies to $c^2$. This leads the cutoff value at $c^2$ to increase and go above $z_{c^2}(s^3)$, resulting in $s^3$ applying to $c^3$, and so on. In the last step of the rejection chain, student $s^k$ applies to $c^1$ and as a result the cutoff value at $c^1$ becomes higher than $z_{c^1}(s^1)$. Importantly, the relationship between these applications and rejections is causal: if $s^1$ does not apply to $c^1$, students would end up matched to the colleges that they are applying to before these events.

A **manipulation via postponing rejection cycles** is a type of harmful manipulation in which, when all the students follow the straightforward behavior, the rejection chain $s^1 \to c^1 \to s^2 \to \cdots \to s^k \to c^1$ is observed and students $s_2, \ldots, s_k$ are matched, at the end of the procedure, to the choices they make within the cycle. Student $s^1$ performs this manipulation by *postponing her application* to $c^1$ to being *after* period $T^* - k - 1$, where $T^*$ is the last period in which the students can make applications before the mechanism produces an outcome. This results in $s^1$ being matched to $c^1$, while she would

[14] We denote subsets by $\subseteq$, and strict subsets by $\subset$.

[15] A closely related notion of harm and its analysis in direct mechanisms is presented in Afacan and Dur (2017).

[16] It is not crucial for our argument that $s_1$ spends two periods without any activity. She could have, for example, applied to other "irrelevant" colleges during that time.

[17] The notion of rejection chains is not original to our paper and was introduced by Kesten (2010).

have matched to a less-preferred alternative if she followed the straightforward behavior, not postponing her application to $c^1$. Notice that this is a harmful manipulation: one of the students in the rejection cycle will be left unmatched. Example 1, therefore, implies the following remark.

**Remark 4.** The Brazilian Mechanism is manipulable via postponing rejection cycles.

The reason why we cannot say that the Inner Mongolia Mechanism is also manipulable via postponing rejection cycles is that there can be no rejection cycle, since all colleges use the same grades to rank the students. If we allowed for the students to have different grades across colleges, and assumed that there is a minimum amount of time that it takes for students to change their applications, then Example 1 could also have been used there, for students within the same tier.

Notice that this type of manipulation consists of a small departure from the straightforward behavior: the order of applications remains the same, but they are simply postponed.

*3.1.2. Manipulations via cutoffs*

A manipulation via cutoffs occurs when a coalition of students artificially increases the cutoff values of some colleges as a way of preventing other students' applications and then, shortly before matchings are finalized, vacate those seats so that students with a lower exam grade, aware of that manipulation, take their places.

**Example 2** *(Manipulation via cutoffs).* Consider the set of students $S = \{s_1, s_2, s_3, s_4\}$ and of colleges $C = \{c_1, c_2, c_3\}$, each with capacity $q_i = 1$ and no minimum scores. Students' preferences and exam grades are as follows:

$$\begin{array}{lllll} P_{s_1} & : & c_1 & c_2 & c_3 \\ P_{s_2} & : & c_1 & c_2 & c_3 \\ P_{s_3} & : & c_1 & c_2 & c_3 \\ P_{s_4} & : & c_2 & c_1 & c_3 \end{array} \qquad \begin{array}{cccc} & z_{c_1}(s) & z_{c_2}(s) & z_{c_3}(s) \\ s_1 & 100 & 100 & 100 \\ s_2 & 200 & 200 & 200 \\ s_3 & 300 & 300 & 300 \\ s_4 & 400 & 400 & 400 \end{array}$$

Suppose that the Brazilian Mechanism is used, and that students present straightforward behaviors. The cutoff values at the end of each day, and the final matching of students to colleges, would then be as follows (the cutoffs at $t = 4$ represent the final allocation cutoffs):

$$\begin{array}{cccc} & c_1 & c_2 & c_3 \\ t = 1 & 300 & 400 & 0 \\ t = 2, 3, 4 & 300 & 400 & 200 \end{array} \qquad \begin{pmatrix} c_1 & c_2 & c_3 & \emptyset \\ s_3 & s_4 & s_2 & s_1 \end{pmatrix}$$

Suppose, however, that students $s_1$ and $s_4$ collude and modify their behavior, acting instead as follows. During $t = 1, 2, 3$, student $s_1$ chooses college $c_3$ and student $s_4$ chooses college $c_1$. On day $t = 4$, student $s_1$ chooses college $c_1$ and student $s_4$ chooses college $c_2$. Assuming that the other students present straightforward behavior, the cutoff values at the end of each day, and the final matching of students to colleges would be as follows:

$$\begin{array}{cccc} & c_1 & c_2 & c_3 \\ t = 1 & 400 & 0 & 100 \\ t = 2 & 400 & 300 & 100 \\ t = 3 & 400 & 300 & 200 \\ t = 4 & 100 & 400 & 200 \end{array} \qquad \begin{pmatrix} c_1 & c_2 & c_3 & \emptyset \\ s_1 & s_4 & s_2 & s_3 \end{pmatrix}$$

Student $s_1$ is strictly better off with this manipulation, while $s_4$ is matched to the same college in both cases. Moreover, $s_3$, who was not part of the manipulating coalition, is strictly worse off.

Consider a college admissions problem, with a set of students $S$ and of colleges $C$, where the outcome of the Brazilian mechanism when every student follows the straightforward behavior is $\mu$. Let $T^*$ be the number of days in which students can make applications in the mechanism, and $\zeta_{c^*}$ be the final value of the cutoff of college $c^* \in C$ after the end of day $T^*$. A **manipulation via cutoffs** is a type of harmful manipulation performed by a coalition $S^C \subset S$ of students, where $S^C = S^H \cup \{s^L\}$, $s^L \notin S^H$, and $s^L$ prefers $c^*$ to her match under $\mu$. Instead of following the straightforward behavior, students in $S^C$ apply to college $c^*$ in every period $t = 1, \ldots, T^* - 1$, and as a result of that, the cutoff at $c^*$ is strictly higher than $\zeta_{c^*}$ in every period $t = 1, \ldots, T^* - 1$. *Because of that increase in the value of the cutoff, some of the students who are matched to $c^*$ under $\mu$ do not apply to it while following the straightforward behavior.* In period $t = T^*$, the students in $S^H$ change their application to their outcome under $\mu$. The matching produced at the end of the procedure is the same as $\mu$ for the students in $S^H$, and $s^L$ is matched to $c^*$—a strictly better outcome. Notice that if we assumed that there is a minimum amount of time that it takes for students to change their applications, Example 2 could also have been used for the Inner Mongolia mechanism, for students within the same tier. We can, therefore, conclude with the following remark.

**Remark 5.** The Brazilian and Inner Mongolia mechanisms are subject to manipulations via cutoffs.

One thing to note is that the coalition $S^H$ must be able to significantly increase the value of the cutoff when they apply to a college. This may not be easy. After all, colleges typically accept hundreds or thousands of students every year, and a coalition of hundreds of high-achieving students performing these potentially risky manipulations does not seem realistic. In many countries (including Brazil and China), however, students apply directly to specific programs at the universities, so even though the universities as a whole accept hundreds or thousands of students, the number of seats at each program is often below 100, and many times lower than 30 or 20. Moreover, even those seats are often subdivided. In China, the seats in each program are partitioned between seats reserved for candidates from specific provinces. In Brazil, federal universities partition the seats in the programs into five sets of seats, reserved for different combinations of ethnic and income characteristics (Aygün and Bó, 2021). Finally, universities sometimes offer only a subset of the total number of seats in a program through the centralized matching process. In fact, the median number of seats offered in each option available during the January 2016 selection process in Brazil, where more than 228,000 seats in public universities were offered, was *five*.[18]

There is evidence that this type of manipulation takes place in real life. In Inner Mongolia, there is evidence that high schools are coordinating students' actions in a behavior consistent with manipulations via cutoffs, as documented by China News[19]:

> "(...) the clearinghouse found that some high scoring students applied to a college with lower cutoff score. (...) On the other hand, some other students, from the same high school often, applied to colleges that their score would not allow them to go initially (...) the clearinghouse noticed that, 2 or 3 min before the deadline, the ranking of students in the system is changing—this was evidence that high schools are organizing their own high scoring students to occupy seats for low scoring students."

Regarding the Brazilian Mechanism, while we could not find any article in the media describing manipulations via cutoffs, there is at least one video on YouTube describing how to perform the manipulation (egoeimididaskalos, 2012).

The Brazilian mechanism's shortcomings are closely related to the motivation behind the introduction of activity rules in the combinatorial clock auction (Ausubel et al., 2006; Ausubel and Baranov, 2014). In these auctions, participants' bids might be inconsistent over time, and agents might have the incentive to use "sniping behavior," in which bidders conceal their true intentions until the very end of the auction. This behavior is somewhat related to the two types of manipulations that we described, where coalitions of students also "conceal" their true preferences until the last period. The solution proposed by the authors is to use "activity rules," which are rules that restrict participants' bids to guarantee that they satisfy consistency properties, such as some Revealed Preference axiom. In the next section, we will introduce the Iterative Deferred Acceptance Mechanism, an iterative mechanism for matching problems that include the ones dealt with by the mechanisms we evaluated. As we will show below, it improves upon the shortcomings that we identified above. For one thing, it differs from the Brazilian and Inner Mongolia mechanisms in that it restricts the participants' behavior in a way that is analogous to the activity rules used in these combinatorial auctions.

More precisely, our proposal introduces commitment to choices: as long as a student would be matched to the last college she applied to—given other students' choices—she is not able to change her choice. This has at least two consequences: first, it induces a stronger alignment between each application and the student's preferences, since each choice might be the last one and therefore her final allocation. Moreover, as long as rejections are final (as they will be in our proposal), this commitment makes the set of allocations that might be produced in the end converge "mechanically": regardless of the choices that students make, they determine allocations that are no longer possible. Activity rules in auctions perform similar roles: limiting bids and quantities to move monotonically also imply that each bid creates a new lower-bound to the price to be paid, eliminating a "cheap talk" nature that intermediate bids would otherwise have. Moreover, they also narrow the space of allocations that might be produced, in terms of both quantities and prices. In other words, one can say that their main role is to "make talk less cheap" and therefore more informative and consequential.

## 4. The iterative deferred acceptance mechanism

In this section, we introduce our proposed mechanism, denoted as the Iterative Deferred Acceptance Mechanism (IDAM). We consider a general setup, in which the admission criteria used by colleges may be more general than one simply based on exam grades, allowing, for example, for the use of affirmative action policies or variations in financial aid.[20] This version also allows for the same students and colleges to be matched under different contractual terms, as in the matching with contracts model introduced by Hatfield and Milgrom (2005). Finally, we also allow for cases where students might submit not only one choice at a time but also rankings over the available options.

[18] Each program was partitioned into five options: one for each combination of characteristics related to affirmative action. Each one was treated like a college in our framework. In other words, the median number of seats in each one of these partitions was five.

[19] Source (in Chinese): https://www.chinanews.com.cn/edu/2014/09-04/6562740.shtml. (Accessed on April 2, 2021.)

[20] See, for example, Hafalir et al. (2013); Aygün and Bó (2021); Shorrer and Sóvágó (2017); Hassidim et al. (2020), and Yenmez (2018).

A special case of IDAM, which considers the same setup of college admissions based on national exam grades, is introduced in subsection 4.1.

A **matching with contracts market** is a tuple $\langle S, C, T, X, P_S, F_C \rangle$:

1. A finite set of **students** $S = \{s_1, \ldots, s_n\}$,
2. A finite set of **colleges** $C = \{c_1, \ldots, c_m\}$,
3. A vector of **contractual terms** $T = (t_1, \ldots, t_\ell)$,
4. A set of **valid contracts** $X \subseteq C \times S \times T$,
5. A list of strict **student preferences** $P_S = \left(P_{s_1}, \ldots, P_{s_n}\right)$ over contracts,[21] and the respectively derived weak preferences $R_S$,
6. A list of **college choice functions** over sets of contracts $F_C = \left(f_{c_1}, \ldots, f_{c_m}\right)$, where for every $c \in C$ and $I \subseteq X$, $f_c : 2^X \to 2^X$, $f_c(I) \subseteq I$, $\left\{(c,s,t), (c,s',t')\right\} \subseteq f_c(I) \implies s \neq s'$ and $\left(c', s, t\right) \in f_c(I) \implies c' = c$.

We say that a contract $x$ is **acceptable** to $s$ if $x \; P_s \; \emptyset$, and to $c$ if $x \in f_c(\{x\})$. For any $I \subseteq X$, $s \in S$, and $c \in C$, denote $I_s \equiv \left\{(c, s', t) \in I : s' = s\right\}$, $I_c \equiv \left\{(c', s, t) \in I : c' = c\right\}$, $s(I) \equiv \{s \in S : \exists (c,s,t) \in I\}$, and $c(I)$ to be defined analogously. We abuse notation and let $c(x)$ and $s(x)$ be the college and student in contract $x$, respectively. For each student $s \in S$, her preferences $P_s$ are defined over the set $X_s$. An **outcome** is a set of contracts $Y \subseteq X$ such that $Y$ contains at most one contract per student, that is, $|Y_s| \leq 1$ for each $s \in S$. Denote by $\mathcal{X}$ the set of all outcomes. An outcome $Y$ is **individually rational** if for every student $s$, $Y_s R_s \emptyset$ and for every college $c$, $Y_c = f_c(Y_c)$. Define by **maximum rank function** a function $\pi : \mathbb{Z}^+ \to \mathbb{N} \cup \{\infty\}$ that defines, for each step $t = 0, \ldots, T^{Max}$, the maximal length of a ranking that a student may submit.

Next, we provide an informal description of the **IDAM mechanism**. Its formal definition can be found in the Appendix. The IDAM consists of the following steps:

$t = 1$ Each student is given an individualized menu of contracts, consisting of the contracts involving said student that colleges deem acceptable and the null option $\emptyset$. Each student who is given a non-empty menu is asked to submit an ordered list with at most $\pi(1)$ contracts from their menu. After all students have submitted their lists (or opted not to), these are used to perform a cumulative offer process (Hatfield and Milgrom, 2005). Students who are offered a menu with contracts but opt not to submit are left unmatched. That is, one student at a time offers her highest-ranked contract on the list submitted to the involved college. The colleges choose among all contracts offered, cumulatively, with their choice functions. Whenever a contract is rejected, the student involved in it offers the next highest-ranked contract, if any. The step ends whenever every student has a contract held by a college or has all of those on the submitted list rejected.

$1 < t \leq T^{Max}$ First, the set of contracts held by colleges by the end of period $t-1$ is broadcast to the students.[22] There are two cases to consider:

- $(i)$ If $\pi(t) = \infty$ and $T^{Max} < \infty$, every student is given an individualized menu of contracts, which consists of the contracts involving said student that the colleges would accept[23]—while having all contracts that were offered in previous steps still available to these colleges— in addition to the null option $\emptyset$.
- $(ii)$ If $\pi(t) \neq \infty$ or $T^{Max} = \infty$, each student who does not have a contract being held by a college is given an individualized menu of contracts, consisting of the contracts involving said student that the colleges would accept—while having all contracts that were offered in previous steps still available to these colleges—in addition to the null option $\emptyset$.

Each student who is given a non-empty menu is asked to submit an ordered list with at most $\pi(t)$ contracts from their menu. As in the previous step, the same cumulative offer process is undertaken, where the lists submitted by the students in this and previous steps are used to offer contracts to colleges, which are considered together with those offered in previous steps. Students who are offered a menu with contracts but opt not to submit are left unmatched.

The process ends after the step $t = T^{Max}$ or whenever the set of contracts held by all colleges does not change from one step to the next. Denote that last step by $T^*$.[24] Notice that menus always include the null option $\emptyset$, so whenever it helps exposition, we will omit it from our examples.

Different choices of the parameters $T^{Max}$ and $\pi$ result in a variety of qualitatively different procedures. For example, if $T^{Max} = 1$ and $\pi(1) = \infty$, the IDAM mechanism reduces to the usual direct revelation version of the cumulative offer mechanism (Hatfield and Kojima, 2008). Another alternative is what we denote the **free-form IDAM**, where $T^{Max} = \infty$ and

---

[21] In some places we abuse notation and also use $P_s$ over sets with only one contract. Here, $\emptyset$ represents the null contract for the student, representing a student remaining unmatched. We also assume that a student's preference is over contracts in which she is involved and $\emptyset$.

[22] Our positive results using the IDAM mechanism do not rely on the information available to players about other players' characteristics and actions. We introduced this information about tentative allocations as more granular counterparts of cutoff grades in college admissions.

[23] We say that a college $c$ would accept a contract $x$ if $x \in f_c(A \cup \{x\})$, where $A$ is the set of contracts that were offered to $c$ in previous steps.

[24] Depending on the properties of the college choice functions, it is possible to have the IDAM mechanism with $T^{Max} = \infty$ that never terminates. However, for all the results that follow, we will make assumptions of these functions that guarantee this will not happen.

for every $t$, $\pi(t) = \infty$. That is, in a free-form IDAM students can, in any period, submit rankings of any length they wish, and are asked to submit a new one only if they would not be matched to any contract in their previously submitted rankings.

Another alternative that a policymaker could adopt is to use a hybrid of the iterative mechanism considered here and the traditional deferred acceptance, which we denote by **IDAM+DA**. It consists of running the IDAM mechanism, with students making only one choice at a time, for a fixed number of steps, and then asking students to submit a ranking over the remaining options. Formally, for a given number of steps $k > 0$, the IDAM+DA is simply defined as the IDAM mechanism in which the maximum rank function is such that $T^{Max} = k + 1$, for all $t \in \{1, \dots, k\}$, $\pi(t) = 1$, and $\pi(k+1) = \infty$.

One of the main advantages of the IDAM+DA is that it ends after a number of steps set by the designer: $k + 1$. It also has some similarities to the German Mechanism: it consists of a dynamic stage followed by a submission of rankings for a deferred acceptance algorithm that terminates the matching. Unlike the German Mechanism, however, it provides incentives for students to follow the simple straightforward strategy, in which a stable matching is produced.

**Example 3.** Consider a matching with contracts problem in which there are four colleges $C = \{c_1, c_2, c_3, c_4\}$, each with one seat available, and four students $S = \{s_1, s_2, s_3, s_4\}$. Colleges may accept students with or without financial aid. Denote by $(c, F)_s$ and $(c, N)_s$ contracts between student $s$ and college $c$ with and without financial aid, respectively. Colleges have a specific criterion for choosing contracts: contracts without financial aid are always preferred to contracts with aid. When comparing contracts that do not differ in that dimension, students with higher exam scores are preferred.[25] Students' grades in the national exam follow their indexes: $s_1$ has the highest grade, $s_2$ the second highest, etc. Let the maximum rank function be such that $\pi(t) = 2$ for every $t \geq 1$ and $T^{Max} = \infty$.

**Step** $t = 0$: In the first step, all students are given the same menu, containing contracts with and without financial aid, for all colleges. Suppose that students' submitted lists are: $(c_1, F)_{s_1} \succ_{s_1} (c_4, F)_{s_1}$, $(c_1, F)_{s_2} \succ_{s_2} (c_2, F)_{s_2}$, $(c_2, F)_{s_3} \succ_{s_3} (c_2, N)_{s_3}$ and $(c_1, F)_{s_4} \succ_{s_4}$. In the first list submitted by student $s_1$, therefore, she lists college $c_1$ first with financial aid followed by $c_4$, also with financial aid. Student $s_4$ opted to submit a list containing only one option, which is, of course, also a valid list.

The lists submitted are used in a cumulative offer process. In it, student $s_1$ has her top contract tentatively accepted, since no contract without financial aid is offered during that process. Moreover, although student $s_3$ has her contract with $c_2$ with financial aid rejected, her contract without financial aid is tentatively accepted, since colleges always prefer those. As a result, $s_2$ has both contracts in her list rejected. Student $s_4$'s only contract is also rejected.

**Step** $t = 1$: In this step, student $s_1$ is tentatively matched to $(c_1, F)_{s_1}$, and $s_3$ to $(c_2, N)_{s_3}$. Only students $s_2$ and $s_4$ are asked to submit lists. Both students are given all contracts with colleges $c_3$ and $c_4$ in their menus. Since $s_4$ has a low exam grade and $c_2$ tentatively holds a student without financial aid, contracts with $c_2$ would no longer be accepted. Also, while the student with the highest exam grade is tentatively matched to college $c_1$, a contract without financial aid is offered in the menu to students $s_2$ and $s_4$. Therefore, for $s_2$, only contracts without financial aid are offered for colleges $c_1$ and $c_2$, and for $s_4$ the only additional contract in the menu is $(c_1, N)_{s_4}$. Suppose that students' submitted lists are: $(c_1, N)_{s_2} \succ_{s_2} (c_4, F)_{s_2}$ and $(c_1, N)_{s_4} \succ_{s_4} (c_4, F)_{s_4}$.

**Step** $t = 2$: In this step, student $s_1$ is tentatively matched to $(c_4, F)_{s_1}$, $s_2$ to $(c_1, N)_{s_2}$, and $s_3$ to $(c_2, N)_{s_3}$. Only student $s_4$, therefore, will be given a menu, with only three contracts: $(c_3, F)_{s_4}$, $(c_3, F)_{s_4}$ and $(c_3, F)_{s_4}$. This happens even though $s_1$ had her previous match to $c_1$ rejected. The mechanism continues down her submitted list at step $t = 0$ and matches her to $c_4$ with financial aid. Suppose that $s_4$'s submitted list is $(c_3, F)_{s_4} \succ_{s_4} (c_4, N)_{s_4}$.

**Step** $t = 3$: In this final step, the final allocation is produced:
$\{(c_4, F)_{s_1}, (c_1, N)_{s_2}, (c_2, N)_{s_3}, (c_3, F)_{s_4}\}$.

The example above highlights some of the main characteristics of the IDAM mechanism. Students only have to submit a list when the absence of that information would not allow the mechanism to determine where their next tentative allocation (if any) should be. Student $s_1$, for example, did not have to submit anything after the first step, despite the fact that her tentative match was rejected after the second step. Contracts that are no longer a possible allocation for a student are not offered in their menus, reducing the number of options to consider. Students are also free to choose the length of the list they submit, up to the limit established by the maximum rank function $\pi$.

Whenever $T^{Max} = \infty$ or there is a $t \leq T^{Max}$ such that $\pi(t) = \infty$, we say that the IDAM is **unbounded**. When the IDAM is unbounded, therefore, a student is able to express, either over time or via a ranking in some steps, a sequence of choices over as many contracts as she wishes.

An outcome $X' \subseteq X$ is **stable** if it is individually rational and there is no college $c$ and set of contracts $X'' \subset X$ such that $X'' \neq f_c(X')$, $X'' = f_c(X' \cup X'')$ and for every $s \in s(X'')$, $X''_s R_s X'_s$. More specifically, we say that a student $s$ and college $c$ form a **blocking pair** under $X'$ if $s$ has a contract in $X'' \backslash X'$. A stable outcome is the **student-optimal stable allocation** if every student weakly prefers it to any other stable outcome.

In order to guarantee that the outcomes and incentives of the IDAM mechanism satisfy desirable properties, it is necessary to impose some restrictions on the colleges' choice functions. The first one comes from Hatfield and Milgrom (2005):

[25] We do not argue that these preferences are typical or realistic. We use them because they are simple but allow for an informative example of the steps of the IDAM mechanism.

**Definition 1.** Contracts in $X$ are **substitutes** for college $c$ under $f_c$ if there do not exist a set of contracts $Y \subseteq X$ and a pair of contracts $x, z \in X \backslash Y$ such that $z \notin f_c(Y \cup \{z\})$, and $z \in f_c(Y \cup \{x, z\})$.

Another condition that we use comes from Aygün and Sönmez (2013):

**Definition 2.** The choice function $f$ satisfies the **irrelevance of rejected contracts (IRC)** if $x \notin f\left(X' \cup \{x\}\right)$ implies $f\left(X' \cup \{x\}\right) = f\left(X'\right)$ for all $X' \subset X$ and $x \in X \backslash X'$.

Finally, the last property that will be used was also introduced in Hatfield and Milgrom (2005):

**Definition 3.** The choice function $f$ satisfies the **law of aggregate demand** if for all $Y \subseteq Z$, $|f(Y)| \leq |f(Z)|$.

**Lemma 1.** *Assume that for every college $c \in C$, $f_c$ satisfies IRC, and contracts in $X$ are substitutes. Then, for every student $s$ and $0 \leq t \leq t' \leq T^*$, if the set of contracts in the menu given to $s$ at step $t$ is non-empty, then all contracts in a menu given at step $t'$ are also in the menu given at step $t$.*

Lemma 1 posits that once a contract becomes unavailable for a given student, that contract will never become available again, regardless of the strategies used by the students. This shows that one piece of information given by the mechanism after each step—the set of acceptable contracts available to the student—constitutes reliable information about the contracts that are not available anymore for a student, as opposed to the Brazilian and Inner Mongolia mechanisms.

Let us now consider the strategies that students might follow when interacting with the IDAM mechanism. A **strategy** $\sigma_s$ for student $s$ maps, for each sequence of sets of contracts held by colleges broadcast in previous periods, and each sequence of menus offered to $s$ up to $t$, an ordered list with at most $\pi(t)$ contracts in the menu $\psi^t$ given to $s$ in period $t$. A **type-strategy** for student $s$ is a function $\Sigma$ that maps each preference $P_s$ over $X_s \cup \{\emptyset\}$ to a strategy.

We define a formal generalization of "straightforward behavior" (Roth and Sotomayor, 1992) when interacting with the IDAM mechanism:

**Definition 4.** A strategy of student $s \in S$ is **straightforward with respect to** $P^*$ if for every step $t$ in which a non-empty menu $\psi_s^t$ is offered by the mechanism, $s$ submits a ranking with the top $k\left(\psi_s^t, t\right)$ options in $\psi_s^t$, ordered as in $P^*$, where $k: 2^X \times \mathbb{N} \to \mathbb{N}$ is a function such that, for every $t$, $1 \leq k(\cdot, t) \leq \pi(t)$, and $k(\cdot, t^\infty) = \left|\psi_s^{t^\infty}\right|$, where $t^\infty$ is the highest value of $t$ such that $\pi(t) = \infty$. A type-strategy $\Sigma_s$ of student $s$ is **straightforward** if for every preference $P_s$, $\Sigma_s(P_s)$ is straightforward with respect to $P_s$

While the definition above refers to a class of type-strategies, whenever there is no ambiguity we will refer to it simply as *straightforward strategies*. Therefore, a strategy is straightforward in the IDAM when, in every step, the student submits either her full preference over the contracts in the menu offered or some *truncation* of her true preference. Given any preference $P^*$, the class of straightforward strategies contain all the different degrees of truncation in the submitted rankings at each step. Moreover, when there are multiple steps in which a student can submit an unbounded ranking over contracts, she **must** rank all the alternatives presented in a menu at some step in which that is allowed.[26] When $\pi(t) = 1$ for all $t$, the definition of a straightforward strategy reduces to one of straightforward behavior in Roth and Sotomayor (1992): every time the student is asked to make a choice, she picks her most preferred alternative with respect to her true preference. When students follow straightforward strategies, the outcome produced by the unbounded IDAM is of a well-known type:

**Proposition 1.** *Assume that, for every college $c \in C$, $f_c$ satisfies IRC and contracts are substitutes. If all students' strategies are straightforward with respect to $P_S$, there is a finite number of steps $T^*$ after which the outcome of the unbounded IDAM mechanism is the student-optimal stable outcome with respect to $P_S$.*

The proof of Proposition 1 is based on the fact, shown in Hirata and Kasuya (2014), that the cumulative offer process that takes place during the IDAM mechanism is *order independent* and when students follow straightforward strategies the outcome is the student-optimal stable outcome. As a result, all combinations of such strategies yield the same result.[27]

When colleges' choice functions satisfy IRC, substitutes, and the law of aggregate demand, a direct mechanism that produces the student-optimal stable outcome is strategy-proof (Aygün and Sönmez, 2013; Hatfield and Kojima, 2010). That is, submitting her true preference ranking over contracts is a weakly dominant strategy for every student. One may be

[26] Note that these include the null option $\emptyset$, and therefore contracts might be deemed unacceptable even when the ranking submitted is supposed to include all options.

[27] This might come as a surprise in light of Ergin (2000), who shows that *consistency* — a property that is indirectly related to this order independence — would be incompatible with stable rules. One important difference is that in IDAM, in every step before the last, allocations are *tentative*. Therefore, different orders in which the proposals are made do not restrict the allocations as in the cases considered in that paper.

tempted to conclude that this will imply that straightforward strategies, which are the equivalent of truth-telling in this dynamic setting, are also dominant under the IDAM mechanism. However, the proposition below shows that not only is this not the case, but that the students may not have any dominant strategy at all.

**Proposition 2.** *A student may not have a weakly dominant strategy under the IDAM mechanism.*

Not following a straightforward strategy may be profitable because, in contrast to the direct mechanism, an agent may influence others' actions by modifying the signals they receive through the interactions with the procedure. So, for example, if a student has a strategy that depends in some way on the sets of tentative allocations, or even on the sequence of menus that are presented or the timing of the rejection in a particular choice, that fact could be exploited. However, we will show that the profile in which students follow straightforward strategies constitutes an intuitive equilibrium.

The solution concept that we use for the equilibrium is that of ex-post equilibrium. A type-strategy profile $(\Sigma_s)_{s\in S}$ is an **ex-post equilibrium** if for every student $s$, preference $P_s$ for $s$, $P_{-s}$ for the remaining students, and strategy $\sigma'$, the outcome obtained by $s$ under the strategy profile $(\Sigma_s(P_s), \Sigma_{-s}(P_{-s}))$ is at least as good for $s$ as the outcome she obtains under the profile $\left(\sigma', \Sigma_{-s}(P_{-s})\right)$. That is, a profile is an ex-post equilibrium if, given a student's preferences, the strategy in its type-strategy is a best-response regardless of the preferences of the other players. Implementation in ex-post equilibrium is considered a robust implementation notion, since students' optimal behavior does not depend on unrealistic assumptions about their knowledge about other players (Bergemann and Morris, 2005).

Our main result shows that, when facing this game, students following straightforward strategies constitute an ex-post equilibrium.[28]

**Theorem 1.** *Assume that, for every college $c \in C$, $f_c$ satisfies IRC and contracts are substitutes. Consider a maximum rank function $\pi$ and the unbounded IDAM mechanism using it. Let $\Sigma^*$ be a type-strategy profile in which all strategies are straightforward with respect to $P_S$. Then $\Sigma^*$ is an ex-post equilibrium of the game induced by the IDAM mechanism.*

The proof of Theorem 1 is based on the following Lemma.

**Lemma 2.** *Consider a maximum rank function $\pi$ and the unbounded IDAM mechanism using it, and $(P_S)$ be a preference profile, and $s$ a student in $S$. If every student $s' \neq s$ follows a straightforward strategy $\sigma^*_{s'}$ with respect to $P_{s'}$, then for any strategy $\sigma_s$ for $s$, there exists a preference $P^*_s$ such that the outcome of the IDAM mechanism given $(P_S)$ is the student-optimal stable matching under preferences $(P^*_s, P_{-s})$, where $P^*_s$ might be different from $P_s$.*

That is, when other players follow straightforward strategies, any deviating strategy is outcome-equivalent to straightforward strategies for some preference. Lemma 2 relies on the fact that although the space of deviating strategies is significantly large, they are all indistinguishable, from the perspective of an observer, from a student following a straightforward strategy for some different preference. This allows us to evaluate deviating strategies in all of their paths, which may include multiple interactions that the student may have with the mechanism. Without this, it would be difficult to determine the final outcome of a generally specified deviation.[29]

If we narrow our attention to the IDAM mechanism when students can choose only one contract from each menu that is given, we are able to show that straightforward strategies constitute the unique equilibrium in which strategies satisfy a mild and reasonable restriction.

More specifically, we say that a strategy $\sigma_s$ for student $s$ satisfies **no irrelevant participation** if for every period $t$ in which $s$ is given a non-empty menu $\psi^t$ containing only unacceptable contracts, $\sigma_s$ determines that she will choose $\emptyset$. In light of the description of the IDAM mechanism and Lemma 1, a student facing a menu containing only unacceptable contracts can never obtain a matching that she prefers to the option of remaining unmatched. This is, therefore, a very mild assumption that eliminates strategies in which students choose to take the risk of being matched to unacceptable contracts, hoping that they will end up unmatched, instead of simply choosing to remain unmatched. Notice that no strategy outside of this domain could ever result in a better outcome for any student. This condition allows us to obtain the result below.

**Theorem 2.** *Assume that, for every college $c \in C$, $f_c$ satisfies IRC and contracts are substitutes. Consider the unbounded IDAM mechanism with a maximum rank function $\pi$ where $\pi(t) = 1$ for every $t$, and assume that students can only use strategies that satisfy no irrelevant participation. A type-strategy profile $(\Sigma_s)_{s\in S}$ is an ex-post equilibrium if and only if for every $s \in S$, $\Sigma_s$ is straightforward.*

That is, for a very mild and sensible restriction on preferences, the IDAM mechanism fully implements the student-optimal stable matching with a simple and intuitive equilibrium strategy: simply choose your most preferred contract

[28] Mackenzie and Zhou (2020) also show that straightforward strategies constitute an ex-post equilibrium in *menu mechanisms* implementing the student-optimal stable outcome. In our setup, this corresponds to the unbounded IDAM mechanism where $\pi(t) = 1$ for every $t$.

[29] In fact, in most sequential matching mechanisms in the literature (for example, Alcalde and Romero-Medina (2005), Triossi (2009), and Romero-Medina and Triossi (2014)) the number of times an agent interacts with the mechanism is either exogenously given or equals one in equilibrium.

anytime a menu is given, regardless of other students' preferences. Notice that if $\pi(t) > 1$ this uniqueness cannot be guaranteed. For example, when the IDAM mechanism implements a serial dictatorship, any strategy for the highest priority student where her most preferred college is at the top of the list is a best-response. Thus, all these non-straightforward strategies with non-truthful rankings below their top choice will be part of the resulting multiple ex-post equilibria.

### *4.1. IDAM in centralized college admissions*

A special case of the IDAM that is especially relevant is the one in which colleges evaluate students based on exam grades, and after each step students are informed about the values of cutoff grades for each college.

More specifically, in a centralized college admissions problem, there is a list of **vectors of exam scores** $z = (z(s_1), \ldots, z(s_n))$, where for each $s \in S$, $z(s) = \left(z_{c_1}(s), \ldots, z_{c_m}(s)\right)$, are the exam scores that student $s$ obtained, respectively, at college $c_1 \ldots, c_m$. We assume that for every $s, s' \in S$ and $c \in C$, $z_c(s) = z_c\left(s'\right) \implies s = s'$. Colleges also have **capacities** $\left(q_{c_1}, \ldots, q_{c_m}\right)$ and **minimum necessary scores** $\underline{Z} = \left(\underline{z}_{c_1}, \ldots, \underline{z}_{c_m}\right)$. The set of **valid contracts** is $X = \left\{(s, c, z_c(s)) : s \in S, c \in C \text{ and } z_c(s) \geq \underline{z}_c\right\}$.

Colleges' choice functions $F_C$ are such that given a set of valid contracts, college $c$ chooses the ones with the highest exam scores, up to its capacity $q_c$. In this setup, therefore, the definition of stability implies that an outcome $I \subseteq X$ is stable if it is individually rational and there is no student-college pair $(s, c)$ such that $c\ P_s\ c(I_s)$, $z_c(s) \geq \underline{z}_c$, and either $(i)$ $|I_c| < q_c$ or $(ii)$ there is a student $s' \in s(I_c)$ for which $z_c(s) > z_c(s')$.

Notice, given that under the IDAM mechanism students are informed about the tentative allocations after each step, we can also consider the idea of them observing the (tentative) cutoff values for each of these colleges. That is, students can be assumed to observe the minimum required exam grade at each college, given the current set of contracts being held. Finally, we let $T^{Max} = \infty$ and for all $t$, $\pi(t) = 1$. That is, in each step in which a student is given a menu, he can *choose* a college (or $\emptyset$) from it.

Notice that, given our assumptions colleges' choice functions trivially satisfy substitutes, IRC, and the law of aggregate demand. This implies that Proposition 1 and Theorem 1 also hold for this application of the IDAM mechanism.

Let $\zeta_c^t$ be the value of the cutoff of college $c$ at step $t$, as defined above. In light of the definition of the colleges' choice functions, Lemma 1 leads to the following conclusion:

**Corollary 1.** *(Cutoff grades never go down) For every* $1 \leq t \leq T^*$ *and* $c \in C$, $\zeta_c^t \geq \zeta_c^{t-1}$.

## 5. Comparing mechanisms

In section 3, we show that the Brazilian and Inner Mongolia mechanisms, which have some clear similarities with instances of the IDAM mechanism, have some shortcomings. In this section, we compare the outcomes and incentives of these mechanisms, and see whether the IDAM mechanism improves upon them in these dimensions. We will, therefore, consider the IDAM mechanism applied for centralized college admissions introduced in section 4.1.

In this section, we will focus on the Brazilian mechanism when describing the arguments and discussing the results, but unless stated otherwise, all of them will also apply for the Inner Mongolia mechanism—when we assume there is a minimum time that it takes for students to revise their choices.

We present our first observation simply as a remark:

**Remark 6.** Consider a centralized college admissions problem, and let $T^*$ be the last step of the IDAM mechanism when all students follow straightforward strategies. Then, the outcome of the Brazilian mechanism when it runs for $T^*$ days and students follow the straightforward behavior is also the student-optimal stable matching.

The remark above should come as no surprise.[30] Students following the straightforward behavior in the Brazilian mechanism will simply reproduce the steps of the IDAM mechanism, and therefore also its outcome if given enough time.

In order to make a reasonable comparison, the differences in incentives and outcomes that we consider should not rely simply on the fact that the number of periods in which students can revise their choices in the Brazilian Mechanism is not "large enough." Therefore, we will only consider instances of the Brazilian mechanism with a number of days large enough so that it is no smaller than $T^*$ as defined above for any problem that might be presented. Notice that the game induced by the Brazilian mechanism is relatively simple: in every period, every student can apply to any college. Outcomes depend only on their choices on the last day. Their strategies, however, can condition their choices on each day to the values of the cutoffs on previous days.

Our next result is also given as a remark:

**Remark 7.** A student might obtain strictly better outcomes in the Brazilian mechanism by not following the straightforward behavior, when other players are following the straightforward behavior.

30 In fact, Gong and Liang (2016) present a similar result when evaluating the Inner Mongolia mechanism.

The above remark can be directly derived from Example 1, where a student postpones her application to a college to prevent a rejection cycle from reaching her back. It implies that students following the straightforward behavior are not a Nash equilibrium, and therefore also not an ex-post equilibrium. In fact, since student $s_1$ has the incentive to postpone her application to $c_1$, student $s_3$ could anticipate her proposal to $c_1$ as well, guaranteeing her place in that college. In other words, the presence of a rejection chain and the ability to postpone it *unravels* a configuration in which students can safely follow the straightforward behavior. Notice, however, that as pointed out in section 3.1.1, since in the Inner Mongolia mechanism students are assumed to have the same grade in every college, there can be no rejection cycles, and therefore Remark 7 does not apply to it.

Perhaps the most prominent strategic aspect of the Brazilian mechanism is that, as long as choices made on the last day remain the same, changes that take place on previous days have no impact whatsoever on the outcomes. This turns every period before the last one essentially into periods in which students engage in cheap talk. As a result, equilibrium outcomes are fundamentally determined by strategic interactions in the last period, and therefore to a multiplicity of equilibrium outcomes, as shown in the proposition below.

**Proposition 3.** *Consider a centralized college admissions problem, and let $\mathcal{X}^{Stab}$ be the set of stable allocations. Let $\mathcal{X}^{NE}$ and $\mathcal{X}^{SPNE}$ be, respectively, the set of Nash equilibrium and Subgame-Perfect Nash equilibrium outcomes of the Brazilian mechanism. Then, $\mathcal{X}^{Stab} = \mathcal{X}^{NE} = \mathcal{X}^{SPNE}$.*

Remark 7 and Proposition 3 show, therefore, a contrast between equilibrium strategies and outcomes between the IDAM mechanism and the ones currently used in Brazil and Inner Mongolia. While the simple and intuitive straightforward strategy constitutes an ex-post equilibrium in IDAM, this might not be the case in the Brazilian mechanism, where students might have to anticipate harmful manipulations by other students, unraveling the straightforward behavior. Moreover, while under a mild selection of strategies, the straightforward strategies are the unique equilibrium strategy profiles in IDAM, any stable outcome is an equilibrium for the Brazilian and Inner Mongolia mechanisms, making the outcome a result of a more complex coordination process.

When it comes to the dynamic harmful manipulations that we identified in section 3.1, IDAM also represents an improvement, as shown in the proposition below.

**Proposition 4.** *The IDAM mechanism is immune to any individual harmful manipulations, and is immune to manipulations via cutoffs.*

## 6. Conclusion

In this paper we introduced the IDAM, an iterative mechanism in which participants are sequentially asked to make choices or submit rankings over menus they are given, and produce stable outcomes in a simple truthful equilibrium.

The IDAM mechanism is flexible in that it allows for different combinations of the number of choices, rankings, and number of steps, but also share with current procedures the dynamic and more transparent way in which the matchings are created.

The IDAM mechanism, and the special cases we introduced, improve upon the shortcomings that we identified in current college admissions mechanisms in Germany, Brazil, and Inner Mongolia (China).

While under the German Mechanism accepting offers during the dynamic stage may lead to wastefulness or justified envy (Remark 1), simply choosing the most preferred college at those steps in any IDAM is part of an ex-post equilibrium that results in a matching without any justified envy or wastefulness. Moreover, unlike the German Mechanism, under the IDAM+DA mechanism students submitting truthful rankings after the dynamic steps is also part of that equilibrium (Remark 2 and Theorem 1), which can also be made essentially unique (Theorem 2).

While the Brazilian and Inner Mongolian mechanisms produce unreliable information about the possibility of being accepted at different colleges through cutoff values (Remark 3), under the IDAM mechanism students can safely ignore colleges that are not within reach at any step (Corollary 1). Finally, unlike these two mechanisms (Remark 5), the IDAM mechanism is not subject to manipulations via cutoff, via postponing rejection cycles, or other individual harmful manipulations (Proposition 4).

We believe that there are still many paths to explore on the subject of iterative stable mechanisms. One of them is to extend the model to use statistic information that the policymaker may have about students' preferences, and optimize the mechanism accordingly. For example, if it is known that a large proportion of the students will have a certain college high in their preferences, an "adaptive" IDAM mechanism could start with a higher initial value for the cutoff at that college. In this case, we conjecture that a stable matching would still be reached, with a reasonably high probability.

Another related question involves the design of optimal menus that minimize the amount of information requested from the students, based on the known grade distribution. When grades are common, for example, the IDAM mechanism may obtain information on the preferences that low-grade students have over "top" colleges, but if high-grade students are asked for their preferences earlier, it would not be necessary for this information to be revealed. Ideas similar to these are explored in Ashlagi et al. (2020).

## Declaration of competing interest

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

## Appendix A

### A.1. Formal definitions

#### A.1.1. The iterative deferred acceptance mechanism (IDAM)

For any $I \subseteq X$, define $\mathcal{A}_c(I) \equiv \{x \in X \backslash I : x \in f_c(I \cup \{x\})\}$ and $\mathcal{A}_c^s(I) \equiv \mathcal{A}_c(I) \cap X_s$. We denote by $\mathcal{A}_c^s(I)$ the set of **available contracts for** $s$ **at** $c$ **under** $I$. The interpretation of this is simple: a contract is available for student $s$ at $c$ under $I$ if college $c$ chooses to accept that contract while holding the set of contracts $I_c$.

Consider a college matching with contracts market $\langle S, C, T, X, P_S, F_C\rangle$, a maximum number of steps $T^{Max} \in \mathbb{N} \cup \{\infty\}$, and a **maximum rank function** $\pi : \mathbb{Z}^+ \to \mathbb{N} \cup \{\infty\}$. The **iterative deferred acceptance mechanism** (**IDAM**) proceeds as follows:

- **Step** $t = 0$: Let $\mathcal{L}^0 = S$, $S^0 = \emptyset$, and for every $c \in C$, $A^0(c) = \emptyset$.[31]
- **Step** $0 < t \le T^{Max}$:

  **[*Menus Setup*]** Let $S^t \equiv \left\{s \in \mathcal{L}^{t-1} | \nexists x \in X_s, c \in C : x \in f_c\left(A^{t-1}(c)\right)\right\}$. There are two cases[32]:
  
  ($i$) If $\pi(t) \neq \infty$ or $T^{Max} = \infty$, for every $s \in S$, let the menu of contracts presented to $s$ be $\psi_s^t \equiv \bigcup_{c\in C} \mathcal{A}_c^s\left(A^{t-1}(c)\right) \cup \{\emptyset\}$ if $s \in S^t$ and $\psi_s^t = \emptyset$ otherwise. Request each student $s \in S^t$ to submit a ranking with at most $\pi(t)$ elements in $\psi_s^t$.
  
  ($ii$) If $\pi(t) = \infty$ and $T^{Max} < \infty$, for every $s \in S$, let the menu of contracts presented to $s$ be $\psi_s^t \equiv \bigcup_{c\in C} \mathcal{A}_c^s\left(A^{t-1}(c)\right) \cup \{\emptyset\}$. Request that each student $s \in \mathcal{L}^{t-1}$ submits a ranking of any size of elements in $\psi_s^t$.

  **[*Cumulative Offer*]** Let, for every student $s'$, $P_{s'}^t$ be the ranking submitted. For every student $s''$ such that $\psi_{s''}^t = \emptyset$, let $P_{s''}^t = P_{s''}^{t-1}$ and, for all $c \in C$, $B^0(c) \equiv A^{t-1}(c)$. Start with $\tau = 0$ and let $\mathcal{L}^t = \mathcal{L}^{t-1}$.
  - Substep $\tau \ge 0$: Some student $s$ in $\mathcal{L}^{t-1}$, who does not have a contract held by any college, proposes her most preferred contract with respect to $P_s^t$, which has not yet been rejected, $x$. If $x = \emptyset$, remove $s$ from $\mathcal{L}^t$ and from further consideration.[33] Otherwise, college $c(x)$ *holds* $x$ if $x \in \mathcal{A}_c\left(B^\tau\right)$, and *rejects* $x$ if $x \notin \mathcal{A}_c\left(B^\tau\right)$. Let $B^{\tau+1}(c) = B^\tau(c) \cup \{x\}$ and for all $c' \neq c$, $B^{\tau+1}\left(c'\right) = B^\tau\left(c'\right)$.[34]
  - Repeat the process above until no student is able to propose a new contract. Let $\tau^*$ be the last step in that process.

  **[*Update before next step*]** For each college $c$, let $A^t(c) = B^{\tau^*}(c)$. If for every $c \in C$ it is the case that $A^t(c) = A^{t-1}(c)$, stop the procedure. Otherwise, broadcast the value of $\bigcup_{c\in C} A^t(c)$, and proceed to the next step.

- Denote by $T^*$ the last step executed in the procedure. Let $X^* = \bigcup_{c\in C} f_c\left(A^{T^*}(c)\right)$. $X^*$ is the outcome of the IDAM procedure.

#### A.1.2. Equilibrium concept

Let $\mathcal{X}^{\mathbb{N}}$ be the set of finite sequences of subsets of $X$, and $\mathcal{P}_s$ be the set of all strict preferences over $X_s \cup \{\emptyset\}$. In general, a **strategy** of student $s$ is a function $\sigma_s : \mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}} \to \mathcal{P}_s$. For a student $s$, the value of $\sigma_s(\langle A^1, A^2, \ldots, A^{t-1}\rangle, \langle \psi_s^1, \psi_s^2, \ldots, \psi_s^t\rangle)$ is the ranking with at most $\pi(t)$ contracts in $\psi_s^t$ that $s$ will submit in period $t$ when facing the menu $\psi_s^t$, after previously receiving the menus $\psi_s^1, \psi_s^2, \ldots, \psi_s^{t-1}$ (some of them possibly empty), and observing the sequence of contracts held by colleges $A^1, A^2, \ldots, A^{t-1}$. For a given IDAM mechanism, we restrict strategies to having as domain only sequences of menus and of contracts held by colleges that can in fact result from its execution, given the set of contracts $X$ and colleges' choice functions. Denote by $\mathcal{O}\left((\sigma_s)_{s\in S}\right)$ the outcome produced by the IDAM mechanism when each student $s$ follows the strategy $\sigma_s$.

**Definition 5.** A strategy $\sigma_s$ of student $s \in S$ is **straightforward with respect to** $P^*$ if there is a function $k : 2^X \times \mathbb{N} \to \mathbb{N}$ such that for every $t$, $1 \le k(\cdot, t) \le \pi(t)$, and $k(\cdot, t^\infty) = \left|\psi_s^{t^\infty}\right|$, such that for every $\langle A^1, A^2, \ldots, A^{t-1}\rangle \in \mathcal{X}^{\mathbb{N}}$, and

[31] Notation clarification: $\mathcal{L}^t$ is the set of students who are still active at the beginning of step $t$, $S^t$ is the set of students who are active and do not have any contracts held by a college at the beginning of that step, and $A^t(c)$ is the set of contracts that were proposed to college $c$ up to step $t$.

[32] These correspond to cases ($i$) and ($ii$) in the informal description in section 4.

[33] Submitting $\emptyset$ as the most preferred contract is interpreted as opting not to submit a ranking.

[34] Notice that a contract being in $B^\tau(c)$ does not imply that the contract is held by college $c$. It simply means that it was offered to $c$ at some step.

$\langle \psi_s^1, \psi_s^2, \ldots, \psi_s^t \rangle \in \mathcal{X}^{\mathbb{N}}$, $\sigma_s(\langle A^1, A^2, \ldots, A^{t-1} \rangle, \langle \psi_s^1, \psi_s^2, \ldots, \psi_s^t \rangle)$ consists of the $k\left(\psi_s^t, t\right)$ most preferred contracts in $\psi_s^t$, ordered according to $P^*$.

A **type-strategy** of student $s$ is a function $\Sigma_s$ that maps each preference ranking in $\mathcal{P}$ to a strategy.

**Definition 6.** A type-strategy profile $(\Sigma_s)_{s \in S}$ is an **ex-post equilibrium** if for every $s \in S$, strategy $\sigma_s'$ for student $s$ and $(P_s, P_{-s}) \in \mathcal{P}^{|S|}$:

$$\left[\mathcal{O}\left(\Sigma_s(P_s), (\Sigma_{s'}(P_{s'}))_{s' \in S \setminus \{s\}}\right)\right]_s \; R_s \; \left[\mathcal{O}\left(\sigma_s', (\Sigma_{s'}(P_{s'}))_{s' \in S \setminus \{s\}}\right)\right]_s$$

*A.2. Proofs*

*Lemma 1*

**Proof.** First, note that if $\psi_s^{t'} = \emptyset$, then the statement is true. Suppose, for the sake of contradiction, that $\psi_s^{t'} \neq \emptyset$ and the statement is false. Then, there is a student $s \in S$, $1 \leq t \leq t' \leq T^*$ and a contract $x^* \in X$ such that $x^* \in \psi_s^{t'}$ but $x^* \notin \psi_s^t$. Since for any $I \subseteq X$, $\mathcal{A}_c(I) \subseteq X_c$, we can separate the violation of the lemma into the contracts available for a student from a single college. There is, therefore, a college $c$ such that:

$$x^* \in \mathcal{A}_c^s\left(A^{t'}(c)\right) \text{ but } x^* \notin \mathcal{A}_c^s\left(A^t(c)\right)$$

Given the definition of $\mathcal{A}_c^s$:

$$x^* \in f_c\left(A^{t'}(c) \cup \left\{x^*\right\}\right) \text{ but } x^* \notin f_c\left(A^t(c) \cup \left\{x^*\right\}\right)$$

For simplicity of notation, denote by $A_s^t$ and $A_{-s}^t$ the set of contracts in $A^t(c)$ involving student $s$ and involving any student but $s$, respectively. We can, therefore, rewrite the expression above as:

$$x^* \in f_c\left(A_s^{t'} \cup A_{-s}^{t'} \cup \left\{x^*\right\}\right) \text{ but } x^* \notin f_c\left(A_s^t \cup A_{-s}^t \cup \left\{x^*\right\}\right) \tag{A.1}$$

Since $f_c$ selects at most one contract per student, IRC implies that we can remove any contracts in $A_s^{t'}$ from the first expression. Moreover, $A^t(c) \subseteq A^{t'}(c)$ by construction, and therefore:

$$x^* \in f_c\left(Y \cup A_{-s}^t \cup \left\{x^*\right\}\right) \text{ but } x^* \notin f_c\left(A_s^t \cup A_{-s}^t \cup \left\{x^*\right\}\right) \tag{A.2}$$

Where $Y \equiv A_{-s}^{t'} \backslash A_{-s}^t$. There are two cases to consider: $(i)$ $f_c\left(A_s^t \cup A_{-s}^t \cup \{x^*\}\right) \cap A_s^t = \emptyset$, and $(ii)$ $f_c\left(A_s^t \cup A_{-s}^t \cup \{x^*\}\right) \cap A_s^t = y^*$. That is, the choice function in the second expression above selects—case $(ii)$—or not— case $(i)$—a contract with student $s$.

In case $(i)$, IRC implies that $A_s^t$ can be removed from the second expression, and (A.2) becomes:

$$x^* \in f_c\left(Y \cup A_{-s}^t \cup \left\{x^*\right\}\right) \text{ but } x^* \notin f_c\left(A_{-s}^t \cup \left\{x^*\right\}\right)$$

This contradicts the substitutes condition for $f_c$. Next, consider case $(ii)$. We can rewrite (A.1) as:

$$x^* \in f_c\left(A_s^{t'} \cup A_{-s}^{t'} \cup \left\{x^*\right\}\right) \text{ but } x^* \notin f_c\left(A_s^t \cup A_{-s}^t \cup \left\{x^*, y^*\right\}\right) \tag{A.3}$$

Since $A_s^t \subseteq A_s^{t'}$, $y^* \in A_s^{t'}$. Applying IRC, we can remove every contract in $A_s^{t'} \backslash \{y^*\}$. Let, moreover, $Y$ be defined as in (A.2). We obtain:

$$x^* \in f_c\left(Y \cup A_{-s}^t \cup \left\{x^*, y^*\right\}\right) \text{ and } x^* \notin f_c\left(A_{-s}^t \cup \left\{x^*, y^*\right\}\right) \tag{A.4}$$

This also contradicts the substitutes condition for $f_c$, finishing the proof. □

*Proposition 1*

**Proof.** First, note that given the description of an unbounded IDAM mechanism and Lemma 1, every time a student is asked to submit a ranking, the set of contracts available under a IDAM mechanism is weakly smaller. Moreover, for any $t, t'$ and $s \in S$ such that $1 \leq t < t' \leq T^*$, $\psi_s^t \neq \emptyset$ and $\psi_s^{t'} \neq \emptyset$, it must be that the set of contracts in $\psi_s^{t'}$ is a strict subset of $\psi_s^t$, since at least the highest-ranked contract submitted by student $s$ at step $t$ must have been rejected by step $t'$. Therefore, at every step the set $\psi_s^t$ is strictly smaller for at least one student. Since $X$ is finite, IDAM will end and will produce an outcome after a finite number of steps.

Next, notice that regardless of which straightforward strategy students use, in all of them students will offer contracts following the order of their preferences, perhaps only skipping those that would not be held by the college associated with the contract, and that the outcome will be produced when every student either chooses $\emptyset$, has a contract held by a college, or reaches the end of the last ranking submitted. According to Hirata and Kasuya (2014), under our assumptions, the cumulative offer process produces the student-optimal stable matching regardless of the order in which students are called to offer contracts, as long as the order in which each student offers her contracts follows their preferences over them. Different straightforward strategies may imply different orders in which students offer contracts, but this does not change the fact that students follow their own preference until the end.[35] Therefore, for any profile of straightforward strategies, the outcome of the IDAM mechanism will always be the student-optimal stable matching. □

*Proposition 2*

**Proof.** For this proof we consider an exam-based college matching market and the IDAM mechanism with $\pi(t) = 1$ for all $t$. Consider the set of students $S = \{s_1, s_2, s_3\}$ and of colleges $C = \{c_1, c_2, c_3\}$, each with capacity $q_i = 1$. Student $s_1$—who will be the player we show as having no dominant strategy—has preferences $c_1 P_{s_1} c_2 P_{s_1} c_3$, and students' exam grades at those colleges are as follows:

| | $c_1$ | $c_2$ | $c_3$ |
|---|---|---|---|
| $s_1$ | 100 | 100 | 100 |
| $s_2$ | 200 | 200 | 200 |
| $s_3$ | 300 | 300 | 300 |

Suppose now that, conditional on the realized preferences and grades of student $s_1$, student $s_3$ follows a straightforward strategy with respect to the preference $c_3 P^3 c_2 P^3 c_1$. Notice that we are not stating that those are the preferences of student $s_3$, we are simply assuming that she will follow the *straightforward strategy with respect to* $P^3$. Next, we consider two strategies for student $s_2$ and show that no strategy for $s_1$ is a common best-response for these two possibilities.

**Scenario 1**

Suppose that student $s_2$'s strategy is as follows: in $t = 1$, choose $c_3$. If at some later point $s_2$ is asked again to make a choice, she will choose the college with the highest cutoff value at that step among the options available. In the event of a tie, she will choose the college with the lowest index number (for example, the index number of $c_2$ is 2). We will show that, given $s_2$ and $s_3$'s strategies, the best-response for $s_1$ involves first choosing $c_2$. The sequence of steps will be as follows:

**Step 1**: Student $s_1$ applies to $c_2$. Students $s_2$ and $s_3$ apply to $c_3$. Student $s_2$ is rejected. Cutoffs $\left(\zeta^1_{c_1}, \zeta^1_{c_2}, \zeta^1_{c_3}\right)$ are $(0, 100, 300)$.

**Step 2**: Since $\zeta^1_{c_2}$ is the highest cutoff among the colleges offered to $s_2$, student $s_2$ applies to $c_2$. Student $s_1$ is rejected. Cutoffs $\left(\zeta^2_{c_1}, \zeta^2_{c_2}, \zeta^2_{c_3}\right)$ are $(0, 200, 300)$.

**Step 3**: Student $s_1$ is left with two options: choose $c_1$ or $s$. If she chooses $s$, she will remain unmatched. If she applies to $c_1$, she will be accepted. Final cutoffs $\left(\zeta^3_{c_1}, \zeta^3_{c_2}, \zeta^3_{c_3}\right)$ would then be $(100, 200, 300)$ and the outcome would be the matching $\mu'$ as follows:

$$\mu = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_2 & s_3 \end{pmatrix}$$

Student $s_1$ can therefore be matched to her most preferred college by first choosing $c_2$. Next, we show that by first choosing $c_1$ or $c_3$, $s_1$ will always be matched to a strictly inferior college. First, let her initially choose $c_1$:

**Step 1**: Student $s_1$ applies to $c_1$. Students $s_2$ and $s_3$ apply to $c_3$. Student $s_2$ is rejected. Cutoffs $\left(\zeta^1_{c_1}, \zeta^1_{c_2}, \zeta^1_{c_3}\right)$ are $(100, 0, 300)$.

**Step 2**: Since $\zeta^1_{c_1}$ is the highest cutoff among the colleges offered to $s_2$, student $s_2$ applies to $c_1$. Student $s_1$ is rejected. Cutoffs $\left(\zeta^2_{c_1}, \zeta^2_{c_2}, \zeta^2_{c_3}\right)$ are $(200, 0, 300)$.

**Step 3**: Student $s_1$ is left with two options: choose $c_2$ or $s$. If she chooses $s$ she will remain unmatched. If she applies to $c_2$, she will be accepted. Final cutoffs $\left(\zeta^3_{c_1}, \zeta^3_{c_2}, \zeta^3_{c_3}\right)$ would then be $(200, 100, 300)$ and the outcome would be the matching $\mu'$ as follows:

$$\mu' = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_1 & s_3 \end{pmatrix}$$

If $s_1$ chooses $c_3$ first instead, the following will happen:

[35] Technically speaking, under the cumulative order process students will always offer contracts following their preferences, even those that would not be accepted by the college in the contract. Since choice functions satisfy IRC, however, this is equivalent to a process that simply skips those contracts that would not be accepted (and are, therefore, not part of the menus offered to the students under the IDAM mechanism).

**Step 1**: Students $s_1$, $s_2$, and $s_3$ apply to $c_3$. Students $s_1$ and $s_2$ are rejected. Cutoffs $\left(\zeta^1_{c_1}, \zeta^1_{c_2}, \zeta^1_{c_3}\right)$ are $(0, 0, 300)$.

**Step 2**: Following her strategy and the fact that college $c_1$'s index is lower than $c_2$, student $s_2$ applies to $c_1$. Student $s_1$ has three options: also choose $c_1$ and therefore be rejected and left to choose between $c_2$ and $s$ at step $t = 2$, choose $c_2$, or choose $s$. In all cases she will end up either remaining unmatched or matched to $c_2$.

**Scenario 2**

Now, suppose that student $s_2$ follows a similar strategy to scenario 1, but where instead of applying to $c_3$ and then to the college with the highest cutoff value, she applies to the college with the *lowest* cutoff value, once again breaking ties based on the index of the college.[36] Following an exercise similar to the one above, it is easy to see that student $s_1$'s strategies that involve choosing $c_2$ or $c_3$ first will lead to her either being unmatched or matched to $c_2$, while choosing $c_1$ will match her to $c_1$, her most preferred college.

Since every best-response strategy under scenario 1 is dominated by different strategies in scenario 2, we show that a student may not have a weakly dominant strategy for the game induced by the IDAM mechanism, and as a consequence also the IDAM mechanism. □

*Lemma 2*

**Proof.** Consider some period $t$ during the execution of the IDAM mechanism. Given other players' strategies $\sigma^*_{-s}$, all of them straightforward, the sequence of events that results from the strategy profile $\left(\sigma_s, \sigma^*_{-s}\right)$ consists of a series of steps in which each student has either no menu given to her or some menu of options $\psi^t_s$ and a maximum rank $\pi(t)$, as described in the definition of the IDAM. Therefore, given our strategy profile and student $s$, we can write down a list of pairs, with menus given to student $s$ and her submitted ranking.

Suppose that the sequence of menus offered and actions taken (i.e., rankings submitted) for a student $s$ up to period $t$ are as follows:

$$\left(\left(\psi^1_s, a^1\right), \left(\psi^2_s, a^2\right), \ldots, \left(\psi^t_s, a^t\right)\right)$$

We show below that menus given to students never include contracts present in any previously submitted ranking.

**Claim.** *If contract $x$ is in $a^t$, then for every $t'$ such that $t' > t$, $x \notin \psi^{t'}_s$.*

**Proof.** Let $c = c(x)$. If $\psi^{t'}_s = \emptyset$, the claim obviously holds. Therefore, we consider the case in which both $\psi^t_s$ and $\psi^{t'}_s$ have a positive number of contracts available. Since $x$ is in $a^t$, $x \in \psi^t_s$. Also, since $x \in a^t$ and the fact that $\psi^{t'}_s \neq \emptyset$, it must be the case that, for any $c$, all the contracts in $a^t$ that involve $c$ are in $A^{t'-1}(c)$ (otherwise the IDAM mechanism would still use the ranking $a^t$ at step $t'$). Also, under the definition of the IDAM, $\nexists y \in X_s, c \in C : y \in f_c\left(A^{t'-1}(c)\right)$, and in particular $x \notin f_c\left(A^{t'-1}(c)\right)$. Therefore, $x \notin \psi^{t'}_s$. □

Therefore, there is no repetition of contracts in $a^i$, $i = 1, \ldots, t$. We will abuse notation and use $a^i$ to represent the student's choice both as a ranking and as a set of contracts. Denote $\psi^i_{s-} \equiv \psi^i_s \setminus \bigcup_{j=i}^t a^j$ and $X^+_s \equiv X_s \setminus \{\emptyset\}$. We will show that this sequence could have been generated by a student's straightforward strategy with a preference relation in the following class of preferences[37]:

$$X^+_s \setminus \psi^1_s \; R^*_s \; a^1 \; P^*_s \; \psi^1_{s-} \setminus \psi^2_{s-} \; R^*_s \; a^2 \; P^*_s \; \psi^2_{s-} \setminus \psi^3_{s-} \; R^*_s \cdots R^*_s \; a^t \; P^*_s \; \psi^t_{s-}$$

The notation above includes a class of strict preferences because some of its elements ($X^+_s \setminus \psi^1_s$, $\psi^1_{s-} \setminus \psi^2_{s-}$, etc.) consist of (possibly empty) sets of contracts. Any strict preference derived from some ordering over the elements of each of those sets belongs to the class of preferences that we are referring to. We will use $P^*_s$ to refer to some arbitrary element of those preferences. The claim below implies that each preference in that class is complete over the set of contracts and that no contract appears more than once.

**Claim.** *$\psi^t_{s-} \subseteq \psi^{t-1}_{s-} \subseteq \cdots \subseteq \psi^1_{s-} \subseteq X_s$, and $a^i \cap \psi^j_{s-} = \emptyset$ for all $i$, $j$.*

36 Although the strategies used in this proof for student $s_2$ may seem very arbitrary, they can be rationalized by two simple stories. Student $s_2$'s strategy in scenario 1 is consistent with a student who knows that her top choice is $c_3$ but has some uncertainty about whether $c_1$ or $c_2$ is her second choice, and sees the cutoff grade as an indication of how competitive acceptance is at those colleges and therefore sees the perceived quality of those. The strategy in scenario 2 could be rationalized by a student who once again knows that her top choice is $c_3$ but who would otherwise prefer to go to a college with low-achieving peers, and uses the low cutoff as an indication of that fact.

37 Note that this class of preferences does not necessarily include all the preferences that are compatible with the rankings submitted.

**Proof.** First, note that $\psi_{s-}^1 = \psi_s^1 \backslash \bigcup_{j=1}^t a^j$. Since by definition $\psi_s^1$ is non-empty, $a^1 \subseteq \psi_s^1$, and since by definition $\psi_s^1 \subseteq X_s$, it follows that $\psi_{s-}^1 \subset X_s$. Under the definition of $\psi_s^t$ and Lemma 1, $\psi_s^k \subset \psi_s^{k-1}$. Therefore:

$$\psi_s^{k-1} \backslash \bigcup_{j=k-1}^{t} a^j = \left(\left(\psi_s^{k-1} \backslash \psi_s^k\right) \cup \psi_s^k\right) \backslash \left(a^{k-1} \cup \bigcup_{j=k}^{t} a^j\right)$$

Consider now any $k > 1$. By definition, all contracts in $a^{k-1}$ are in $\psi_s^{k-1}$, and according the first claim in this proof, no contract in $a^{k-1}$ is in $\psi_s^k$. Therefore:

$$\psi_{s-}^{k-1} = \psi_s^{k-1} \backslash \bigcup_{j=k-1}^{t} a^j = \left(\left(\left(\psi_s^{k-1} \backslash \psi_s^k\right) \backslash a^{k-1}\right) \cup \psi_s^k\right) \backslash \bigcup_{j=k}^{t} a^j = \psi_{s-}^k \cup \left(\left(\psi_s^{k-1} \backslash \psi_s^k\right) \backslash a^{k-1}\right)$$

That is, $\psi_{s-}^{k-1} = \psi_{s-}^k \cup \left(\left(\psi_s^{k-1} \backslash \psi_s^k\right) \backslash a^{k-1}\right)$, which implies that $\psi_{s-}^k \subseteq \psi_{s-}^{k-1}$. Finally, for every $j \geq i$, it follows from the definition of $\psi_{s-}^i$ that $a^j \cap \psi_{s-}^i = \emptyset$. Suppose instead that there is a $i > j$ such that $a^j \cap \psi_{s-}^i = I$, for some non-empty set of contracts $I$. In that case, the definition of $\psi_{s-}^i$ implies that $I \subseteq \psi_s^i$. But in that case, we have that the contracts in $I$ were submitted in a ranking by the student in step $j$ and were available in the menu at step $i > j$, which contradicts the first claim in the proof. □

Next, take some of the menus that were offered, $\psi_s^i$. We now show that for all $a \in \psi_s^i$ where $a \notin a^i$, $a^i P_s^* a$. For this, it suffices to show that:

$$a \in \bigcup_{j=i+1}^{t} a_j \cup \bigcup_{j=i}^{t-1} \psi_{s-}^j \backslash \psi_{s-}^{j+1} \cup \psi_{s-}^t$$

That is, we will show that $a$ must be at some element to the right of those in $a^i$ in the definition of $P_s^*$. Since $a \notin a^i$, this is equivalent to:

$$a \in \bigcup_{j=i}^{t} a_j \cup \bigcup_{j=i}^{t-1} \psi_{s-}^j \backslash \psi_{s-}^{j+1} \cup \psi_{s-}^t$$

Since we defined $\psi_{s-}^i \equiv \psi_s^i \backslash \bigcup_{j=i}^t a^j$, we can rewrite the condition as:

$$a \in \underbrace{\psi_s^i \backslash \psi_{s-}^i}_{(i)} \cup \underbrace{\bigcup_{j=i}^{t-1} \psi_{s-}^j \backslash \psi_{s-}^{j+1}}_{(ii)} \cup \underbrace{\psi_{s-}^t}_{(iii)}$$

Suppose not. Then $a$ cannot be in $(i)$, $(ii)$, or $(iii)$. According to $(i)$, it must be that $a \notin \psi_s^i \backslash \psi_{s-}^i$. Since $a \in \psi_s^i$, that implies $a \in \psi_{s-}^i$. By $(ii)$, since $a \notin \psi_{s-}^i \backslash \psi_-^{i+1}$, it must then be that $a \in \psi_-^{i+1}$. This reasoning can be repeated until finding that it must be that $a \in \psi_{s-}^t$. But this is $(iii)$, which leads to a contradiction.

The sequence $\left(\left(\psi_s^1, a^1\right), \left(\psi_s^2, a^2\right), \ldots, \left(\psi_s^t, a^t\right)\right)$ is consistent with student $s$ having a preference over contracts $P_s^*$ and following a straightforward strategy that at each step $k \leq t$ submits a ranking with the top $\left|a^k\right|$ contracts among those available, with respect to her preference.

This implies that, since all other students follow straightforward strategies, every deviating strategy for student $s$ is outcome-equivalent to following a straightforward strategy for some preference over contracts that is not necessarily that student's real preference $P_s$. Proposition 1, therefore, shows that the outcome produced will be the student-optimal stable matching with respect to the preference profile $\left(P_s^*, P_{-s}\right)$. □

*Theorem 1*

**Proof.** Hatfield and Milgrom (2005), combined with Aygün and Sönmez (2013), show that since colleges' choice functions satisfy substitutes and the law of aggregate demand, submitting a true ranking is always a best-response when using a direct mechanism. In light of Lemma 2, when other students follow straightforward strategies, any deviating strategy for $s$ is outcome-equivalent to a deviating strategy in the direct mechanism, and as a result is not profitable. Therefore, for any preference profile, given that other players' strategies $\sigma_{-s}^*$ are straightforward, following any straightforward strategy is a best-response for student $s$.

We have therefore concluded that, for any preference profile, a student following the straightforward type-strategy is a best-response to the other students also following the straightforward type-strategy. That is, every student following the straightforward type-strategy is an ex-post equilibrium. □

*Theorem 2*

**Proof.** Under Theorem 1, if $(\Sigma_s)_{s\in S}$ is such that for every $s \in S$, $\Sigma_s$ is straightforward, then $(\Sigma_s)_{s\in S}$ is an ex-post equilibrium.

Consider now a type-strategy profile $(\Sigma_s)_{s\in S}$ that forms an ex-post equilibrium, let $s^*$ be any student in $S$, and let $t$ be any period in which $s^*$ is given a non-empty menu $\psi^t_{s^*}$. Consider now two sets of students: let $\hat{S}$ be the set of all students, except for $s^*$, who were also given a menu in period $t$, and $\bar{S} \equiv S\setminus\left(\hat{S} \cup \{s^*\}\right)$. Consider now the strategy profile $P^*$, where:

1. For every $\hat{s} \in \hat{S}$ and menu $\psi^t_{\hat{s}}$ given to her in period $t$, for every $x \in \psi^t_{\hat{s}}$, $\emptyset\ P^*_{\hat{s}}\ x$.
2. For every $\bar{s} \in \bar{S}$ and last menu $\psi_{\bar{s}}$ given to her and her choice from it $x^* \in \psi_{\bar{s}}$, for every $x \in \psi_{\bar{s}}\setminus\{x^*\}$, $\emptyset\ P^*_{\bar{s}}\ x$.
3. $P^*_{s^*}$ is any preference over $X_s$.

Under Lemma 1 and the definition of the IDAM mechanism, there is no continuation after period $t$ such that the contract matched to $s^*$ is not in $\psi^t_{s^*}$. Moreover, by the way in which menus are constructed, if $s^*$ chooses any contract $x \in \psi^t_{s^*}$ and no other player chooses any contract from college $c(x)$ in any period after $t$, then $s^*$ will end up matched to college $c(x)$ under the contract $x$.

Since players' strategies satisfy no irrelevant participation, every player in $\hat{S}$ will choose the option $\emptyset$ in their menus and remain unmatched. There are two cases to consider. First, the case in which every student who has a contract being held by some college in period $t$ is also held by a college in $t+1$.[38] In this case, no student is given a menu in period $t+1$, and the IDAM mechanism ends, matching all students with contracts held by colleges under these terms, including $s^*$, who is matched to the contract she chose at period $t$. In the second case, there is at least one student who was held by some college in period $t$, but was not anymore at $t+1$. In this case, these students (all of them in $\bar{S}$) are each given a menu which, by construction of $P^*$, contains only unacceptable contracts. With no irrelevant participation, they will all choose $\emptyset$, and therefore by the end of period $t+2$ the IDAM mechanism will end, also matching all students with contracts held by colleges under these terms, including $s^*$, who is matched to the contract she chose in period $t$.

Since under the profile $P^*$, when facing the menu $\psi^t_{s^*}$, $s^*$ ends up matched to any contract she chooses, it must be that in an ex-post equilibrium her type-strategy $\Sigma_{s^*}$ is such that, for any $P_s$, $\Sigma_{s^*}(P_s)$ chooses the most preferred contract in $\psi^t_{s^*}$.

Since the argument above holds for an arbitrary period in which a student is given a non-empty menu, it must hold for any student and any time she is given a non-empty menu. In other words, if an ex-post equilibrium exists, type-strategies must be straightforward, finishing the proof. □

*Proposition 3*

**Proof.** Let $X^*$ be an unstable outcome, and let $(s^*, c^*)$ be a blocking pair, that is, $c^*\ P_{s^*}\ c(I_{s^*})$, $z_{c^*}(s^*) \geq \underline{z}_{c^*}$, and either $(i)$ $|I_{c^*}| < q_{c^*}$ or $(ii)$ there is a student $s' \in s(I_{c^*})$ for which $z_{c^*}(s^*) > z_{c^*}(s')$.

Consider now the last day of the Brazilian mechanism and assume that $X^*$ is the outcome produced given students' strategies. This implies that on the last day, every student $s$ matched to a college under $X^*$ applies to college $c(X^*_s)$, including $s^*$, who applies to $c(I_{s^*})$. If $s^*$, however, changes her strategy to apply to $c^*$ on the last day, she will end up matched to $c^*$, and will therefore obtain a strictly better outcome. For the Inner Mongolia mechanism, let $s^* \in S_k$ — that is, $s^*$ is in the $k$-th tier. We need to consider two cases: $(i)$ $s' \in S_k$ and $(ii)$ $s' \in S_{k'}$ for some $k' > k$. For case $(i)$, the scenario is equivalent to the Brazilian mechanism: if $s^*$ changes her application in the last instant in which a student might make a choice without leaving enough time for other students in $S_k$ to react to $c^*$, she will be matched to $c^*$ and will be strictly better off. In case $(ii)$, the fact that $s'$ is matched to $c^*$ implies that by the end of the time slot in which students in $S_k$ make their final choices, $c^*$ will still have empty seats. This then implies that $s^*$ changing her application without leaving enough time for other students in $S_k$ to react to $c^*$, will be matched to $c^*$ and will be strictly better off. We conclude, therefore, that $X^*$ cannot be either a Nash equilibrium or Subgame Perfect Nash equilibrium outcome, and $X^* \notin \mathcal{X}^{NE}$ and $X^* \notin \mathcal{X}^{SPNE}$.

Let now $X^{**}$ be any stable outcome, and consider a strategy profile, for both mechanisms, in which every student $s$, matched to a college under $X^{**}$, applies to $c(X^{**}_s)$ in every period in which she is able to, regardless of any changes in cutoffs they might observe. Denote that strategy profile by $\sigma$ in both mechanisms. Note that for both mechanisms, the outcome for the strategy profile $\sigma$ is $X^{**}$.

First, consider the Brazilian Mechanism. Since outcomes are produced solely based on the applications made on the last day, and given the construction of $\sigma$, any student $s$ and deviating strategy $\sigma'_s$ involving different choices only on the days before the last one does not change the outcome, since in this case choices made on the last day are not affected. Therefore, $s$ can only affect her outcome with strategies that change her application on the last day. Given the stability of $X^{**}$ and the fact that under $\sigma$ all the other students apply to their outcomes in $X^{**}$ on the last day, no college that $s$ prefers to $c(X^{**}_s)$ would accept her application on the last day. Therefore, $\sigma$ constitutes both a Nash equilibrium and a Subgame-Perfect Nash equilibrium.

[38] That is, no student was rejected as a consequence of the choice made by student $s^*$.

Finally, consider the Inner Mongolia mechanism, some student $s$ and deviating strategy $\sigma_s'$. Let $S_k$ be the tier which $s$ is in. As in the Brazilian mechanism, given the construction of $\sigma$, only deviations that involve changes in the application at the last instant in which a student might make a choice without leaving enough time for other students in $S_k$ to react might change her outcome. Suppose for a contradiction that $\sigma_s'$ is such that $s$ is matched to $c^*$, a college for which $c^* \ P_s \ c(X_s^{**})$. The stability of $X^{**}$ implies that either $(i)$ $z_{c^*}(s) < \underline{z}_{c^*}$— in which case there is no circumstance in which $s$ is matched to $c^*$— or $(ii)$ every student $s' \in s(X_{c^*}^{**})$ is such that $z_{c^*}(s) < z_{c^*}(s')$. For case $(ii)$, given that all of them have grades higher than $s$, then either all of them are in "higher" tiers than $s$, in which case by the time $s$ is matched college $c^*$ has no seats left, or some of them are in $S_k$, but given $\sigma$ they apply to $c^*$ until the last minute and therefore $\sigma_s'$ cannot change $s$'s match to $c^*$. Therefore, no student $s$ has a profitable deviating strategy starting from any period, and therefore $\sigma$ is both a Nash equilibrium and a Subgame-Perfect Nash equilibrium, implying that $X^{**} \in \mathcal{X}^{NE}$ and $X^{**} \in \mathcal{X}^{SPNE}$. □

*Proposition 4*

**Proof.** First, note that manipulations via cutoffs rely on bringing the cutoff value of some college up and later down (in the last period). Corollary 1, however, shows that this can never happen in the IDAM mechanism. Therefore, IDAM is immune to manipulations via cutoffs.

Lemma 2 shows that, for a problem in which preferences are $(P_s, P_{-s})$, when all the other players follow straightforward strategies, any deviating strategy for a student $s$ will make the outcome be the student-optimal matching for the preference profile $(P_s', P_{-s})$, for some preference $P_s'$ for $s$. Since colleges' choice functions satisfy IRC, substitutes, and the law of aggregate demand, a direct mechanism that produces the student-optimal stable outcome is strategy-proof (Aygün and Sönmez, 2013; Hatfield and Kojima, 2010). This implies that $s$ has no deviating strategy that could result in an allocation that she strictly prefers to the one she obtains by following the straightforward strategy. Therefore, by its definition, $s$ has no individual harmful manipulations. □

## References


Abdulkadiroğlu, Atila, Sönmez, Tayfun, 2003. School choice: a mechanism design approach. Am. Econ. Rev. 93 (3), 729–747.

Abdulkadiroglu, Atila, Pathak, Parag, Roth, Alvin E., Sonmez, Tayfun, 2006. Changing the Boston school choice mechanism. Technical report. National Bureau of Economic Research.

Abdulkadiroğlu, Atila, Pathak, Parag A., Roth, Alvin E., 2009. Strategy-proofness versus efficiency in matching with indifferences: redesigning the NYC high school match. Am. Econ. Rev. 99 (5), 1954–1978.

Afacan, Mustafa Oguz, Dur, Umut Mert, 2017. When preference misreporting is Harm[less]ful? J. Math. Econ. (ISSN 0304-4068) 72, 16–24.

Alcalde, José, Romero-Medina, Antonio, 2000. Simple mechanisms to implement the core of college admissions problems. Games Econ. Behav. 31 (2), 294–302.

Alcalde, José, Romero-Medina, Antonio, 2005. Sequential decisions in the college admissions problem. Econ. Lett. 86 (2), 153–158.

Ashlagi, Itai, Gonczarowski, Yannai A., 2018. Stable matching mechanisms are not obviously strategy-proof. J. Econ. Theory 177, 405–425.

Ashlagi, Itai, Braverman, Mark, Kanoria, Yash, Shi, Peng, 2020. Clearing matching markets efficiently: informative signals and match recommendations. Manag. Sci. 66 (5), 2163–2193.

Ausubel, Lawrence M., 2004. An efficient ascending-bid auction for multiple objects. Am. Econ. Rev. 94 (5), 1452–1475.

Ausubel, Lawrence M., 2006. An efficient dynamic auction for heterogeneous commodities. Am. Econ. Rev. 96 (3), 602–629.

Ausubel, Lawrence M., Baranov, Oleg V., 2014. Market design and the evolution of the combinatorial clock auction. Am. Econ. Rev. 104 (5), 446–451.

Ausubel, Lawrence M., Cramton, Peter, Milgrom, Paul, 2006. The clock-proxy auction: a practical combinatorial auction design. In: Handbook of Spectrum Auction Design, pp. 120–140.

Aygün, Orhan, Sönmez, Tayfun, 2013. Matching with contracts: Comment. Am. Econ. Rev. 103 (5), 2050–2051.

Aygün, Orhan, Bó, Inácio, 2021. College admission with multidimensional privileges: the Brazilian affirmative action case. Am. Econ. J. Microecon. 13 (3), 1–28.

Baccara, Mariagiovanna, İmrohoroğlu, Ayşe, Wilson, Alistair J., Yariv, Leeat, 2012. A field study on matching with network externalities. Am. Econ. Rev. 102 (5), 1773–1804.

Balinski, Michel, Sönmez, Tayfun, 1999. A tale of two mechanisms: student placement. J. Econ. Theory 84 (1), 73–94.

Bergemann, Dirk, Morris, Stephen, 2005. Robust mechanism design. Econometrica 73 (6), 1771–1813.

Biró, Péter, 2011. University admission practices – Hungary. MiP Country Profile 5.

Bó, Inácio, Hakimov, Rustamdjan, 2020. Iterative versus standard deferred acceptance: experimental evidence. Econ. J. 130 (626), 356–392.

Chen, Li, Pereyra, Juan Sebastián, 2016. Time-constrained school choice. ECARES Working Papers.

Chen, Li, Pereyra, Juan Sebastián, 2019. Self-selection in school choice. Games Econ. Behav. 117 (C), 59–81.

Chen, Yan, Kesten, Onur, 2017. Chinese college admissions and school choice reforms: a theoretical analysis. J. Polit. Econ. 125 (1), 99–139.

Chen, Yan, Sönmez, Tayfun, 2006. School choice: an experimental study. J. Econ. Theory 127 (1), 202–231.

Ding, Tingting, Schotter, Andrew, 2019. Learning and mechanism design: an experimental test of school matching mechanisms with intergenerational advice. Econ. J. 129 (623), 2779–2804.

Dur, Umut, Kesten, Onur, 2019. Sequential versus simultaneous assignment systems and two applications. Econ. Theory 68 (2), 251–283.

Dur, Umut, Hammond, Robert G., Morrill, Thayer, 2018. Identifying the harm of manipulable school-choice mechanisms. Am. Econ. J. Econ. Policy 10 (1), 187–213.

Echenique, Federico, Wilson, Alistair J., Yariv, Leeat, 2016. Clearinghouses for two-sided matching: an experimental study. Quant. Econ. 7 (2), 449–482.

egoeimididaskalos, 2012. Sisu 2012 como manipular [video] Youtube. https://www.youtube.com/watch?v=WTDxkW0cIGQ.

Ergin, Haluk I., 2000. Consistency in house allocation problems. J. Math. Econ. 34 (1), 77–97.

Gale, D., Shapley, L.S., 1962. College admissions and the stability of marriage. Am. Math. Mon. 69 (1), 9–15.

Gong, Binglin, Liang, Yingzhi, 2016. A dynamic college admission mechanism in Inner Mongolia: Theory and experiment. Working paper.

Grenet, Julien, He, Yinghua, Kübler, Dorothea, 2019. Decentralizing centralized matching markets: Implications from early offers in university admissions. Mimeo.

Haeringer, Guillaume, Iehlé, Vincent, 2021. Gradual college admission. J. Econ. Theory 198, 105378.
Hafalir, Isa E., Yenmez, M. Bumin, Yildirim, Muhammed A., 2013. Effective affirmative action in school choice. Theor. Econ. 8 (2), 325–363.
Hakimov, Rustamdjan, Kübler, Dorothea, 2021. Experiments on centralized school choice and college admissions: a survey. Exp. Econ. 24 (2), 434–488.
Hakimov, Rustamdjan, Raghavan, Madhav, 2020. Transparency in centralised allocation. Available at SSRN.
Hakimov, Rustamdjan, Kübler, Dorothea, Pan, Siqi, 2020. Costly information acquisition in centralized matching markets. In: Available at SSRN.
Hassidim, Avinatan, Romm, Assaf, Shorrer, Ran I., 2020. The limits of incentives in economic matching procedures. Manag. Sci.
Hatfield, John William, Kojima, Fuhito, 2008. Matching with contracts: comment. Am. Econ. Rev. 98 (3), 1189–1194.
Hatfield, John William, Kojima, Fuhito, 2010. Substitutes and stability for matching with contracts. J. Econ. Theory 145 (5), 1704–1723.
Hatfield, John William, Milgrom, Paul R., 2005. Matching with contracts. Am. Econ. Rev. 95 (4), 913–935.
Hirata, Daisuke, Kasuya, Yusuke, 2014. Cumulative offer process is order-independent. Econ. Lett. 124 (1), 37–40.
Kagel, John H., Levin, Dan, 2009. Implementing efficient multi-object auction institutions: an experimental study of the performance of boundedly rational agents. Games Econ. Behav. 66 (1), 221–237.
Kagel, John H., Harstad, Ronald M., Levin, Dan, 1987. Information impact and allocation rules in auctions with affiliated private values: a laboratory study. Econometrica 55 (6), 1275–1304.
Kesten, Onur, 2010. School choice with consent. Q. J. Econ. 125 (3), 1297–1348.
Klaus, Bettina, Klijn, Flip, 2017. Non-revelation mechanisms for many-to-many matching: equilibria versus stability. Games Econ. Behav. 104, 222–229.
Klijn, Flip, Pais, Joana, Vorsatz, Marc, 2019. Static versus dynamic deferred acceptance in school choice: theory and experiment. Games Econ. Behav. 113, 147–163.
Li, Shengwu, 2017. Obviously strategy-proof mechanisms. Am. Econ. Rev. 107 (11).
Mackenzie, Andrew, Zhou, Yu, 2020. Menu mechanisms. Working Paper.
Pais, Joana, Pintér, Ágnes, 2008. School choice and information: an experimental study on matching mechanisms. Games Econ. Behav. 64 (1), 303–328.
Pathak, Parag A., Sönmez, Tayfun, 2013. School admissions reform in Chicago and England: comparing mechanisms by their vulnerability to manipulation. Am. Econ. Rev. 103 (1), 80–106.
Rees-Jones, Alex, 2018. Suboptimal behavior in strategy-proof mechanisms: evidence from the residency match. Games Econ. Behav. 108, 317–330.
Romero-Medina, Antonio, Triossi, Matteo, 2014. Non-revelation mechanisms in many-to-one markets. Games Econ. Behav. 87, 624–630.
Roth, Alvin E., 1982. The economics of matching: stability and incentives. Math. Oper. Res. 7 (4), 617–628.
Roth, Alvin E., Sotomayor, Marilda A. Oliveira, 1992. Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis. Cambridge University Press.
Shorrer, Ran I., Sóvágó, Sándor, 2017. Obvious mistakes in a strategically simple college-admissions environment. Working paper.
Triossi, Matteo, 2009. Hiring mechanisms, application costs and stability. Games Econ. Behav. 66 (1), 566–575.
Yenmez, M. Bumin, 2018. A college admissions clearinghouse. J. Econ. Theory 176, 859–885.