2016/33

A Criterion to Compare Mechanisms when Solutions Are not Unique , with Applications to Constrained School Choice

BENOÎT DECERF AND MARTIN VAN DER LINDEN

50 YEARS OF
CORE
DISCUSSION PAPERS

CORE

Voie du Roman Pays 34, L1.03.01

Tel (32 10) 47 43 04

Fax (32 10) 47 43 01

Email: immaq-library@uclouvain.be

<http://www.uclouvain.be/en-44508.html>

A criterion to compare mechanisms when solutions are not unique, with applications to constrained school choice

Benoit Decerf

Université de Namur

benoit.decerf(at)unamur.be

Martin Van der Linden

Department of Economics, Vanderbilt University

martin.van.der.linden(at)vanderbilt.edu

October 4, 2016

Abstract

We introduce a new criterion to compare the properties of mechanisms when the solution concept used induces multiple solutions. Our criterion generalizes previous approaches in the literature. We use our criterion to compare the stability of constrained versions of the Boston (*BOS*) and deferred acceptance (*DA*) school choice mechanisms in which students can only rank a subset of the schools they could potentially access. When students play a *Nash equilibrium*, we show that there is a stability cost to increasing the number of schools students can rank in *DA*. On the other hand, when students only play *undominated strategies*, increasing the number of schools students can rank increases stability. We find similar results for *BOS*. We also compare *BOS* and *DA*. Whatever the number of schools students can rank, we find that *BOS* is more stable than *DA* in Nash equilibrium, but less stable in undominated strategies.

JEL Classification: C78, D47, D82, I 20.

Keywords: Multiple solutions, School choice, Stability, Boston mechanism, Deferred acceptance mechanism, Nash equilibrium, Undominated strategy.

1 Introduction

Mechanism design studies the institutions through which agents interact in order to choose between different outcomes. Agents hold private information – their *types* – that is relevant to the decision maker and they behave strategically when asked to report this information. Typically, an institution is modeled as a (direct) *mechanism* which describes how reported types translate into a choice of outcome. The goal of the mechanism designer is to construct mechanisms that satisfy key properties. For example, a mechanism should incentivize agents

to truthfully report their types. The chosen outcomes should also be in some sense fair and efficient.

When comparing the outcomes of two mechanisms K and L , it is standard to adopt the following three-stage procedure which we call the *standard technique*. First, choose a solution concept C that *uniquely* predicts the types agents will report given their true types in both K and L . Second, compute the *solution outcomes* in mechanisms K and L , that is the outcomes that would prevail when agents report the types predicted by solution concept C . Third, determine whether the solution outcomes satisfy some property of interest X (e.g. Pareto efficiency or stability) for *every* true type profile in a relevant domain. When the third stage holds true for L but not for K , one often says that L satisfies X , that K does not satisfy X , and that L is therefore *more* X than K .

For example, consider the school choice problem in which seats at schools have to be allocated among potential students. Let the solution concept be “dominant strategy” and X be “stability”, a fairness property guaranteeing that no student legitimately envies the seat held by another student. Two famous mechanisms for school choice problems are the *deferred acceptance* and *top trading cycle* mechanisms. The dominant strategy outcome of deferred acceptance is always stable whatever the type profile, whereas the dominant strategy outcome of top trading cycle is not necessarily stable (Abdulkadiroğlu and Sönmez, 2003). From this observation one typically concludes that deferred acceptance is stable, that top trading cycle is not, and that deferred acceptance is therefore *more* stable than top trading cycle.

As elegant as the standard technique may be, it presents two major limitations that makes it applicable to only a handful of comparisons between mechanisms. First, most mechanisms cannot be *uniquely* solved using a convincing solution concept. In particular, comparisons in dominant strategies are known to severely restrict the set of comparable mechanisms (Gibbard, 1973). We refer to this limitation as the *multiple solutions* issue.

Second, it is often the case that *both* K and L fail to satisfy property X for some (but not all) of the type profiles in the domain. When this is true, K and L cannot be compared in terms of X using the standard technique. We refer to this limitation as the *subdomain violation* issue.

Many approaches have been developed to go beyond the limitations of the standard technique. In this paper, we introduce a simple but powerful criterion that generalizes some of these approaches and enables the comparison of a wide variety of mechanisms. Informally, we will say that a mechanism L is *at least as* X *as* K under some solution concept if for every type profile in which *all* solution outcomes satisfy X in K , all solution outcomes also satisfy X in L . Section 2 presents a formal definition of our criterion and a description of its relation to other approaches in the literatures.

In the rest of the paper, we illustrate the usefulness of our criterion by using it to compare the stability of *constrained* school choice mechanisms (Haeringer and Klijn, 2009). As observed by Pathak and Sönmez (2013) and others, it is rare for a school district to allow students to report preferences on all the schools they could potentially be assigned to.¹ A school choice mechanism is

¹ For instance, at the time Haeringer and Klijn (2009) was written, the authors reported that the New York City school district allowed students to rank only 12 programs, while the district had more than 500 different programs available.

constrained if students can only report preferences over a limited number of schools.

Constrained school choice mechanisms are well-suited to showcase the use of our criterion because they combine both the multiple solutions and the subdomain violation issues. In general, constrained school choice mechanisms have no dominant-strategy and cannot be uniquely solved using reasonable solution concepts (multiple solutions issue). Also, for many combinations of constrained school choice mechanisms and solution concepts, the resulting outcomes are stable under *some* type profiles only (subdomain violation issue).

Pathak and Sönmez (2013) argue that in many constrained school choice mechanisms, increasing the number of schools that students can rank decreases the risk of students misreporting their preferences. It is then natural to ask whether this decreased risk of misreport comes at the cost of a decrease in stability. Because constrained school choice mechanisms are subject to the multiple solutions and subdomain violation issues, the standard technique cannot be used to answer this question. The criterion we propose goes beyond these two limitations and enables us to provide clear analytical comparisons of the stability of constrained school choice mechanisms.

We focus on the constrained versions of the so-called Boston (*BOS*) and deferred acceptance (*DA*) mechanisms. We apply our criterion using two different solution concepts which illustrate how the choice of a solution concept can influence the comparisons. Roughly, we show that if students play Nash equilibria, increasing the number of schools that students can rank actually *decreases* the stability of *DA*. Constrained *BOS* are also more stable than constrained *DA* in Nash equilibrium. If students play undominated strategies however, both conclusions are reversed. In undominated strategies, increasing the number of schools that students can rank *increases* the stability of *DA* and constrained *DA* mechanisms are more stable than constrained *BOS* mechanisms.

The paper is organized as follows. Section 2 presents our criterion and its relation with other comparison approaches in the literature. Section 3 introduces the two constrained *DA* and *BOS* mechanisms. The stability comparisons for *DA* and *BOS* are in Section 4 and Section 5.

2 A general criterion for comparing direct mechanisms

In this section, we give a general formulation of our criterion for direct mechanisms. We illustrate the definitions with examples from school choice. In the examples, we also introduce some terminology we use throughout the paper. Appendix A describes the school choice model in more details.

There is a finite set of players I , with typical element i , and a finite set of outcomes A . Each player is associated with a **type** $y_i \in Y_i$, where Y_i is the set of possible types for player i . A list of types for every agent $y := (y_i)_{i \in I}$ is a **type profile** (sometimes *profile* for short). The set of possible type profiles is $\mathcal{Y} = \times_{i \in I} Y_i$.

Example 1 (School choice profile).

The set of players I is composed of a set of students and a set of schools. Students' types are determined by their preferences over schools and themselves. School s is **acceptable** for student t if t prefers being assigned to school s to her outside option of being unassigned (formally, the outside option corresponds to t being assigned to herself). Schools' types are determined by a priority ordering over students and a capacity which specifies the number of students the school can accept. The sets Y_i represent possible preferences when i is a student and possible pairs of priorities and capacities when i is a school. The set of outcomes A contains all assignments of students to schools and themselves (some students may be assigned to no school under $a \in A$).

A (direct) **mechanism** is a function $M : Y \rightarrow A$ which associates every type profile in $Y \subseteq \mathcal{Y}$ with a single *feasible* outcome in A . If players report profile $y \in Y$, the outcome under mechanism M is denoted by $M(y)$.

Example 2 (School choice mechanisms).

Examples of school choice mechanisms are the aforementioned *BOS* and *DA*, which will be formally described in Section 3. In *k-constrained* school choice mechanisms, Y is the set of profiles in which students report preferences with no more than k acceptable schools. Together, a school choice profile and a school choice mechanism define a preference revelation game called a *game of school choice* (Ergin and Sönmez, 2006). In a game of school choice, the feasible outcomes are the assignments of students to schools in which no school exceeds its capacity.

For every mechanism M and every type profile $y \in \mathcal{Y}$, a **solution concept** C returns a subset $C(M, y) \subseteq Y$ of type profiles. Any profile $\hat{y} \in C(M, y)$ is interpreted as a profile that players could report when their true type profile is y and the mechanism is M . Conversely, any $\hat{y} \notin C(M, y)$ is a profile we do *not* expect players to report in the same circumstances. Given some $y \in \mathcal{Y}$, we refer to $C(M, y)$ as the set of **C-profiles**. Similarly, the set of **C-outcomes** under M is $M(C(M, y))$, i.e., the set of outcomes $M(\hat{y})$ for all possible $\hat{y} \in C(M, y)$.

Example 3 (Nash equilibrium and undominated strategies in school choice).

In school choice mechanism, it is standard to assume that the schools always report their priorities and capacities truthfully.² Therefore, Nash equilibrium *profiles* are school choice profiles in which schools report their priorities and capacities truthfully and no student can improve her assignment by reporting different preferences. Nash equilibrium *outcomes* are the assignments of students to schools that result from Nash equilibrium profiles. Profiles of undominated strategies and undominated strategies outcomes are defined similarly.

In general, we are interested in knowing whether, given some type profile y , the C -outcomes of mechanism M satisfy a property X , such as stability or efficiency. For the sake of this paper, a property is a correspondence $X : Y \rightrightarrows A$ which specifies for every type profile y the set of outcomes $X(y) \subseteq A$ that satisfy X .³

For a given $y \in Y$, if mechanism M has a *unique* C -outcome $M(C(M, y)) =$

² See Kesten (2011) for an analysis of capacity manipulation by schools.

³ In general, properties could also depend on the domain of types, or on the set of feasible outcomes. Our criterion and formalism extend naturally to these more general kinds of properties.

$\{a\}$, things are simple. Either a satisfies X in which case M satisfies X , or a does not satisfy X in which case M does not satisfy X . When confronted with the multiple solutions issue however, it is often the case that only *some* of the C -outcomes satisfy X given y , whereas others do not. In these cases, we need to choose the condition under which we consider that M satisfies X and the condition under which we think that M does not. In this paper, we take a conservative approach and say that M satisfies X in C given y if *all* C -outcomes of M satisfy X given y .

Definition 1 (M satisfies X in C given y).

Mechanism M satisfies property X in solution concept C given type profile y if

$$M(C(M, y)) \subseteq X(y).$$

In the language of implementation theory, property X is a social choice correspondence and M satisfies X in C given y if M weakly implements X in C on the domain of profiles made of y only (see, e.g., [Jackson \(2001\)](#)).

Example 4 (Constrained efficient in Nash equilibrium).

Example 2 in [Ergin and Sönmez \(2006\)](#) shows a type profile for which one Nash equilibrium outcome of BOS is constrained efficient (i.e. Pareto efficient among the stable assignments), whereas the other is not. For this profile, we say that BOS is *not* constrained efficient in Nash equilibrium. For other type profiles however, all the Nash equilibrium outcomes of BOS are constrained efficient.⁴ For these profiles, we say that BOS is constrained efficient in Nash equilibrium.

As described in the Introduction, we then have the following comparison criterion. Mechanism L is at least as X as mechanism K in solution concept C if for any type profile for which K satisfies X in C , L also satisfies X in C .

Definition 2 (At least as X as).

Mechanism L is at least as X as mechanism K in solution concept C if

$$\{y \in Y \mid K(C(K, y)) \subseteq X(y)\} \subseteq \{y \in Y \mid L(C(L, y)) \subseteq X(y)\}.$$

In the language of implementation theory, L is at least as X as K in C if the *largest* domain on which K weakly implements X in C is a subset of the *largest* domain on which L weakly implements X in C . The corresponding strict comparison follows naturally.

Definition 3 (More X than).

A mechanism L is more X than mechanism K in C , if

- (i) L is at least as X as K in C , and
- (ii) there exists a type profile for which L satisfies X in C but K does not satisfy X in C .

⁴ A trivial example is when every student has the same priority at each and every school (e.g., student 1 has the highest priority in every school, student 2 the second highest priority in every school, and so on). Then there is only one Nash equilibrium outcome which is the outcome of a serial dictatorship and is therefore (constrained) efficient.

Relation to other comparison techniques in the literature

Let us first consider the standard technique. In the standard technique,

- (i) the set of C -outcomes needs to be a singleton for every type profile and
- (ii) L is more X than K if the C -outcome of L satisfies X for *all* type profiles, whereas the C -outcome of K violates X for *some* type profiles.

Clearly our criterion encompasses the standard technique in the sense that

- a) every pair of mechanisms that can be compared using C in the standard technique can also be compared using C with our criteria and
- b) our criterion agrees with any comparison made via the standard technique.

When the set of C -outcomes is a singleton but the two mechanisms face the subdomain violation issue, a common approach consists in replacing (ii) by

- (ii)' L is more X than K if for every type profile in which the C -outcome of K satisfies X , the C -outcome of L also satisfies X (and the converse is not true).

Conditions (i) and (ii)' define the approach known as “profile by profile”. The profile by profile approach is for instance used in [Gerber and Barberà \(2016\)](#) and [Dasgupta and Maskin \(2008\)](#).⁵ Again, our criterion encompasses this approach in the sense of a) and b).

Our approach bears some similarities with the recent literature comparing the degree of manipulability of mechanisms.⁶ Our problem differs however because we focus on the properties of mechanisms' *outcomes* rather than on the extent to which mechanisms are manipulable. Comparing the properties of outcomes requires us to solve the game associated with the mechanisms of interest, whereas measures of manipulability can often be defined without explicit attention to game theoretic considerations. In particular, manipulability comparisons do not have to confront the multiple solutions issue.

Our criterion can also be used to rationalize informal arguments in the literature. For instance, [Ergin and Sönmez \(2006\)](#) prove that the set of Nash equilibrium outcomes of BOS is the set of stable assignments. As noted by [Ergin and Sönmez \(2006\)](#), this provides an argument in favor of DA because the outcome of DA is efficient among the stable assignments when students play dominant strategies (whereas the Nash equilibrium outcome of BOS can be a stable outcome which is Pareto dominated by another stable outcome). Using our criterion, this argument can be reformulated as follows : DA is more

⁵ In [Gerber and Barberà \(2016\)](#), C is “iterated elimination of weakly dominated strategy” and X is the possibility of agenda manipulation in two sequential voting rules. In [Dasgupta and Maskin \(2008\)](#), C is “truthful revelation” and X is a collection of five properties in a standard voting model.

⁶ Besides the papers already cited, see [Aleskerov and Kurbanov \(1999\)](#), [Maus et al. \(2007a\)](#), [Maus et al. \(2007b\)](#), [Andersson et al. \(2014b\)](#), [Andersson et al. \(2014a\)](#), [Arribillaga and Massó \(2015a\)](#), [Arribillaga and Massó \(2015b\)](#), and [Fujinaka and Wakayama \(2015\)](#), among others).

constrained efficient than *BOS* when students play Nash equilibria in undominated strategies. According to our criterion, when students play Nash equilibria in undominated strategies, *DA* is also at least as efficient as *BOS*.⁷

Chen and Kesten (2015) propose a different criterion for outcome comparison based on the nestedness of solutions. The criterion in Chen and Kesten (2015) and our's do not encompass each other in the sense of a) and b). One feature of the nestedness criterion in Chen and Kesten (2015) is that it does not generalize the standard technique. In general, nestedness criteria are applicable to a smaller set of problems than our criterion. For example, nestedness criteria cannot always be used to compare the stability of constrained school choice mechanisms, as we illustrate in Section 4. Nestedness criteria may also lead to counter-intuitive comparisons when all solutions are not equally likely.⁸

We now turn to an application of our criterion to stability comparisons for constrained *BOS* and *DA*, considering Nash equilibrium and undominated strategies as solution concepts.

3 Two classes of competing mechanisms

In this section we describe the two classes of school choice mechanisms that we focus on. These classes were identified by Haeringer and Klijn (2009) and correspond to constrained versions of *BOS* and *DA*. Recall that a school s is acceptable for student t if t prefers being assigned to s to being assigned to no school (i.e., being assigned to herself). We first describe the well known *unconstrained BOS*.

Input : A (reported) school choice profile (see Example 1).

Round 1: Students apply to the school they reported as their most preferred acceptable school (if any). Every school that receives more applications than its capacity starts rejecting the lowest applicants in its priority ranking, up to the point where it meets its capacity. All other applicants are *definitively accepted* at the schools they applied to and capacities are adjusted accordingly.

⋮

Round ℓ : Students who are not yet assigned apply to the school they reported as their ℓ th acceptable school (if any). Every school that receives more *new* applications in round ℓ than its *remaining* capacity starts rejecting the lowest *new* applicants in its priority ranking, up to the point where it meets its capacity. All other applicants are definitively accepted at the schools they applied to and capacities are adjusted accordingly.

The algorithm terminates when all acceptable schools of the reported profile have been considered, or when every student is assigned to a school. The

⁷ Assume all Nash equilibrium outcomes of *BOS* are efficient. Ergin and Sönmez (2006) tells us that all Nash equilibrium outcomes of *BOS* are stable. Thus one of the stable assignments is efficient. This means the optimal stable assignment is also efficient. But the optimal stable assignment is precisely the unique undominated strategies outcome of *DA*.

⁸See an earlier version of this paper in Decerf (2015) for more details.

constrained versions of BOS which we will denote by BOS^k are identical to BOS except that no student is allowed more than k acceptable schools in her reported preferences.

We now turn to DA . Again, we first describe the famous *unconstrained* DA .

Input : A (reported) school choice profile.

Round 1: Students apply to the school they reported as their most preferred acceptable school (if any). Every school that receives more applications than its capacity *definitively rejects* the lowest applicants in its priority ranking, up to the point where it meets its capacity. All other applicants are *temporarily accepted* at the schools they applied to (this means they could still be rejected in a later round).

⋮

Round ℓ : Students who were rejected in round $\ell - 1$ apply to their next acceptable school (if any). Every school considers the new applicants of round ℓ *together with* the students it temporarily accepted. If needed, each school *definitively rejects* the lowest students in its priority ranking, up to the point where it meets its capacity. All other applicants are *temporarily accepted* at the schools they applied to (this means they could still be rejected in a later round).

The algorithm terminates when all acceptable schools of the reported profile have been considered, or when every student is assigned to a school. The constrained versions of DA which we will denote by DA^k are identical to DA except that no student is allowed more than k acceptable schools in her reported preferences.

4 Comparing DA^k for different values of k

4.1 Nash equilibrium

It was recognized early on that not all Nash equilibria of the unconstrained DA mechanisms are stable (Sotomayor, 1998). This turns out to be true for every DA^k when $k \geq 2$. Example 6.2 in Haeringer and Klijn (2009) illustrates this for DA^2 . In the example, a student t_1 has a higher priority than student t_2 at a school s . Also, t_1 likes s more than her assignment. Nevertheless, student t_2 is assigned to school s and t_1 is not.

It is somewhat counter-intuitive that such an assignment emerges from a Nash equilibrium. One may wonder whether t_1 could benefit from changing her reported preferences and “claiming” a seat at school s (for instance by pretending that s is her only acceptable school). Indeed, when t_1 claims a seat at s , t_2 is rejected from s , and t_1 is *temporarily* assigned to s . But this also means that t_2 will apply to her next best school, which may induce the rejection of yet another student from that school. Overall, t_1 ’s claim of a seat at s can induce a *rejection chain* (Kesten, 2010) which results in a student t^* applying to s . If t^* has a higher priority than t_1 at s , then t_1 is rejected from s when t^* applies to s and t_1 does not benefit from her claim of a seat at s (see Example 5).

Of course, not all Nash equilibria of DA^k rely on potential rejection chains. For any profile and any stable assignment, consider for example the reported profile in which all students report as their only acceptable school the school they are assigned to under the stable assignment. This reported profile does not feature potential rejection chains and is a stable Nash equilibrium in DA^k . In fact, there exist school choice profiles for which *all* Nash equilibria of DA^k are stable (e.g., when all schools have the same priorities, see Footnote 4). In the terminology of the Introduction, stability comparisons between DA^k mechanisms in Nash equilibrium combine both the multiple solutions and the subdomain violation issues. They are therefore good candidates to showcase the usefulness of our criterion.

In the case of Nash equilibrium in DA^k , a result from [Haeringer and Klijn \(2009\)](#) is instrumental to the application of our criterion. Theorem 5.3 in [Haeringer and Klijn \(2009\)](#) shows that the Nash equilibria of DA^k are *nested*, in the sense that for any k , every Nash equilibrium in DA^k is also a Nash equilibrium in DA^{k+1} .⁹ This tells us right away that whenever all Nash equilibrium outcomes of DA^{k+1} are stable, all Nash equilibrium outcomes of DA^k are also stable. Hence, DA^k is at least as stable as DA^{k+1} .

To complete our analysis, we must determine whether the converse is true. The next example shows that DA^3 is not at least as stable as DA^2 in Nash equilibrium. In all our examples, t_i 's are students and s_i 's are schools. Also, *each school always has one seat* unless stated otherwise. The *reported* preferences of student t_i are Q_i and her *true* preferences are R_i . The priorities at school s_i are F_i . The leftmost panel represents reported preferences for DA^3 . The boxed schools correspond to the outcome of DA^3 under the reported profile, and the starred outcome is the optimal stable assignment. Empty parentheses () mean that the rest of the ranking is arbitrary. Parenthesis containing a crossed-out element () mean that the rest of the ranking cannot contain s , but is otherwise arbitrary.

Example 5 (DA^3 not at least as stable as DA^2 in Nash equilibrium).

$Q_1 :$	s_3	s_1	()	$R_1 :$	s_3	s_1^*	()	$F_1 :$	t_1	t_3	t_2	()
$Q_2 :$	s_1	s_3	s_4	$R_2 :$	s_1	s_2^*	s_3 s_4	$F_2 :$	t_2	()		
$Q_3 :$	(s_1)			$R_3 :$	s_1	()		$F_3 :$	t_4	t_2	t_1	()
$Q_4 :$	s_4	s_3	()	$R_4 :$	s_4	s_3^*		$F_4 :$	t_2	t_5	t_4	()
$Q_5 :$	(s_4)			$R_5 :$	s_4^*	()						

In the example, the outcome of the reported profile is unstable and admits two blocking pairs: (t_3, s_1) and (t_5, s_4) . Nevertheless, profile Q is a Nash equilibrium since neither t_3 nor t_5 would obtain a better assignment by ranking their blocking school. For instance, if t_3 claimed a seat at s_1 , it would trigger the following *rejection chain*:

⁹ Intuitively, it is easier to understand the contrapositive. Consider a reported profile in which no student ranks more than k schools. Suppose that this profile is *not* a Nash equilibrium in DA^{k+1} . Then we argue that the profile is not a Nash equilibrium in DA^k either. Because the profile is *not* a Nash equilibrium in DA^{k+1} , some student would benefit from reporting different preferences and obtaining s^* instead of her assignment. The key is that this student could then also be assigned to s^* by ranking s^* as her *only* acceptable school. These preferences can be reported no matter the value of k . In particular, they can be reported in DA^k . Hence, the original profile is not a Nash equilibrium of DA^k .

1. t_3 claims a seat at s_1 .
2. Because t_3 has a higher priority at s_1 than the incumbent t_2 , t_2 is rejected and applies to s_3 .
3. Because t_2 has a higher priority at s_3 than the incumbent t_1 , t_1 is rejected and applies to s_1 .
4. Because t_1 has a higher priority at s_1 than the incumbent t_3 , t_3 is rejected from s_1 .

An important feature of the above example is that t_2 is involved in the rejection chains for both (t_3, s_1) and (t_5, s_4) . What is more, it can be shown that for these two rejection chains to co-exist, t_2 *must* be able to report at least three schools (see Appendix C.1). Thus this assignment cannot be reproduced as a Nash equilibrium in DA^2 . As it turns out, DA^2 is in fact stable (again, see Appendix C.1), which yields the desired counter-example for DA^2 and DA^3 .

In Appendix C.1 we show that Example 5 can be generalized to any k . For any k , there exists a profile for which DA^{k+1} is *not* at least as stable as DA^k . We therefore have the following proposition.

Proposition 1. *For all $k < m$, DA^k is more stable than DA^{k+1} in Nash equilibrium.*

Proposition 1 suggests that when agents have sufficient information on each other's preferences and coordinate on a Nash equilibrium, there is a stability cost to increasing k . This cost in terms of stability contrasts with the decrease in manipulability identified by Pathak and Sönmez (2013).¹⁰

The stability cost to increasing k suggested by Proposition 1 obviously depend on the assumption that students play a Nash equilibrium. Because school choice mechanisms are typically one-shot mechanisms in which students have imperfect information about each other's preferences, Nash equilibrium need not always be a good approximation of the way students play.¹¹ This motivates the next subsection, in which we repeat the above analysis while only assuming that students play undominated strategies.

4.2 Undominated strategy

Undominated strategies is an interesting solution concept for school choice mechanisms. One reason is that, in some cases, the only characteristic that officials may attribute to students' strategies is that they are undominated. Suppose students seek the advice of experts on the best strategy to adopt. When little information about other students' preferences is available, experts might only be

¹⁰ There is no direct relation between the results in Pathak and Sönmez (2013) and ours, however, as our comparison criterion differs from the one in Pathak and Sönmez (2013).

¹¹ This is especially true for small k . Suppose we view Nash equilibria as the result of a process of iterated elimination of dominated strategies. When k is small, there is typically few to no dominated strategies. Therefore, even at the end of the iterated deletion process, all students will be left with many possible strategies, and the chance that they coordinate on a Nash equilibrium is small.

able to recommend that students avoid dominated strategies.¹² Thus, officials can anticipate that students will play undominated strategies but may have a hard time guessing which undominated strategies are more likely than others.

Undominated strategies is a relatively weak solution concept but in school choice mechanisms it is far from vacuous. In DA^k in particular, eliminating dominated strategies does constrain significantly the set of possible reported profiles. First, Lemma 1 in Appendix B.1 shows that, in essence, undominated strategies force students to (i) report schools in the order of their true preferences, (ii) report as many acceptable schools as they can, and (iii) not report any unacceptable schools as acceptable. Also, Proposition 2 in Decerf and Van der Linden (2016) suggests that, as k grows, more and more students will have a unique undominated strategy (equivalent to their dominant strategy in the unconstrained version of DA). In fact, the next example shows that many strategies are dominated even when k is small and (i), (ii) and (iii) are satisfied.

Example 6 (Safe set of schools).

For simplicity, the example is presented for DA^3 in an environment with 4 schools but the example can easily be extended to more schools.

$$\begin{array}{ll}
 R_1 : & () \\
 R_2 : & () \\
 R_3 : & () \\
 R_4 : & s_1 \quad s_2 \quad s_3 \quad s_4 \quad () \\
 F_1 : & t_1 \quad t_2 \quad t_4 \quad () \\
 F_2 : & t_1 \quad t_4 \quad () \\
 F_3 : & t_2 \quad t_4 \quad () \\
 F_4 : & t_4 \quad ()
 \end{array}$$

At first glance, it may seem that for t_4 , ranking $Q_4 : s_1 \ s_2 \ s_4$ is undominated. By ranking s_4 where t_4 has the highest priority, t_4 makes sure that if t_1 and t_2 obtain the unique seats at s_1 and s_2 , t_4 would not end up unassigned. But note that if t_1 and t_2 are assigned to s_1 and s_2 , t_2 cannot at the same time be assigned to s_3 . If t_2 is assigned to either s_1 or s_2 , student t_4 has the highest priority at t_3 among the remaining students. Thus, for t_4 , reporting $Q'_4 : s_1 \ s_2 \ s_3$ dominates reporting $Q_4 : s_1 \ s_2 \ s_4$.

We call $\{s_1, s_2, s_3\}$ a *safe set* (of schools) for t_4 because when ranking s_1 , s_2 , and s_3 , student t_4 is certain to be assigned to a school she reported (weakly) above the school she reports last in (s_1, s_2, s_3) whatever other students report. In general, the existence of safe sets can make a lot of strategies dominated. For example, a student who has a safe set made of her 4th and 6th favorite schools will never report as acceptable two schools ranked past her 7th favorite school in an undominated strategy of DA^k with $2 \leq k \leq 7$.

Interestingly, if students play undominated strategies, the conclusion of Proposition 1 is reversed.

Proposition 2. *For all $k < m$, DA^{k+1} is more stable than DA^k in undominated strategies.*

The next example illustrates a situation in which DA^2 is stable in undominated strategies, while DA^1 is not.

¹² Even when students have probabilistic beliefs on other students' reported preferences, experts might only be able to recommend the play of *stochastically* undominated strategies (see Roth and Rothblum (1999)). Of course, an undominated strategy can always be viewed as a stochastically undominated strategies for some particular set of probabilistic beliefs.

Example 7 (DA^1 not at least as stable as DA^2 in undominated strategies).

$Q_1 :$	$\boxed{s_2}$	$Q'_1 :$	$\boxed{s_1}$	$R_1 :$	s_2	s_1^*	$()$	$F_1 :$	t_1	t_3	t_2
$Q_2 :$	$\boxed{s_1}$	$Q'_2 :$	s_1	$R_2 :$	s_1	s_2^*	$()$	$F_2 :$	t_2	t_1	$()$
$Q_3 :$	$\boxed{s_3}$	$Q'_3 :$	$\boxed{s_3}$	$R_3 :$	s_1	s_3^*	$()$	$F_3 :$	t_3	$()$	

Profiles Q and Q' are both undominated in DA^1 and their outcomes are unstable. The blocking pairs are (t_3, s_1) for Q and (t_2, s_2) for Q' .

For each of the three students, the two preferred schools form a safe set. Hence, there exists a unique profile of undominated strategies in DA^2 (the profile in which each student report her two preferred schools as acceptable without switch) and this profile yields a stable outcome. Note that the stability of DA^2 does not come from 2 being the maximum number of acceptable schools among the students (all students could have up to three acceptable schools in the example).

The proof that DA^k is *not* at least as stable as DA^{k+1} in undominated strategies uses a generalization of Example 7 to every k (see the end of Appendix C.2).

The proof that DA^{k+1} is at least as stable as DA^k in Appendix C.2 is structured as follows. First, show that for any profile of undominated strategies in DA^{k+1} , one can truncate the profile by removing the worst school ranked as acceptable for every student who rank $k+1$ acceptable schools and obtain a profile of undominated strategies in DA^k . Then consider the extension of the truncated profile in DA^k in which the next best acceptable school (if any) is added to the reported preferences of every student. We show that, provided the truncated profile is stable in DA^k , its extension is also stable in DA^{k+1} .

Thus, if all profiles of undominated strategies are stable in DA^k , the truncated profile is stable in DA^k . Also, by construction, the extension of the truncated profile is in fact the original profile in DA^{k+1} itself. Hence, when all profiles of undominated strategies are stable in DA^k , the original profile in DA^{k+1} is also stable, which concludes the proof.

Note that, as opposed to the argument for Nash equilibrium in Proposition 1, the above argument does not rely on the nestedness of undominated strategies between DA^{k+1} and DA^k . The argument does not rely either on the nestedness of undominated strategies *outcomes*. In fact, the next example shows that undominated strategies outcomes are *not* nested in DA^k . This example illustrates one limitation of comparison based on the nestedness of outcomes and equilibria such as the one in Chen and Kesten (2015).

Example 8 (Undominated strategies outcomes not nested in DA^k).

The following example shows that an undominated strategies outcome in DA^1 is not necessarily an undominated strategies outcome in DA^2 :

$Q_1 :$	$\boxed{s_1}$	$R_1 :$	s_2^*	s_1	$F_1 :$	t_1	t_2	t_3
$Q_2 :$	$\boxed{s_2}$	$R_2 :$	s_1^*	s_2	$F_2 :$	t_2	t_1	t_3
$Q_3 :$	s_2	$R_3 :$	$()$					

The squared outcome is an undominated strategies outcome in DA^1 . Intuitively, t_1 does not apply to s_2 because t_1 is worried that t_2 will apply to s_2 (which would leave t_1 unassigned if t_1 was to apply to s_2 too). In DA^2 however, the unique outcome is the optimal stable assignment with t_1 assigned to

s_2 and t_2 assigned to s_1 . Here, uniqueness follows from both t_1 and t_2 having a dominant strategy in DA^2 .

In the next example, the unique undominated strategies outcome of DA^3 is not an undominated strategies outcome in DA^2 :

$Q_1 :$	s_2	$\boxed{s_1}$	$()$	$R_1 :$	s_2	s_1^*	$()$	$F_1 :$	t_1	t_5	t_2
$Q_2 :$	s_1	$\boxed{s_2}$	$()$	$R_2 :$	s_1	s_2^*	$()$	$F_2 :$	t_2	t_1	
$Q_3 :$	s_4	$\boxed{s_3}$	$()$	$R_3 :$	s_4	s_3^*	$()$	$F_3 :$	t_3	t_5	t_4
$Q_4 :$	s_3	$\boxed{s_4}$	$()$	$R_4 :$	s_3	s_4^*	$()$	$F_4 :$	t_4	t_3	
$Q_5 :$	s_1	s_3	$\boxed{s_5}$	$R_5 :$	s_1	s_3	s_5^*	$()$	$F_5 :$	t_5	

All five students have a dominant strategy in DA^3 displayed in the leftmost panel. The unique undominated strategy outcome of DA^3 is the optimal stable assignment. Each of the first two schools reported in Q_5 triggers a different rejection chain. The rejection chain triggered when t_5 applies to s_1 in Q_5 swaps schools s_1 and s_2 between students t_1 and t_2 . Similarly, the rejection chain triggered when t_5 applies to s_3 in Q_5 swaps schools s_3 and s_4 between students t_3 and t_4 . In DA^2 , students t_1 to t_4 have dominant strategies that are equivalent to their dominant strategies in DA^3 . Student t_5 on the other hand cannot report her dominant strategy in DA^3 anymore. In order to generate the unique undominated outcome of DA^2 (the optimal stable assignment), t_5 must report s_5 as acceptable. But then t_5 cannot report both s_1 and s_3 as acceptable in DA^2 , and hence the optimal stable assignment is not supported by an undominated profile in DA^2 .

Proposition 2 suggests that when students have little information about each other's preferences and can only resort to undominated strategies, increasing the number of acceptable schools that students can report increases stability.¹³

One reason we observe this reversal between undominated strategies and Nash equilibrium is that, given the preferences reported by the other students, Nash equilibrium forces a student to take as much advantage as she can of her priorities. On the other hand, a profile of undominated strategies can combine undominated strategies which are based on incompatible beliefs in which students don't benefit fully from their priorities.

For example, in the first part of Example 8, the profile of reported preferences $Q_1 : s_1$, $Q'_2 : s_1$, $Q_3 : s_2$ would also be a profile of undominated strategies in DA^1 . In this profile, t_1 again applies to s_1 instead of applying to s_2 because t_1 is worried that t_2 will apply to s_2 (which would leave t_1 unassigned if she was to apply to s_2 too). But differently from the profile in Example 8, t_2 is confident that t_1 will apply to s_2 and that the seat at s_1 will be available. Therefore, t_2 takes the risk of applying to s_1 instead of playing a safe strategy and ranking s_2 . This reported profile leads to t_2 being unassigned and not taking advantage of her priority at s_2 . The reported profile however relies on t_2 having incorrect beliefs about t_1 's strategy.

Such incompatible beliefs are possible under a profile of undominated strategies and lead to some students not taking full advantage of their priorities. Incompatible beliefs cannot occur in Nash equilibrium however because a Nash equilibrium supposes that students correctly anticipate each other's strategies.

¹³ In this case, the effect goes in the same direction as in Pathak and Sönmez (2013).

To understand how this impacts the stability of DA^k as k decreases, remember the reason why some Nash equilibria are unstable. For a student, claiming a seat at a school at which she has a high priority may not be worth it if this induces a rejection chain which leads to the student being eventually rejected from this school anyways. Now, the construction of these rejection chains is easier when k is large. In the limit, when $k = 1$, it is in fact impossible to construct rejection chains and the Nash equilibrium outcomes of DA^1 are always stable. Because there are less and less rejection chains as k decreases, more and more students are forced to claim those seats they have a high priority at, hence increasing the likeliness of a stable outcome.

The increased ability to take advantage of priorities when rejection chains are more rare however supposes an accurate knowledge of the way other students will play. Because undominated strategies allow students to be mistaken about each other's reported preferences, this channel for increased stability as k decreases does not play out in undominated strategies. This explains part of the difference between the comparisons in undominated strategies and in Nash equilibrium.

5 Comparing BOS^k for different values of k

5.1 Nash equilibrium

The comparison of BOS^k for different values of k follows again from previous results in the literature. Theorem 6.1 in [Haeringer and Klijn \(2009\)](#) provides a straightforward generalization to BOS^k of a result from [Ergin and Sönmez \(2006\)](#) that shows that the unrestricted BOS is stable in Nash equilibrium. In Nash equilibrium, all BOS^k are indeed strategically equivalent to one another. Whatever the value of k , if student t has a higher priority at school s than a student assigned to s , student t can report s as her first choice and be assigned to s . Hence, the reported profile cannot be a Nash equilibrium if t likes s better than her assignment, ruling out any unstable Nash equilibria.

Proposition 3. *For all $k < m$, BOS^k and BOS^{k+1} are equally stable in Nash equilibrium.*

Proposition 3 suggests that when students can coordinate on a Nash equilibrium, the number of acceptable schools students are allowed to report does not affect the stability of BOS^k .

5.2 Undominated strategy

To analyze undominated strategies in BOS^k , we assume that seats are in short-supply, that is, no set of less than k schools can offer a seat to each and every student. Unless stated otherwise, all our results for BOS^k in undominated strategies rely on this assumption. From a strategic point of view, in both BOS^k and DA^k , a student who ranks an over-supplied set of schools is certain to be assigned (in the above terminology, an over-supplied set of schools is a safe set).

Short-supply conditions are weak and common in the literature (see for example [Pathak and Sönmez \(2013\)](#)). In many cases, the condition is satisfied because no set of schools can accept all potential students. This is true in many

public school districts in the United States. It can be due to the existence of outside options (such as private schools), or to the segmentation of public high schools into different groups of schools, each with their separate assignment procedure.¹⁴

Even when there happens to be an over-supplied set of schools, this set must contain no more than k schools for our short-supply condition to be violated. This is again unlikely because – as mentioned in the Introduction – k is usually much smaller in practice than the total number of schools m . Would the condition be violated, it would most likely be in a district where k is very high. In these cases, the impact of increasing k should be rather marginal.¹⁵

The result for BOS^k in undominated strategies contrasts again with the situation in Nash equilibrium.

Proposition 4. *For all $3 \leq k < m$, BOS^{k+1} is more stable than BOS^k in undominated strategies.*

Proposition 4 parallels Proposition 2 for DA^k . When agents can only resort to undominated strategies, increasing the number of acceptable schools that students can report increases the stability of BOS^k . The reason is that under the short-supply assumption, the undominated strategies *outcomes* of BOS^{k+1} are nested into the undominated strategies *outcomes* of BOS^k (Claim 4 in Appendix C.3). That is, for every undominated strategies outcome in BOS^{k+1} , there exists a profile of undominated strategies in BOS^k with the same outcome. Therefore, whenever all profiles of undominated strategies in BOS^k yield stable assignments, all profiles of undominated strategies in BOS^{k+1} necessarily yield stable assignments too.

In the proof in the Appendix, the constructions used to link undominated strategies outcomes in BOS^k and BOS^{k+1} are more complex than the truncations used in DA^k . In particular, if we remove the worst reported schools (among the schools reported as acceptable) from a profile of undominated strategies in BOS^{k+1} , the resulting profile is *not* necessarily a profile of undominated strategies in BOS^k .

This is due to profiles of undominated strategies in BOS^{k+1} featuring non-trivial switches. Although profiles of undominated strategies in DA^k may contain switches, these switches are trivial in the sense that they never influence the outcome (Lemma 1 in Appendix B.1). This is not true in BOS^k . When profiles of undominated strategies in BOS^k feature non-trivial switches, the school a student likes best among the school she reports as acceptable can for example be reported as the worst acceptable school by the student. In such cases, there is no guarantee that the truncated strategies are still undominated strategies in BOS^k .

The next example illustrates a situation in which BOS^k is *not* at least as stable as BOS^{k+1} in undominated strategies. In most cases, Q_i is undominated in BOS^k if it contains k acceptable schools, whatever the order in which the schools are reported.¹⁶ When t_i has the highest priority at her most preferred

¹⁴ Pathak and Sönmez (2013) report for instance that in 2009, there were over 14,000 applicants in the procedure assigning seats at 9 selective public high schools in Chicago, with the 9 schools only having 3,040 seats as a whole.

¹⁵ This being said, whether our results involving BOS^k in undominated strategies hold when the short-supply condition is relaxed remains an open question.

¹⁶ See Lemma 8 in Appendix B.2 for a more detailed analysis of undominated strategies in BOS^k .

school however, ranking the preferred school first is a dominant strategy. We rely on this observation to construct the next example.

Example 9 (BOS^2 not at least as stable as BOS^3 in undominated strategies).

$Q_1 :$	$\boxed{s_1}$	$()$	$R_1 :$	s_1^*	$()$	$F_1 :$	t_1	$()$
$Q_2 :$	$\boxed{s_2}$	$()$	$R_2 :$	s_2^*	$()$	$F_2 :$	t_2	$()$
$Q_3 :$	$\boxed{s_3}$	s_1	$R_3 :$	s_1	s_4^*	s_3	$F_3 :$	$t_1 \quad t_3 \quad t_4$
$Q_4 :$	s_3	s_1	$R_4 :$	s_1	s_3^*	s_4	$F_4 :$	$t_1 \quad t_4 \quad t_3$
$Q_5 :$	$()$		$R_5 :$	$(\cancel{s_3})$	$(\cancel{s_4})$			

Profile Q is undominated in BOS^2 and its outcome admits the blocking pair (t_4, s_4) .

The example is such that all undominated strategies outcomes in BOS^3 are stable. Students t_1 and t_2 must report their preferred school first in any undominated strategy as they have the highest priority at their preferred school. Students t_3 and t_4 only have three acceptable schools and must report all three as acceptable in any undominated strategy of BOS^3 . As t_5 does not find s_3 or s_4 acceptable, these schools go to either t_3 or t_4 in any undominated strategies outcome. But any distribution of s_3 and s_4 among t_3 and t_4 leads to a stable outcome. Again the stability of BOS^3 does not come from 3 being the maximum number of acceptable schools among the students (t_1 and t_2 could have up to four acceptable schools in the example).

The proof of Proposition 4 uses a generalization of Example 9 to all k (see the end of Appendix C.3).

6 Comparing BOS^k and DA^k

6.1 Nash equilibrium

Using Theorem 6.1 in Haeringer and Klijn (2009), we directly obtain that BOS^k is at least as stable as DA^k . The converse is not true for $k \geq 2$. As noted by Haeringer and Klijn (2009, p.1930), DA^1 and BOS^1 are formally equivalent. However for $k \geq 2$, there exist unstable Nash equilibria in DA^k , as shown in Example 5 (see also Example 6.2 in Haeringer and Klijn (2009)). Therefore we have the following proposition.

Proposition 5. *For all $k \in \mathbb{N}$, BOS^k is at least as stable as DA^k in Nash equilibrium. For all $k \geq 2$, BOS^k is more stable than DA^k in Nash equilibrium.*

6.2 Undominated Strategy

Once again, Proposition 5 contrasts with its counterpart in undominated strategies.

Proposition 6. *For all $k \geq 2$, DA^k is more stable than BOS^k in undominated strategies.*

As for BOS^k and BOS^{k+1} , the reason why DA^k is at least as stable as BOS^k in undominated strategies is that the undominated strategies outcomes

of DA^k are nested in the undominated strategies outcomes of BOS^k . From any profile of undominated strategies in DA^k , one can construct a profile of undominated strategies which induces the same outcome in BOS^k as follows :

- (i) do not change the report of students who are *unassigned* in DA^k , and
- (ii) make every student who is assigned in DA^k reports the school she is assigned to in DA^k as her favorite school.

The resulting profile is a profile of undominated strategies in BOS^k and yields the same outcome as the original profile in DA^k .

The next two examples present profiles for which DA^k is stable in undominated strategies whereas BOS^k is not. The difference between BOS^k and DA^k in the examples below comes from the fact that BOS^k mostly ignores the priorities of students at schools they do not report first (Ergin and Sönmez, 2006). Assume that, in the course of BOS^k , a student applies to a school s after having been rejected from her favorite reported school. If all seats at s have been allocated in earlier rounds of BOS^k , this student will be rejected from s *even if* she has a higher priority at s than the students that have been previously assigned to s . This is a consequence of BOS^k *immediately* assigning seats in each round. On the other hand, DA^k avoids these instabilities by assigning seats temporarily in each round.

Example 10 (BOS^2 not at least as stable as DA^2 in undominated strategies because of immediate assignment).

$$\begin{array}{lll}
Q_1 : \boxed{s_1} & () & R_1 : s_1^* & () & F_1 : t_1 & () \\
Q_2 : s_1 & s_2 & R_2 : s_1 & s_2^* & () & F_2 : t_2 & () \\
Q_3 : \boxed{s_2} & s_3 & R_3 : s_2 & s_3^* & s_1 & F_3 : t_3 & ()
\end{array}$$

Profile Q – which is the truncation of students’ preferences after their second preferred school – is an undominated profile in both BOS^2 and DA^2 . Its outcome in BOS^2 (boxed) is unstable with blocking pair (t_2, s_2) . The fact that t_2 has a higher priority than t_3 at s_2 has been denied by BOS^2 because t_2 did not report s_2 as her favorite school. Profile Q is the only profile of undominated strategies in DA^2 and its outcome (stared) is the most efficient stable assignment (all students have a safe set covering their two preferred schools).

Another source of instabilities in BOS^k comes from the fact that profiles of undominated strategies in BOS^k may contain non-trivial switches. As mentioned above, this cannot happen in DA^k . As it turns out, non-trivial switches induce additional instabilities. This is illustrated in the next example, where Q_3 is the strategy containing a non-trivial switch.

Example 11 (BOS^2 not at least as stable as DA^2 in undominated strategies because of non-trivial switches).

$$\begin{array}{lll}
Q_1 : \boxed{s_1} & () & R_1 : s_1^* & () & F_1 : t_1 & () \\
Q_2 : s_1 & s_2 & R_2 : s_1 & s_2^* & F_2 : t_1 & t_2 & t_3 \\
Q_3 : \boxed{s_2} & s_1 & R_3 : s_1 & s_2 &
\end{array}$$

Profile Q is undominated in BOS^2 and leads to the boxed unstable outcome, with blocking pair (t_2, s_2) . In the example, the unique profile of undominated strategies in DA^2 is R itself, which leads to the stared stable outcome.

DA^k	$>_{NE}$ $<_{US}$	DA^{k+1}
$\wedge_{NE} \vee_{US}$		$\wedge_{NE} \vee_{US}$
BOS^k	$=_{NE}$ $<_{US}$	BOS^{k+1}

Figure 1: Summary of the results. The notation $L >_C K$ reads “ L is more stable than K when students play according to solution concept C ”, where NE stands for Nash equilibrium and US for undominated strategies. Some of the comparisons in the figure require mild restrictions on k (see the propositions in the text).

The proof of Proposition 6 presents a general example valid for all k (see the end of Appendix C.5).

7 Conclusion

We have proposed a new criterion to compare mechanisms when the solution concept induces multiple solutions. We applied our criterion to school choice mechanisms and obtained stability comparisons using Nash equilibrium and undominated strategies as solution concepts. The results are summarized in Figure 1.

Although undominated strategies may sometimes be more empirically relevant than Nash equilibrium, mechanisms are rarely studied in terms of undominated strategies. The obvious reason is that undominated strategies allows for a larger number solutions, and hence a large number of potential outcomes. Because comparing many potential outcomes can be cumbersome, Nash equilibrium is usually preferred as a solution concept as it tends to induce fewer potential outcomes. The criterion we introduce is one possible answer to the difficulty of comparing multiple potential outcomes. It should therefore facilitate comparisons based on “weaker” solution concepts such as undominated strategies.

In a sense, undominated strategies and Nash equilibrium are extreme solution concepts. As is well known, a Nash equilibrium can be viewed as an iteratively undominated profile in which players correctly anticipate each other’s actions (Mas-Colell et al., 1995, Chapter 8). In undominated strategies on the other hand, neither do players have enough information on each others’ preferences to iteratively eliminate dominated strategies, nor do they correctly anticipate each other’s action.

There is of course room for a wide variety of intermediate information and

anticipation structures. One could for instance consider a solution concept US^+ in which students know the profile of priorities and also have *some* information on each others' preferences. This may be useful in mimicking the features of some actual school choice problems.

For example, it is often common knowledge that some schools are highly demanded. To match this feature, it may be useful to assume that students know each others' x first choice(s), or know the number of students who have some school s among their x preferred school(s).¹⁷ Whether any of these alternative solution concepts would yield different stability comparison than the one we obtained in undominated strategies and Nash equilibrium is an open question.

Ideally, the solution concept to use to analyze school choice mechanism should depend on which solution concept best describes students behavior. This is an empirical question which would benefit from further investigations, for example through lab experiments.

Acknowledgments

We are grateful to Estelle Cantillon, Caterina Calsamiglia, Eun Jeong Heo, François Maniquet and John Weymark for helpful discussions and comments. We also thank Tommy Andersson, Paul Edelman, Jordi Massó, William Phan, Yves Sprumont, Myrna Wooders and participants to presentations at Université catholique de Louvain, Katholieke Universiteit Leuven, Université Saint-Louis, Vanderbilt University, the 9th Conference on Economic Design and the 13th Meeting of the Society for Social Choice and Welfare for useful questions and suggestions. Martin Van der Linden thanks the CEREC and the CORE for their hospitality during visits. Support from the ERC under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement No. 269831, the Kirk Dornbush summer research grant, the National Science Foundation grant IIS-1526860 and the Fond National de la Recherche Scientifique (Belgium, mandat d'aspirant FC 95720) is gratefully acknowledged. Benoit Decerf thanks CORE, the institution at which he worked on this paper.

Appendix

A The school choice model, terminology and notations

The model is similar to [Haeringer and Klijn \(2009\)](#).¹⁸ There is a finite set of schools $S := \{s_1, \dots, s_m\}$ and a finite set of students $T := \{t_1, \dots, t_n\}$.

¹⁷ Another interesting middle point between undominated strategies and Nash equilibrium is the case in which agents know each other's preferences and iteratively eliminate undominated strategy a certain number of times but (i) do not necessarily iteratively eliminate *ad infinitum* or (ii) do not necessarily anticipate each other's action correctly (i.e., players do not necessarily coordinate on a Nash equilibrium).

¹⁸ The model in this paper differs slightly from the model of the companion paper [Decerf and Van der Linden \(2016\)](#). Among other differences, the model here accounts for changes in priorities, quotas, and preferences (whereas these are fix throughout most of [Decerf and Van der Linden \(2016\)](#)).

A generic school is denoted s_i , or sometimes s . Every school $s_i \in S$ has a capacity q_i and a priority profile F_i . A capacity $q_i \in \mathbb{N}_+$ represents the number of seats available at school $s_i \in S$. A priority F_i is a linear ordering of the students in T .¹⁹ A **priority profile** $F = (F_1, \dots, F_m)$ is a list of priorities for every $s_i \in S$ and a **capacity profile** $q = (q_1, \dots, q_m)$ a list of capacities for every $s_i \in S$.

A set of school $S^* \subseteq S$ is **over-supplied** if there are enough seats in S^* to host all students. The short-supply assumption corresponds to profiles in which there is no over-supplied set of schools S^* with more than k schools, i.e., for all $S^* \subseteq S$ with $\#S^* \leq k$, we have $\sum_{s_i \in S^*} q_i \leq \#T$, where for any set A , $\#A$ denotes the cardinality of A .

A generic student is denoted t_i , or sometimes t . Every student t_i has preferences R_i which are a linear ordering of $S \cup \{t_i\}$. A **preference profile** $R := (R_1, \dots, R_n)$ is a list of the preferences of every $t_i \in T$. For a given preference profile R , the list containing the preferences of everyone but t_i is R_{-i} .

A strict preference of t_i for s over s' is denoted $s P_i s'$, while $s R_i s'$ denotes a weak preference, allowing for $s = s'$. A school $s \in S$ is **acceptable** for t_i if $s R_i t_i$. For simplicity, we abuse the notation and write $s \in R_i$ when s is acceptable according to R and $\#R_i$ for the number of acceptable schools in R_i . By the same token, a subset of schools $S' \subseteq S$ is acceptable for t_i if all the schools in S' are acceptable, which we denote $S' \subseteq R_i$. For all x , school $R_i(x)$ is the school ranked in x -th position in R_i .

An **assignment** is a function $\mu : T \rightarrow S \cup T$ which distributes the seats among students. In an assignment, every student is assigned to a school or to herself ($\mu(t) \in S \cup \{t\}$). An assignment is **feasible** if no school exceeds its capacity, i.e. for all $s_i \in S$, $\#\{t \in T \mid \mu(t) = s_i\} \leq q_i$.

School choice mechanisms and games of school choice

A triple (F, q, R) is a **school choice profile**. A (school choice) **mechanism** M associates every reported school choice profile in some domain with a feasible assignment μ . In a **constrained mechanism** M^k , the domain only contains school choice profiles in which students report *no* more than $k \leq m$ acceptable schools.

Because we assume that schools report F and q truthfully, we suppress the reference to reported priorities and capacities and focus on reported preference profiles $Q := (Q_1, \dots, Q_n)$. The outcome of mechanism M under reported preferences profile Q is $M(Q)$.

For any Q and any t_i , the school t_i is assigned to under the assignment $M(Q)$ is $M_i(Q)$. Student t_i is **assigned** in M given Q if $M_i(Q) \neq t_i$ and **unassigned** if $M_i(Q) = t_i$.

A quadruple (M, F, q, R) defines a normal form game known as a game of school choice (Ergin and Sönmez, 2006). Therefore, we sometimes refer to reported preferences Q as **strategy profiles**.

The notations and terminology for preferences extend to reported preferences : (i) $s Q_i s'$ means that t_i reports s *weakly* before s' in Q_i , allowing for $s = s'$, (ii) school $s \in S$ is **ranked** by t_i in Q_i if $s Q_i t_i$ (iii) $s \in Q_i$ if s is ranked in Q_i ,

¹⁹ An ordering is a complete, reflexive and transitive binary relation. A linear ordering $\succ$ is an ordering that is antisymmetric, that is, $a \succ b$ and $b \succ a$ implies $a = b$.

(iv) $\#Q_i$ is the number of ranked schools in Q_i , (v) $S \subseteq Q_i$ if all the schools in S are ranked in Q_i (vi) for all x , $Q_i(x)$ is the school ranked in x -th position in Q_i , and (vii) Q_{-i} is the list of reported preferences in Q for every student but t_i .

For reported preferences Q_i and preferences R_i , we abuse the notation and write $Q_i = R_i$ when Q_i is of the form $Q_i : R_i(1) \dots R_i(\#R_i) t_i \dots$. Similarly $Q_i \neq R_i$ means that Q_i is *not* of the form $Q_i : R_i(1) \dots R_i(\#R_i) t_i \dots$. For any two strategies Q_i and Q'_i , we also write $Q_i = Q'_i$ if both strategies share the same set of ranked schools and report those schools in the same order.

Properties of strategies and ranked (sets of) schools

Consider some quadruple (M, F, q, R) . Strategy Q_i is an **undominated strategy** if for all Q'_i ,

$$M_i(Q_i, Q_{-i}) = M_i(Q'_i, Q_{-i}), \quad \text{for all } Q_{-i}, \text{ or} \quad (1)$$

$$M_i(Q_i, Q_{-i}) P_i M_i(Q'_i, Q_{-i}), \quad \text{for some } Q_{-i}. \quad (2)$$

Strategy Q_i is **dominated** if both (1) and (2) are false (i.e. there exists a strategy Q'_i such that $M_i(Q'_i, Q_{-i}) R_i M_i(Q_i, Q_{-i})$ for all Q_{-i} and $M_i(Q'_i, Q_{-i}) P_i M_i(Q_i, Q_{-i})$ for some Q_{-i}). A profile of **undominated strategy** is any strategy profiles Q for which Q_i satisfies (1) or (2) for all $t_i \in T$.

A strategy profile Q is a **Nash equilibrium** if for all $t_i \in T$

$$M_i(Q_i, Q_{-i}) R_i M_i(Q'_i, Q_{-i}), \quad \text{for all } Q'_i. \quad (3)$$

Strategy Q_i is a **dominant strategy** if

$$M_i(Q_i, Q_{-i}) R_i M_i(Q'_i, Q_{-i}), \quad \text{for all } Q_{-i} \text{ and all } Q'_i.$$

For a given mechanism M , a set of schools $S^S \subseteq S$ is a **safe set** for t_i if for any Q_i in which S^S is ranked, t_i is at least assigned to the worst school of S^S according to Q_i . Formally, take any set of schools $S^* \subseteq S$. Let Q_i^* be any strategy that contains all the schools in S^* , and let $s^* \in S^*$ be the last reported school in Q_i^* among the schools in S^* (formally $s Q_i^* s^*$ for all $s \in S^*$). Then $S^* \subseteq S$ is a safe set for t_i in M if $M_i(Q_i^*, Q_{-i}) Q_i^* s^*$ for every Q_{-i} .

For any student t_i and any $S^* \subseteq S$, we denote by w^{S^*} the worst school in S^* according to R_i (the reference to student t_i is dropped in the notation for the sake of simplicity). Finally we say that school s_j is **safe-if-favorite** for t_i if t_i is among the q_j -students with highest priority in school s_j .²⁰

Properties of assignments

A student-school pair (t, s_j) is **blocking** in assignment μ if t prefers s_j to $\mu(t)$ and there exists t' with $\mu(t') = s_j$ and $t F_j t'$. An assignment μ is **stable** if it satisfies

(No unjustified envy) μ contains no blocking pairs,

(Individual rationality) no student t_i is assigned to an unacceptable school and

²⁰ The name “safe-if-favorite” refers to mechanism BOS^k in which, if school s is safe-if-favorite for t_i and $Q_i(1) = s$, we have $BOS_i^k(Q_i, Q_{-i}) = s$ for all Q_{-i} .

(*Non wastefulness*) no student prefers a school with an available seat to her assignment, that is there exists no $t_i \in T$ and $s_j \in S$ such that $s_j P_i \mu(t_i)$ and there are less than q_j students assigned to s_j in μ .

A feasible assignment μ is (Pareto) **efficient** if there exists no other feasible assignment μ' such that $\mu'(t_i) R_i \mu(t_i)$ for all $t_i \in T$ and $\mu'(t_j) P_j \mu(t_j)$ for some $t_j \in T$. For any school choice profile, there exists a stable assignment (Gale and Shapley, 1962). Furthermore, one of the stable assignments is deemed at least as good as any other stable assignment by all students (Gale and Shapley, 1962). This Pareto optimal assignment among the stable assignments is called the **most efficient stable assignment**.

B Preliminary results on undominated and dominant strategies in DA^k and BOS^k

This section introduces several results from Decerf and Van der Linden (2016), a companion paper which contains characterizations of undominated and dominant strategies for DA^k and BOS^k . All the proofs can be found in Decerf and Van der Linden (2016).

B.1 Undominated and dominant strategies in DA^k

Two strategies Q_i and Q'_i are **equivalent** in a mechanisms M if they always yield the same assignment for i in M , whatever the preferences reported by other students. A strategy Q_i is **clean** if it

- (i) does not rank unacceptable schools,
- (ii) features no switches, where a switch is a situation in which schools s and s' are ranked in Q_i and $s Q_i s'$ although $s' P_i s$ and
- (iii) ranks as many acceptable schools as possible, i.e., $\#Q_i = \min\{k, \#R_i\}$.

Lemma 1. *For all $k \in \mathbb{N}$ and any undominated strategy Q_i of DA^k , there exists an equivalent strategy Q'_i of DA^k that is clean.*

Lemma 2 (Characterization of dominant strategies in DA^k).

For all $k \in \mathbb{N}$, strategy Q_i is dominant in DA^k if and only if either

- (i) *all the acceptable schools in R_i are ranked without switches in Q_i and these are the only schools ranked in Q_i , i.e., $Q_i : R_i(1) \dots R_i(\#R_i) t_i$, or*
- (ii) *for some $q \leq \min\{k, \#R_i\}$, the q most preferred schools in R_i form a safe set that is ranked first in Q_i and there is no switch among those q schools in Q_i , i.e., $Q_i : R_i(1) \dots R_i(q) ()$.*

A **minimal safe set** S^{MS} is a safe set for which $S^{MS} \setminus \{w^{S^{MS}}\}$ is not a safe set. That is, removing the worst school $w^{S^{MS}}$ in S^{MS} yields a set of school that is not safe.

Lemma 3 (Characterization of undominated strategies in DA^k).

For all $k \in \mathbb{N}$, Q_i is an undominated strategy in DA^k if and only if either

- (i) Q_i is a dominant strategy in DA^k , or
- (ii) k acceptable schools are ranked without switches in Q_i and for any minimal safe set S^{MS} with $\#S^{MS} \leq k$ and $S^{MS} \not\subseteq Q_i$

$$s P_i w^{S^{MS}} \quad \text{for some } s \in Q_i \text{ with } s \notin S^{MS}.$$

B.2 Undominated and dominant strategies in BOS^k

Lemma 4 (Characterization of safe strategies in BOS^k).

For all $k \in \mathbb{N}$, a strategy Q_i is safe in BOS^k if and only if

- (i) $Q_i(1)$ is safe-if-favorite, or
- (ii) there exists an over-supplied set of schools $O \subseteq Q_i$ with $\#O \leq k$.

Lemma 5 (All assignment possible when unsafe).

Let Q_i be an unsafe strategy of BOS^k , where $k \in \mathbb{N}$. For any school $s \in Q_i$, there exists Q_{-i}^* such that

$$BOS_i^k(Q_i, Q_{-i}^*) = s.$$

Lemma 6 (Unsafe strategy not dominated by another unsafe strategy).

For all $k \in \mathbb{N}$, if Q_i' ranks $\min\{k, \#R_i\}$ schools all of which are acceptable and if Q_i dominates Q_i' , then Q_i is safe.

Lemma 7 (Characterization of dominant strategies in BOS^k).

For all $k \in \mathbb{N}$, a strategy Q_i is a dominant strategy in BOS^k if and only if $Q_i(1) = R_i(1)$ and either

- (i) $R_i(1)$ is safe-if-favorite, or
- (ii) $\#R_i = \#Q_i = 1$.

Lemma 8 (Characterization of undominated strategies in BOS^k).

For all $k \in \mathbb{N}$, Q_i is an undominated strategy in BOS^k if and only if either

- (i) $Q_i(1)$ is t_i 's preferred safe-if-favorite and acceptable school, or
- (ii) $Q_i(1)$ is not safe-if-favorite and Q_i contains $\min\{k, \#R_i\}$ acceptable schools, one of which t_i prefers to all of her safe-if-favorite and acceptable schools.

C Proofs of the propositions

C.1 Proposition 1 (DA^k more stable than DA^{k+1} in Nash equilibrium)

We prove in the text that DA^k is at least as stable as DA^{k+1} in Nash equilibrium. Below, we show that DA^{k+1} is *not* at least stable as DA^k in Nash equilibrium. The required profile for DA^3 and DA^2 is provided in Example 5. We now provide a generic example for every $3 \leq k < m - 1$, which is based on extending Example 5 recursively.

$R_1 :$	s_3	s_1^*				$F_1 :$	t_1	t_3	t_2		
$R_2 :$	s_1	s_2^*	s_3	$\dots$	s_{k+2}	$F_2 :$	t_2				
$R_3 :$	s_1					$F_3 :$	t_4	t_2	t_1		
$R_4 :$	s_4	s_3^*				$F_4 :$	t_6	t_2	t_5	t_4	
$R_5 :$	s_4					$F_5 :$	t_8	t_2	t_7	t_6	
$\vdots$						$\vdots$					
$R_{2k} :$	s_{k+2}	s_{k+1}^*				$F_{k+1} :$	t_{2k}	t_2	t_{2k-1}	t_{2k-2}	
$R_{2k+1} :$	s_{k+2}^*					$F_{k+2} :$	t_2	t_{2k+1}	t_{2k}		

In Example 5, we showed that DA^{k+1} is less stable than DA^k for $k = 2$. The example showing that DA^k is more stable than DA^{k+1} is constructed recursively. With respect to the example showing that DA^{k-1} is more stable than DA^k , we add an extra school s_{k+2} and two extra students t_{2k} and t_{2k+1} :

- School s_{k+2} is attached at the end of preferences R_2 ,
- The priority ordering F_{k+1} is modified in order to give higher priority to t_{2k} than to t_2 and
- F_{k+2} , R_{2k} and R_{2k+1} are as shown above.

The stated assignment is the most efficient stable assignment.

The ranked profile given below constitutes a Nash equilibrium in DA^{k+1} and leads to the boxed outcome. The pairs (t_3, s_1) , (t_5, s_4) , $\dots$ and (t_{2k+1}, s_{k+2}) are blocking in this assignment although the profile is a Nash equilibrium as there exists a rejection chain for each of these pairs.

$Q_1 :$	$\boxed{s_3}$	s_1			
$Q_2 :$	$\boxed{s_1}$	s_3	s_4	$\dots$	s_{k+2}
$Q_3 :$	(s_1)				
$Q_4 :$	$\boxed{s_4}$	s_3			
$Q_5 :$	(s_4)				
$Q_6 :$	$\boxed{s_5}$	s_4			
$Q_7 :$	(s_5)				
$\vdots$					
$Q_{2k} :$	$\boxed{s_{k+2}}$	s_{k+1}			
$Q_{2k+1} :$	(s_{k+2})				

We prove that there exists no Nash equilibrium of DA^q with $q \in \{1, \dots, k\}$ leading to an unstable outcome. As Nash equilibria of DA^q are nested in Nash equilibria of DA^{q+1} , we only need to prove this result for $q = k$.

The proof is based on the following claim: in all Nash equilibrium outcome, any student t_i having the highest priority in one of her acceptable school s must be assigned to a school s' with $s' R_i s$.²¹ Such agent can never end up unassigned. By construction, all students t_i with $i \in \{1, 2, 4, 6, \dots, 2k-2, 2k\}$ are in this situation. Two cases can arise for Q_2 :

²¹ The reason being that the singleton $\{s\}$ constitutes a safe set for t_i when she has the highest priority at s . Therefore, if $Q_i : s$, then we have $DA_i^k(Q_i, Q_{-i}) = s$ for all Q_{-i} . Thus t_i can always guarantee herself a school at least as good as s and there can be no Nash equilibrium in which t_i is assigned a worse school than s .

Case 1: $s_2 Q_2 s_{k+2}$ or $s_{k+2} \notin Q_2$.

This case is such that t_2 cannot be assigned to s_{k+2} . As a result, t_{2k+1} is assigned to s_{k+2} if she ranks this school as her favorite since she has the highest priority in s_{k+2} after t_2 . This shows, in this case, that student t_{2k+1} is assigned to s_{k+2} in any Nash equilibrium outcome. Since t_{2k} has the highest priority in an acceptable school, t_{2k} cannot end up unassigned in Nash equilibrium and t_{2k} is therefore assigned to s_{k+1} . As a consequence, students t_{2k-1} and t_{2k-2} are not assigned to s_{k+1} . Applying the same reasoning recursively, we have that t_{2k-2} is assigned to s_k , t_{2k-4} is assigned to s_{k-1} , $\dots$ until t_4 is assigned to s_3 and t_1 is assigned to s_1 . This shows that t_2 is assigned to s_2 and the assignment obtained is the most efficient stable assignment.

Case 2: $s_{k+2} Q_2 s_2$.

We show by contradiction that such reported preferences are never part of a Nash equilibrium in DA^k . As t_2 has the highest priority at school $R_2(2) = s_2$, student t_2 is assigned to a school she deems at least as desirable as $R_2(2)$ for any Nash equilibrium outcome. There are two subcases:

Subcase 1: Student t_2 is assigned to s_1 in the Nash equilibrium outcome.

As t_1 and t_4 must be assigned in any Nash equilibrium outcome, student t_1 is assigned to s_3 and consequently student t_4 is assigned to s_4 . This reasoning can be pursued until t_{2k-2} is assigned to s_{k+1} and t_{2k} is assigned to s_{k+2} . The pairs (t_3, s_1) , (t_5, s_4) , $\dots$ and (t_{2k+1}, s_{k+2}) are blocking in such an assignment. As we assumed that this assignment is a Nash equilibrium outcome, there must exist a rejection chain for those blocking pairs. Given the priority structure, such a rejection chain exists for (t_3, s_1) only if $Q_1 : s_3 s_1$ and $Q_2 : s_1 s_3$ (). The rejection chain exists for (t_5, s_4) only if $Q_4 : s_4 s_3$ and $Q_2 : s_1 s_3 s_4$ (). This reasoning is extended recursively until we conclude that the rejection chain for (t_{2k-1}, s_{k+1}) exists only if $Q_{2k-2} : s_{k+1} s_k$ and $Q_2 : s_1 s_3 s_4 \dots s_{k+1}$, where $\#Q_2 = k$. Therefore $s_{k+2} \notin Q_2$ because Q_2 can contain at most k schools in DA^k , which implies that there is no rejection chain for (t_{2k+1}, s_{k+2}) , contradicting the hypothesis that this assignment is a Nash equilibrium outcome.

Subcase 2: Student t_2 is assigned to a school s with $s_2 P_2 s$ in the Nash equilibrium outcome.

This cannot be a Nash equilibrium since t_2 could profitably deviate by ranking s_2 before s .

C.2 Proposition 2 (DA^{k+1} more stable than DA^k in undominated strategies)

Part 1. For all $k < m$, DA^{k+1} is at least as stable as DA^k in undominated strategies.

Let a **clean US** be an undominated strategy that is clean. A **clean US profile** is a profile of clean US strategies.

Claim 1 (Enough to prove the proposition for clean US).

For all $k < m$, DA^{k+1} is at least as stable as DA^k in undominated strategies if and only if DA^{k+1} is at least as stable as DA^k in clean US.

Proof. Remember that DA^{k+1} is at least as stable as DA^k in undominated strategies if whenever all undominated strategies outcomes of DA^k are stable, all undominated strategies outcomes of DA^{k+1} are stable too. But by Lemma 1, all undominated strategies outcomes of DA^k are stable if and only if all clean US outcomes of DA^k are stable, which proves the claim. $\blacksquare$

Claim 2 (Constructing clean US in DA^k from clean US in DA^{k+1}).

Let Q_i^{k+1} be a clean US in DA^{k+1} . If Q_i^{k+1} ranks $k+1$ schools, let Q_i^k be obtained from Q_i^{k+1} by deleting the last school of Q_i^{k+1} . If Q_i^{k+1} ranks less than k schools, let $Q_i^k := Q_i^{k+1}$. Then Q_i^k is a clean US in DA^k .

Proof. There are two cases:

Case 1: $\#R_i < k+1$.

By Lemma 2, student t_i has a dominant strategy in DA^{k+1} . The only clean US in DA^{k+1} is $Q_i^{k+1} = R_i$. Clearly, $Q_i^k = Q_i^{k+1} = R_i$ is also a dominant strategy in DA^k , and is a clean US, the desired result.

Case 2: $\#R_i \geq k+1$.

In this case, Q_i^{k+1} ranks $k+1$ acceptable schools without switches. By Lemma 3, because Q_i^k ranks k acceptable schools without switches (by construction), it is enough to prove that for all minimal safe set S^{MS^k} with $\#S^{MS^k} \leq k$ and $S^{MS^k} \not\subseteq Q_i^k$ we have

$$s^k P_i w^{S^{MS^k}} \quad \text{for some } s^k \in Q_i^k \text{ with } s^k \notin S^{MS^k}. \quad (4)$$

Take any such S^{MS^k} . Because Q_i^k ranks k schools, $S^{MS^k} \not\subseteq Q_i^k$ implies that there exist $s^k \in Q_i^k$ with $s^k \notin S^{MS^k}$.

If $S^{MS^k} \subseteq Q_i^{k+1}$, because $S^{MS^k} \not\subseteq Q_i^k$, we have $w^{Q_i^{k+1}} \in S^{MS^k}$. But in this case, because Q_i^{k+1} is without switch $s^k P_i w^{Q_i^{k+1}} \in S^{MS^k}$ and we are done.

So assume $S^{MS^k} \not\subseteq Q_i^{k+1}$. By cases (i) and (ii) in Lemma 3, for every minimal safe set $S^{MS^{k+1}}$ with $\#S^{MS^{k+1}} \leq k+1$ and $S^{MS^{k+1}} \not\subseteq Q_i^{k+1}$ we have

$$s^{k+1} P_i w^{S^{MS^{k+1}}} \quad \text{for some } s^{k+1} \in Q_i^{k+1} \text{ with } s^{k+1} \notin S^{MS^{k+1}}. \quad (5)$$

In particular,

$$s^{k+1} P_i w^{S^{MS^k}} \quad \text{for some } s^{k+1} \in Q_i^{k+1} \text{ with } s^{k+1} \notin S^{MS^k}. \quad (6)$$

If $s^{k+1} \neq w^{Q_i^{k+1}}$, then by construction of Q_i^k , $s^{k+1} \in Q_i^k$ and we are done. Thus, assume that $s^{k+1} = w^{Q_i^{k+1}}$. By assumption, $S^{MS^k} \not\subseteq Q_i^{k+1}$. But because Q_i^{k+1} ranks $k+1$ schools, this means that there exists *at least two* schools ranked in Q_i^{k+1} which are not in S^{MS^k} . In particular, there must exist $s' \notin S^{MS^k}$ with $s' \in Q_i^{k+1}$ and $s' \neq s^{k+1}$. Because Q_i^{k+1} is without switch and $s^{k+1} = w^{Q_i^{k+1}}$, we have

$$s' P_i s^{k+1} P_i w^{S^{MS^k}},$$

and by construction of Q_i^k , $s' \in Q_i^k$, the desired result.

■

With the two above claims, we are equipped to prove the “at least as stable” part of the main proposition.

By Claim 1, it is enough to prove the claim for clean US. By Claim 2, for any clean US profile Q^{k+1} in DA^{k+1} , there exists a clean US profile Q^k in DA^k constructed by removing the worst school of every agent ranking $k+1$ schools in Q^{k+1} . Let μ^k be the assignment obtained from Q^k in DA^k . We proceed by showing that if μ^k is stable, the assignment under Q^{k+1} is also stable. Then, whenever all clean US profile Q^k in DA^k lead to a stable assignment, μ^k is stable, and Q^{k+1} also yields a stable assignment, which proves that DA^{k+1} is at least as stable as DA^k .

For any student t_i , let w_i^{k+1} be the last school that t_i ranks in Q_i^{k+1} . Let T^A be the set of assigned students in μ^k . The complement set $T \setminus T^A$ contains the students who are *unassigned* in μ^k . Because μ^k is stable and w_i^{k+1} is acceptable for every $t_i \in T \setminus T^A$ (Q_i^{k+1} is a US^*), all the seats at w_i^{k+1} are assigned in μ^k . Furthermore these seats are assigned to students with a higher priority at w_i^{k+1} than t_i . Thus profile Q_*^{k+1} constructed from Q^k by adding w_i^{k+1} to the ranked profile of every $t_i \in T \setminus T^A$ yields the same assignment as Q^k (that is $DA^{k+1}(Q_*^{k+1}) = \mu^k$).

Now construct Q_{**}^{k+1} from Q_*^{k+1} by adding w_h^{k+1} to the ranked strategy of every $t_h \in T^A$ for which $w_h^{k+1} \notin Q_h^k$. By construction, w_h^{k+1} is ranked after w_h^k , the last school that t_h ranks in Q_h^k . Because $t_h \in T^A$, this means t_h is assigned to a better school than w_h^{k+1} in μ^k and never applies to w_h^{k+1} . Again, this implies that Q_{**}^{k+1} yields the same assignment μ^k as Q_*^{k+1} .

But notice that in constructing Q_{**}^{k+1} , we have added w_i^{k+1} back to the ranked profile of every agent t_i for whom Q_i^k was constructed from Q_i^{k+1} by removing w_i^{k+1} . Hence, $Q_{**}^{k+1} = Q^{k+1}$, and we have shown that $DA^{k+1}(Q^{k+1}) = \mu^k$, which proves that $DA^{k+1}(Q^{k+1})$ is stable, the desired result.

Part 2. For all $k < m$, DA^k is *not* at least as stable as DA^{k+1} in undominated strategies.

The required profile for DA^2 and DA^1 is provided in Example 7. We now provide a generic profile for every $2 \leq k < m$.

$$\begin{array}{ll}
 R_1 : & s_1^* \quad () \\
 R_2 : & s_2^* \quad () \\
 \vdots & \\
 R_{k+1} : & s_1 \quad s_2 \quad \dots \quad s_{k+1}^*
 \end{array}
 \qquad
 \begin{array}{ll}
 F_1 : & t_1 \quad () \\
 F_2 : & t_2 \quad () \\
 \vdots & \\
 F_{k+1} : & t_{k+1} \quad ()
 \end{array}$$

All the students except t_{k+1} have the highest priority at their preferred school. By Lemma 2, all students except t_{k+1} have a dominant strategy and are assigned to their favorite school in any US profile outcome. Student t_{k+1} likes k schools better than s_{k+1} , the school she is assigned to in the unique stable assignment (stared). Mechanism DA^k is unstable in undominated strategies because, by Lemma 3, it is an undominated strategy for t_{k+1} to rank her k preferred schools (and hence not rank s_{k+1}). For any profile of undominated strategies in DA^k ranked by the other students, if t_{k+1} does not rank s_{k+1} then the outcome is unstable because t_{k+1} is unassigned whereas s_{k+1} has an

available seat (in violation of non-wastefulness). On the other hand, mechanism DA^{k+1} is stable in undominated strategies since t_{k+1} has a safe set covering her $k+1$ preferred schools, which by Lemma 2 implies that her only undominated strategy is a truncation of R_i after s_{k+1} . The unique profile of undominated strategies in DA^{k+1} leads to the unique stable assignment.

C.3 Proposition 4 (BOS^{k+1} more stable than BOS^k in undominated strategies)

Part 1. For all $3 \leq k < m$, BOS^{k+1} is at least as stable as BOS^k in undominated strategies.

As explained in the text, the proof is based on showing that the undominated strategies outcomes of BOS^k are nested in the undominated strategies outcomes of BOS^{k+1} (Claim 4 below). The following definition will be useful in the proof. For any strategy Q_i , let $Q_i^{(r)}$ be the strategy obtained from Q_i by deleting the school ranked at rank r in Q_i . Formally, $Q_i^{(r)}$ is defined as follows:

- (i) $Q_i^{(r)}(q) := Q_i(q)$ for all $q \in \{1, \dots, r-1\}$ and
- (ii) $Q_i^{(r)}(q) := Q_i(q+1)$ for all $q \in \{r, \dots, \#Q_i-1\}$.

Claim 3. Let $\mu := BOS^{k+1}(Q)$ with Q a profile of undominated strategies in BOS^{k+1} and $3 \leq k < m$. For all $t_i \in T$, if $\#Q_i = k+1$, then there exists $r \in \{\#Q_i-2, \#Q_i-1, \#Q_i\}$ such that $Q_i(r) \neq \mu(t_i)$ and $Q_i^{(r)}$ is undominated in BOS^k .

Proof.

Case 1 : $Q_i(1)$ is safe-if-favorite.

By Lemma 8, as Q_i is undominated in BOS^{k+1} , $Q_i(1)$ is t_i 's favorite acceptable school among t_i 's safe-if-favorite schools. If $r = \#Q_i$, then by Lemma 8 we have that $Q_i^{(r)}$ is undominated in BOS^k as $Q_i^{(r)}(1)$ is t_i 's favorite acceptable safe-if-favorite school. Furthermore, $Q_i(r) \neq \mu(t_i)$ as $Q_i(1)$ is safe-if-favorite. Hence $Q_i(1) = \mu(t_i)$ and as we assumed $k \geq 3$, we have $2 \leq \#Q_i-2 \leq r$.

Case 2 : $Q_i(1)$ is not safe-if-favorite.

Case 2.1 : $\mu(t_i) = t_i$.

By Lemma 8, as Q_i is undominated in BOS^{k+1} , all schools ranked in Q_i are acceptable and at least one school ranked is preferred to t_i 's favorite acceptable safe-if-favorite school. Let s^* be the favorite school ranked in Q_i . Let $r := \#Q_i$ if $Q_i(\#Q_i) \neq s^*$ and $r := \#Q_i-1$ otherwise. By construction, as we assumed $\#Q_i = k+1$, strategy $Q_i^{(r)}$ contains k acceptable schools among which s^* that is preferred to the favorite acceptable safe-if-favorite. By Lemma 8, $Q_i^{(r)}$ is undominated in BOS^k .

Case 2.2 : $\mu(t_i) \neq t_i$.

The reasoning is the same as for the Case 2.1. The only difference lies in the construction of r . Rank r is any element in $\{\#Q_i-2, \#Q_i-1, \#Q_i\}$ such that $Q_i(r) \neq \mu(t_i)$ and $Q_i(r) \neq s^*$ where s^* is the favorite school ranked in Q_i .

■

Claim 4. For all $3 \leq k < m$, the US-outcomes of BOS^{k+1} are nested in the the US-outcomes of BOS^k .

Proof. Take any $3 \leq k < m$ and any profile of undominated strategies Q in BOS^{k+1} . Let $\mu := BOS^{k+1}(Q)$. We construct a profile of undominated strategies Q' in BOS^k and show that $\mu = BOS^k(Q')$.

Step 1: Construction of the profile of undominated strategies Q' .

In a nutshell, if $\mu(t_i) \in S$ and $BOS_i^{k+1}(Q) \neq Q_i(1)$, then the strategy Q'_i is constructed from Q_i by setting $Q'_i(1) := Q_i(1)$ and $Q'_i(2) := \mu(t_i)$. The detailed construction of Q'_i covering the cases $\mu(t_i) \notin S$ and $BOS_i^{k+1}(Q) = Q_i(1)$ goes as follows.

Case 1 : $Q_i(1)$ is safe-if-favorite.

Let $Q'_i : Q_i(1) t_i \dots$. By Lemma 8, as Q_i is undominated in BOS^{k+1} , $Q_i(1)$ is the favorite acceptable safe-if-favorite school. Strategy Q'_i is undominated in BOS^k by Lemma 8 as $Q'_i(1)$ is the favorite acceptable safe-if-favorite school.

Case 2 : $Q_i(1)$ is not safe-if-favorite.

For the sake of this case, we introduce the transformation Q_i^{s-2} of strategy Q_i that swaps the ranks of schools $Q_i(2)$ and s , where $s \in Q_i$. Let r^* be the rank of school s in Q_i . Formally, Q_i^{s-2} is defined from Q_i as follows:

- $Q_i^{s-2} := Q_i$ if $r^* = 1$ or $r^* = 2$,
- $Q_i^{s-2}(2) := s$, $Q_i^{s-2}(r^*) := Q_i(2)$ and for all $r \in \{1, \dots, \#Q_i\} \setminus \{2, r^*\}$ we have $Q_i^{s-2}(r) := Q_i(r)$ otherwise.

Case 2.1 : $\#Q_i < k + 1$.

By Lemma 8, if $Q_i(1)$ is not safe-if-favorite, $\#Q_i < k + 1$ and Q_i is undominated in BOS^{k+1} , then we have

- $\#Q_i = \#R_i$ and strategy Q_i contains all the acceptable schools and
- $R_i(1)$ is not safe-if-favorite.

By Lemma 8, Q_i is therefore undominated in BOS^k . Let $Q'_i := Q_i$ if $\mu(t_i) = t_i$ and $Q'_i := Q_i^{\mu(t_i)-2}$ otherwise. In the latter case, strategy Q'_i is undominated in BOS^k by Lemma 8 because $Q_i^{\mu(t_i)-2}(1)$ is not safe-if-favorite and $Q_i^{\mu(t_i)-2}$ contains all acceptable schools.

Case 2.2 : $\#Q_i = k + 1$.

By Claim 3, there exists $r^* \in \{\#Q_i - 2, \#Q_i - 1, \#Q_i\}$ such that $Q_i(r^*) \neq \mu(t_i)$ and $Q_i^{(r^*)}$ is undominated in BOS^k . Thus if $H_i := Q_i^{(r^*)}$, strategy H_i is undominated in BOS^k . Let $Q'_i := H_i$ if $\mu(t_i) = t_i$ and $Q'_i := H_i^{\mu(t_i)-2}$ otherwise. In the latter case, strategy Q'_i is undominated in BOS^k by Lemma 8 because $H_i^{\mu(t_i)-2}(1)$ is not safe-if-favorite and $H_i^{\mu(t_i)-2}$ contains k acceptable schools among which one is preferred to the favorite acceptable safe-if-favorite.²² Indeed, by construction, strategy Q'_i ranks the same schools as the undominated strategy H_i .

²² $H_i^{\mu(t_i)-2}(1)$ is not safe-if-favorite since $H_i^{\mu(t_i)-2}(1) = Q_i(1)$ by construction.

Step 2: $BOS^k(Q') = \mu$.

By the construction of Q' , we have $Q'_i(1) = Q_i(1)$ for all $t_i \in T$. As a result, the first round of BOS^k given profile Q' is the same as the first round of BOS^{k+1} given Q . Therefore, from now on, we can focus exclusively on students who are not assigned in the first round of BOS^k .

In the second round of mechanism BOS^k given profile Q' , all students for whom $\mu(t_i) \in S$ apply to $\mu(t_i)$ and all students for whom $\mu(t_i) = t_i$ apply to $Q_i(2)$.²³ We show that in round 2 of BOS^k (given profile Q') (i) all students for whom $\mu(t_i) = t_i$ are rejected and (ii) all students who rank $\mu(t_i) \in S$ are assigned to $\mu(t_i)$.

(i) $\mu(t_i) = t_i$.

In this case, t_i is rejected in the second round of BOS^{k+1} given Q . This implies that, given Q , at least $q_{Q_i(2)}$ students with higher priority than t_i at $Q_i(2)$ apply to $Q_i(2)$ in the first two rounds of BOS^{k+1} . By construction, these $q_{Q_i(2)}$ students also apply to $Q_i(2)$ in the two first rounds of BOS^k given Q' (and none of them is assigned to another school in round 1 as the first round is unchanged). As a consequence, t_i is rejected from $Q'_i(2)$ in round 2 of BOS^k .

(ii) $\mu(t_i) \in S$.

By construction we have $Q'_i(2) = \mu(t_i)$. Either there are at most $q_{\mu(t_i)}$ students applying to school $\mu(t_i)$ during the first two rounds of BOS^k given Q' (and t_i is again assigned to $\mu(t_i)$), or there are more than $q_{\mu(t_i)}$ such students. The latter case happens only if all students assigned to $\mu(t_i)$ were assigned during the first two rounds of BOS^{k+1} given Q . This implies that t_i is among the $q_{\mu(t_i)}$ students with highest priority in the subset of students who apply to $\mu(t_i)$ during the first two rounds of BOS^{k+1} given profile Q . By construction, t_i is still among the $q_{\mu(t_i)}$ students with highest priority among the subset of students who apply to $\mu(t_i)$ during the first two rounds of BOS^k given profile Q' . Therefore t_i is assigned to $\mu(t_i)$.

No assignment takes place in the later rounds of BOS^k given Q' . The students who are unassigned after round 2 are those for whom $\mu(t_i) = t_i$. In later rounds, they apply to schools from which they were rejected by BOS^{k+1} given Q . This implies that all the schools have accepted a number of students equal to their quota. After round 2 of BOS^k given Q' , these schools are also full. Therefore students t_i for whom $\mu(t_i) = t_i$ ends up unassigned. This completes the proof showing that $BOS^k(Q') = BOS^{k+1}(Q)$. ■

Claim 4 shows that for any $3 \leq k < m$, any undominated strategies outcomes in BOS^{k+1} is also an undominated strategies outcome in BOS^k . As a result, if all undominated strategies outcomes in BOS^k are stable, all undominated

²³ When $k = 3$, the construction of Q'_i described above does not guarantee that $Q'_i(2) = Q_i(2)$ for students for whom $\mu(t_i) = t_i$. Nevertheless, students for whom $\mu(t_i) = t_i$ have another undominated strategy Q_i^* such that both $Q_i^*(1) = Q_i(1)$ and $Q_i^*(2) = Q_i(2)$. Strategy Q_i^* can be constructed in this way even if $k = 3$ by taking either $Q_i^* := Q_i^{(3)}$ or $Q_i^* := Q_i^{(4)}$ depending on the rank of the preferred school ranked in Q_i (see case 2.1 in Claim 3).

strategies outcomes in BOS^{k+1} are stable as well. This implies that BOS^{k+1} is at least as stable as BOS^k .

Part 2. For all $3 \leq k < m$, BOS^k is *not* at least as stable as BOS^{k+1} in undominated strategies.

Example 9 in the text provides the required profile for the comparison of BOS^2 and BOS^3 . We now provide the generic profile for all $3 \leq k < m$.

$$\begin{array}{llllll}
R_1 : & s_1^* & () & & & F_1 : & t_1 & () \\
R_2 : & s_2^* & () & & & F_2 : & t_2 & () \\
\vdots & & & & & \vdots & & \\
R_{k-1} : & s_{k-1}^* & () & & & F_{k-1} : & t_{k-1} & () \\
R_k : & s_1 & \dots & s_{k-1} & s_{k+1}^* & s_k & & \\
R_{k+1} : & s_1 & \dots & s_{k-1} & s_k^* & s_{k+1} & & \\
R_{k+2} : & (\cancel{s_k}) & (\cancel{s_{k+1}}) & & & & &
\end{array}$$

All students except t_k, t_{k+1} and t_{k+2} have the highest priority at their preferred school. By Lemma 7, all students except t_k, t_{k+1} and t_{k+2} must rank their preferred school first in all undominated strategies. Students t_k and t_{k+1} prefer $k-1$ schools to the school they are assigned to in the most efficient stable assignment μ^e (stared). The only stable assignment that differs from μ^e is obtained from μ^e by letting t_k and t_{k+1} exchange s_k and s_{k+1} . Because there exists profiles of undominated strategies in BOS^k for which Q_k and Q_{k+1} are such that $s_k \notin Q_{k+1}$ and $s_k \notin Q_k$, BOS^k is unstable in undominated strategies (see Lemma 8). All profile of undominated strategies for which neither t_k nor t_{k+1} rank s_k lead to unstable outcomes, with one of the two students unassigned and the acceptable school s_k having an empty seat. On the other hand, BOS^{k+1} is stable since, by Lemma 8, in all the undominated strategies of BOS^{k+1} students t_k and t_{k+1} either

- rank both s_k and s_{k+1} (they have exactly $k+1$ acceptable schools), or
- play a safe strategy with $Q_k(1) = s_k$ and $Q_{k+1}(1) = s_{k+1}$.

C.4 Proposition 5 (BOS^k more stable than DA^k in Nash equilibrium)

We prove in the text that DA^k is at least as stable as BOS^k . Next, we show that DA^k is more stable than BOS^k . ? provide an example of an unstable Nash equilibrium in DA^2 . We now provide the generic profile for all $k \in \{3, \dots, m\}$.

$$\begin{array}{llllll}
R_1 : & s_2 & \boxed{s_1} & () & & F_1 : & t_1 & () \\
R_2 : & \boxed{s_3} & s_2 & () & & F_2 : & t_2 & t_1 & t_3 & () \\
R_3 : & \boxed{s_2} & s_3 & () & & F_3 : & t_3 & t_2 & () \\
\vdots & & & & & \vdots & & &
\end{array}$$

Any profile in DA^k such that

- $Q_1 : s_1$ (), $Q_2 : s_3$ s_2 (), $Q_3 : s_2$ s_3 () and
- $Q_{-1,2,3}$ is a Nash equilibrium of DA^k for the sub-profile $(F \setminus \{F_1, F_2, F_3\}, q \setminus \{q_1, q_2, q_3\}, R \setminus \{R_1, R_2, R_3\})$

is a Nash equilibrium and leads to an outcome $DA^k(Q)$ (boxed) for which the pair (t_1, s_2) is blocking. Such a Nash equilibrium in the sub-profile always exists and the profile $Q := (Q_1, Q_2, Q_3, Q_{-1,2,3})$ is a Nash equilibrium in the profile (F, q, P) . Indeed, t_2 and t_3 are assigned to their favorite school, t_1 is assigned to her second favorite school and there exists a rejection chain preventing t_1 to be assigned to s_2 if she ranked s_2 . Furthermore, no $t_i \in T \setminus \{t_1, t_2, t_3\}$ can obtain a school in $\{s_1, s_2, s_3\}$ given Q_1 , Q_2 and Q_3 and F .

C.5 Proposition 6 (DA^k more stable than BOS^k in undominated strategies)

Part 1. For all $k \in \{1, \dots, m\}$, DA^k is at least as stable as BOS^k in undominated strategies.

As explained in the text, the proof is based on showing that the undominated strategies outcomes of DA^k are nested in the undominated strategies outcomes of BOS^k (Claim 5 below).

Claim 5 (US outcomes of DA^k nested in US outcomes of BOS^k).

For all $k \in \mathbb{N}$ and all undominated strategy profile Q of DA^k , there exists an undominated strategy profile Q'' of BOS^k such that $DA^k(Q) = BOS^k(Q'')$.

Proof. First, note that by Lemma 1, it is enough to prove that $BOS^k(Q'') = DA^k(Q')$ for some clean profile Q' that is equivalent to Q (recall that a profile is clean if for all $t_i \in T$, Q'_i ranks exactly $\min\{k, \#R_i\}$ acceptable schools without switch, and no other schools – in particular no unacceptable schools – are ranked).

The proof consists in constructing a profile of undominated strategies Q'' in BOS^k such that $BOS^k(Q'') = DA^k(Q')$. The profile Q'' can be any profile satisfying the two following conditions:

- (i) for all t_i who is *unassigned* in $DA^k(Q')$, $Q''_i := Q'_i$ and
- (ii) for all t_i who is assigned in $DA^k(Q')$, $Q''_i(1) := DA^k_i(Q')$ and Q''_i ranks $\min\{k, \#R_i\}$ acceptable schools, *including all the (acceptable) schools in Q'_i .*²⁴

We show below that any strategy in profile Q'' is undominated in BOS^k . It is then easy to see that $BOS^k(Q'') = DA^k(Q')$ (the proof is omitted) which yields the desired result.

We break the rest of the proof into two parts.

(i) yields an undominated strategy for all t_i *unassigned* in $DA^k(Q')$.

We first show that Q''_i is unsafe. Consider the two cases in Lemma 4. By assumption, case (ii) is ruled out. Therefore, Q''_i is safe if and only if $Q''_i(1)$ is

²⁴ Such a strategy is feasible since $DA^k_i(Q')$ is acceptable by construction of Q'_i .

safe-if-favorite. But because $Q_i''(1) = Q_i'(1)$, if $Q_i''(1)$ was safe-if-favorite, we would have $DA_i^k(Q) = Q_i(1)$, contradicting the assumption that t_i is unassigned in $DA^k(Q)$. Therefore, Q_i'' is unsafe.

Now, in order to derive a contradiction, assume that there exists a strategy Q_i''' that dominates Q_i'' in BOS^k . By construction, $Q_i'' = Q_i'$ ranks $\min\{k, \#R_i\}$ acceptable schools. Hence, by Lemma 6, Q_i'' cannot be dominated by Q_i''' if this latter strategy is unsafe.

We obtain a contradiction by showing that Q_i''' cannot be safe either. Again by Lemma 4 and the assumption on oversupplied schools, if Q_i''' is safe, then $Q_i'''(1)$ is safe-if-favorite. If $Q_i'''(1)$ is safe-if-favorite, then we cannot have $Q_i'''(1) \in Q_i''$ (with $Q_i'' = Q_i'$ by construction). Indeed, as $Q_i'''(1)$ is safe-if-favorite, student t_i is among the students having the highest priority at $Q_i'''(1)$. If $Q_i'''(1) \in Q_i'$, then t_i would not be rejected from school $Q_i'''(1)$ in the course of DA^k given Q' . Therefore, t_i would not be unassigned in $DA^k(Q')$ as we assumed. Hence, we have $Q_i'''(1) \notin Q_i''$.

As Q_i'' is unsafe, Lemma 5 applies and for all $\ell \in \{1, \dots, \#Q_i''\}$, there exists Q_{-i}^ℓ such that

$$BOS_i^k(Q_i'', Q_{-i}^\ell) = Q_i''(\ell).$$

But because $Q_i'''(1)$ is safe-if-favorite and because Q_i''' dominates Q_i'' in BOS^k , the last displayed equality implies

$$Q_i'''(1) R_i Q_i''(\ell), \quad \text{for all } \ell \in \{1, \dots, \#Q_i''\}.$$

Finally, because $Q_i'''(1)$ is not ranked in Q_i'' , the last displayed relation implies

$$Q_i'''(1) P_i Q_i''(\ell), \quad \text{for all } \ell \in \{1, \dots, \#Q_i''\}.$$

Because $Q_i'' = Q_i'$, this means that Q_i''' would also dominate Q_i' in DA^k . Indeed, as $Q_i'''(1)$ is safe-if-favorite we have that $DA_i^k(Q_i''', Q_{-i}) = Q_i'''(1)$ for all Q_{-i} . This contradicts the fact that Q_i' is an undominated strategy of DA^k . Hence, Q_i''' cannot be safe.

As Q_i''' is neither safe nor unsafe, we have derived a contradiction.

(ii) yields an undominated strategy for all t_i assigned in $DA^k(Q')$.

Case 1 : $DA_i^k(Q')$ is safe-if-favorite.

By construction, $DA_i^k(Q')$ is acceptable for t_i . This case is such that Q_i'' is safe in BOS^k by construction. Therefore, Q_i'' can only be dominated by a safe strategy. But by Lemma 4 and the no oversupply assumption, any safe strategy Q_i''' that dominates Q_i'' must be such that $Q_i'''(1)$ is safe-if-favorite. This in turn implies that there exists a safe-if-favorite school $Q_i'''(1)$ such that

$$Q_i'''(1) P_i Q_i''(1).$$

But because $Q_i''(1) = DA_i^k(Q')$ and Q' is without switch by construction, this implies $Q_i'''(1) \notin Q_i'$ as otherwise, we would have $DA_i^k(Q') = Q_i'''(1) \neq Q_i''(1)$, a contradiction. Overall, this implies that Q_i' is dominated in DA^k by a strategy Q_i^* constructed from Q_i' by only replacing school $Q_i''(1)$ by $Q_i'''(1)$, contradicting the assumption that Q_i' is undominated in DA^k .

Case 2 : $DA_i^k(Q')$ is not safe-if-favorite.

Again, by Lemma 4 and the no oversupply assumption, Q_i'' is unsafe. But because Q_i'' ranks $\min\{k, \#R_i\}$ acceptable schools, Lemma 6 applies and any Q_i''' dominating Q_i'' must be a safe strategy. Now, by the same argument as in (i), this implies that $Q_i'''(1)$ is a safe-if-favorite school that t_i strictly prefers to all the schools in Q_i' , contradicting the assumption that Q_i' is an undominated strategy of DA^k .

■

Claim 5 shows that undominated strategies outcomes in DA^k are also undominated strategies outcomes in BOS^k . As a result, if all undominated strategies outcomes in BOS^k are stable, then all undominated strategies outcomes in DA^k are stable as well. This implies that DA^k is at least as stable as BOS^k .

Part 2. For all $k \in \{2, \dots, m\}$, BOS^k is not at least as stable as DA^k in undominated strategies.

Examples 10 and 11 in the text provide the required profiles for DA^2 and BOS^2 . We now provide the generic profile for all $k \in \{3, \dots, m\}$.

$$\begin{array}{ll}
 R_1 : & s_1^* \quad \dots \quad s_k \\
 R_2 : & s_1 \quad \dots \quad s_k \\
 \vdots & \\
 R_k : & s_1 \quad \dots \quad s_k^* \\
 R_{k+1} : & s_k
 \end{array}
 \qquad
 \begin{array}{ll}
 F_1 : & t_1 \quad \dots \quad t_k \quad t_{k+1} \\
 F_2 : & t_1 \quad \dots \quad t_k \quad t_{k+1} \\
 \vdots & \\
 F_k : & t_1 \quad \dots \quad t_k \quad t_{k+1}
 \end{array}$$

All students except t_{k+1} have the same preferences. All schools have the same priority rankings. The unique stable outcome is such that for all $i \in \{1, \dots, k\}$, student t_i is assigned to s_i and t_{k+1} is unassigned.

One reason why BOS^k is unstable in undominated strategies is that there exist profiles of undominated strategies in which Q_k is such that $Q_k(1) \neq s_k$. If this profile is ranked, because $Q_{k+1} : s_k$ is the only undominated strategy of t_{k+1} (by Lemma 8), t_k ends up unassigned although it has a higher priority at s_k than t_{k+1} . On the other hand, DA^k is stable in undominated strategies since any t_i with $i \in \{1, \dots, k\}$ must first rank her i preferred schools without switches in any undominated strategy. This follows from Lemma 2 since t_i has a safe set covering her i preferred schools.

References

- Abdulkadiroğlu, A., Sönmez, T., 2003. School choice: a mechanism design approach. *American economic review* 93, 729–747.
- Aleskerov, F., Kurbanov, E., 1999. Degree of manipulability of social choice procedures, in: Alkan, P.A., Aliprantis, P.C.D., Yannelis, P.N.C. (Eds.), *Current Trends in Economics*. Springer Berlin Heidelberg, number 8 in *Studies in Economic Theory*, pp. 13–27.

- Andersson, T., Ehlers, L., Svensson, L.G., 2014a. Budget balance, fairness, and minimal manipulability: Budget balance, fairness, and minimal manipulability. *Theoretical Economics* 9, 753–777.
- Andersson, T., Ehlers, L., Svensson, L.G., 2014b. Least manipulable envy-free rules in economies with indivisibilities. *Mathematical Social Sciences* 69, 43–49.
- Arribillaga, R.P., Massó, J., 2015a. Comparing generalized median voter schemes according to their manipulability. *Theoretical Economics* Volume 11, 547–586.
- Arribillaga, R.P., Massó, J., 2015b. Comparing voting by committees according to their manipulability. Unpublished manuscript .
- Chen, Y., Kesten, O., 2015. Chinese college admissions and school choice reforms: a theoretical analysis. *Journal of Political Economy* Forthcoming.
- Dasgupta, P., Maskin, E., 2008. On the robustness of majority rule. *Journal of the European Economic Association* 6, 949–973.
- Decerf, B., 2015. Welfare and behavior micro-analysis of economies with agents exhibiting non-classical preferences. Ph.D. thesis. Université Catholique de Louvain/Bielefeld Universitat.
- Decerf, B., Van der Linden, M., 2016. Manipulability and tie-breaking in constrained school choice. SSRN Working Paper No. 2809566.
- Ergin, H., Sönmez, T., 2006. Games of school choice under the Boston mechanism. *Journal of Public Economics* 90, 215–237.
- Fujinaka, Y., Wakayama, T., 2015. Maximal manipulation of envy-free solutions in economies with indivisible goods and money. *Journal of Economic Theory* 158, Part A, 165–185.
- Gale, D., Shapley, L.S., 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 9–15.
- Gerber, A., Barberà, S., 2016. Sequential voting and agenda manipulation. *Theoretical Economics* Forthcoming.
- Gibbard, A., 1973. Manipulation of voting schemes: a general result. *Econometrica* , 587–601.
- Haeringer, G., Klijn, F., 2009. Constrained school choice. *Journal of Economic Theory* 144, 1921–1947.
- Jackson, M.O., 2001. A crash course in implementation theory. *Social Choice and Welfare* 18, 655–708.
- Kesten, O., 2010. School choice with consent. *The Quarterly Journal of Economics* 125, 1297–1348.
- Kesten, O., 2011. On two kinds of manipulation for school choice problems. *Economic Theory* 51, 677–693.
- Mas-Colell, A., J.R., Whinston, M.D., 1995. *Microeconomic theory*. Oxford University Press.
- Maus, S., Peters, H., Storcken, T., 2007a. Anonymous voting and minimal manipulability. *Journal of Economic Theory* 135, 533–544.
- Maus, S., Peters, H., Storcken, T., 2007b. Minimal manipulability: anonymity and unanimity. *Social Choice and Welfare* 29, 247–269.
- Pathak, P.A., Sönmez, T., 2013. School admissions reform in chicao and england : comparing mechanisms by their vulnerability to manipulation. *American Economic Review* 103, 80–106.
- Roth, A.E., Rothblum, U.G., 1999. Truncation atrategies in matching markets—in search of advice for participants. *Econometrica* 67, 21–43.

Sotomayor, M.A.O., 1998. The strategy structure of the college admissions stable mechanisms. Unpublished manuscript, Universidade de São Paulo .