Controlled School Choice*

Atila Abdulkadiroğlu
Department of Economics
Duke University
Durham, NC 27708

Lars Ehlers
Department of Economics and CIREQ
Université de Montréal
Montréal, QC H3C 3J7

April 2006 (revised November 2006)

*Part of this paper was previously circulated in a working paper “Controlled Choice in Public Schools” by the first author. The first author acknowledges financial support from the Alfred P. Sloan Research Foundation and from the NSF. The second author acknowledges financial support from the SSHRC (Canada).
Abstract

Controlled choice over public schools is common practice in the United States. It attempts giving choice to parents while maintaining the racial and ethnic balance at schools. In education there is a large literature evaluating racial segregation and its consequences, and giving advice for desegregation guidelines. These guidelines may be court-ordered like in Missouri, Florida, and New York. However, none of these papers (and other papers in school choice) describe how in practice to assign students to schools while complying with these desegregation guidelines. In this paper, we provide a foundation for controlled choice. We present a model that captures the essential features of controlled school choice programs. We develop a natural notion of fairness, which is the central issue in student assignment in any school choice program and show the following results: on the positive side assignments, which are fair for same type students and constrained non-wasteful, always exist in controlled choice problems; a “controlled” version of the Gale-Shapley algorithm always finds such an assignment which is also weakly Pareto-optimal; on the negative side it is impossible to design a mechanism that is fair for same type students, constrained non-wasteful and incentive compatible. Our controlled version of the Gale-Shapley algorithm may provide a practical solution for controlled school choice.

JEL C78, D61, D78, I20.
1 Introduction

A central issue in school choice is diversity in schools. Controlled school choice in the United States attempts to give parents the opportunity to choose the school their child will attend while maintaining the racial and ethnic balance at schools. Traditionally, children were assigned public schools in their neighborhood and wealthy parents used to be able to choose a good school by moving to the neighborhood of that school whereas parents without such means did not have any choice of school, regardless of the school quality or appropriateness for the children. As a result of these concerns, public school choice programs have become increasingly popular across the United States. In 1987, Minnesota became the first state to oblige all its districts to establish a choice plan that allows parents to send their children to public schools in areas outside their resident districts. Then the students admitted at a school may not reflect the racial balance of the neighborhood in which the school is located. At the extreme only black (or only white) students may be admitted at a school located in a white (or in a black) neighborhood.

In order to avoid such segregation, some states limit school choice by court-ordered desegregation guidelines. In Missouri, St. Louis and Kansas City must observe strict racial guidelines for the placement of students in city schools. Similarly, Section 228.057 of Florida Statutes requires each school district in the state to design a choice plan that allows parents to send their children to public schools in areas outside their resident districts. It then the students admitted at a school may not reflect the racial balance of the neighborhood in which the school is located. At the extreme only black (or only white) students may be admitted at a school located in a white (or in a black) neighborhood.

In order to avoid such segregation, some states limit school choice by court-ordered desegregation guidelines. In Missouri, St. Louis and Kansas City must observe strict racial guidelines for the placement of students in city schools. Similarly, Section 228.057 of Florida Statutes requires each school district in the state to design a choice plan that allows parents to send their children to public schools in areas outside their resident districts. It then the students admitted at a school may not reflect the racial balance of the neighborhood in which the school is located. At the extreme only black (or only white) students may be admitted at a school located in a white (or in a black) neighborhood.

1 Other types of control are also present in several districts. In New York City, “Educational Option” (EdOpt) schools have to accept students across different ability range. In particular, 16 percent of students that attend an EdOpt school must score above grade level on the standardized English

\[\text{See http://wpcsd.k12.ny.us/1info/index.html}\]
Language Arts test, 68 percent must score at grade level, and the remaining 16 percent must score below grade level.²

There is a large literature in education evaluating and estimating the effects of segregation across schools on students’ achievements (Hanushek, Kain, and Rivkin (2002), Guryan (2004), Card and Rothstein (2005), and others).³ These papers give advice on how to measure segregation and how to determine optimal desegregation guidelines. School segregation can be purely racial or, as in Echenique, Fryer and Kaufman (2006), school segregation is measured according to the spectral segregation index of Echenique and Fryer (2006) which uses the intensity of social interactions among the members of a group (see also Cutler and Glaeser (1997)). Once certain desegregation guidelines are implemented, none of these papers discusses the problem of how in practice to assign students to schools and this is exactly what our paper does.

In many school districts, controlled choice constraints are implemented by imposing racial quotas at public schools. These quotas may be flexible or they may be rigid. For example, in Minneapolis, a district is allowed to go above or below the district-wide average enrollment rates by up to 15 percent points in determining the racial quotas. Cambridge applies a similar policy not only on racial diversity but on socioeconomic diversity as well. Accordingly, each school is sub-divided in two socioeconomic categories, and ±15% range is applied to each category. In White Plains, from 1964 to 1988, the rule of a differential of no more than 20 percentage points between the school with the lowest black enrollment and that with the highest guided the district’s student assignment procedures. After 1988, the Board aimed to achieve at each elementary school a mix among the black, Hispanic, and “other” students that is within ±5% points of the district average for each of these groups in each of the grade levels (Yanofsky and Laurette, 1992).

Fairness appears to be the important criterion in student assignment. Quoting from Weaver (1992), “although controlled-choice districts cannot assign all students to their first-choice schools, districts try to avoid subjective and unfair assignments by establishing clear assignment criteria. This process

²There are similar constraints in other countries as well. For example in England, City Technology Colleges are required to admit a group of students from across the ability range and their student body should be representative of the community in the catchment area (Donald Hirch, 1994, page 120).
³We will refer the interested reader to Echenique, Fryer, and Kaufman (2006) for an illuminating account of this literature.
is often as simple as prioritizing factors such as whether a family has other
cchildren in the chosen school, what a student’s racial/ethnic background is,
where a family lives, and when the application was turned in.”Indeed, a key
feature of most school choice programs (not only controlled choice programs)
is to give some students priority at certain schools. For example, some state
and local laws require that students who live in the attendance area of a
school must be given priority for that school over students who do not live in
the school’s attendance area; siblings of students already attending a school
must be given priority, and students requiring a bilingual program must be
given priority in schools that offer such programs.

Is fairness compatible with controlled choice? That is, given a controlled
school choice program, can one guarantee fair assignment of students? With-
out a positive answer to this question, it would not be possible to talk about
controlled choice and fairness. In this paper we provide a foundation for the
controlled school choice programs in the United States.

Abdulkadiroğlu and Sönmez (2003) has introduced the problem of stu-
dent assignment in school choice programs as an application of matching
theory. This theory is capable of incorporating students’ preferences as well
as schools’ preferences, which may reflect a true preference relation or a pri-
ority ordering at schools. Depending on how binding schools preferences are,
one can apply either the theory of two-sided matching, where stability arises
as a key notion, or the theory of one-sided matching, where efficiency arises
as a key notion. However, their theory does not capture one dimension of
the problem to the full extent: controlled choice.

As mentioned in Abdulkadiroğlu and Sönmez (2003), a natural point of
departure for school choice is a closely related problem, namely the college ad-
missions problem (Gale and Shapley, 1962). The college admissions problem
has been extensively studied (see Roth and Sotomayor (1990) for a survey)
and the theory built on this problem has been the basis in designing British
and American entry-level labor markets (see Roth (1984,1991,2002) and Roth
and Peranson (1999)). The central difference between college admissions and
school choice is that in college admissions, schools themselves are strategic
agents which have preferences over students, whereas in school choice, schools
are merely objects to be consumed by the students, even though schools may

4One of the key obstacles identified by the critics of school choice concerns student
selection to overdemanded schools (Hirch 1994, p. 14). Because of this reason, the design
of a student assignment mechanism remains to be an important issue in school choice
programs, whether it is controlled or not.
have priority orderings over students. Despite this important difference between the two models, school preferences and school priorities are similar mathematical objects. Consequently a potential link occurs between these two classes of problems.

In particular, since all agents in a college admissions problem are strategic, an unmatched student-school pair, say student $s$ and school $c$, may threaten stability of an assignment if student $s$ prefers school $c$ to her assignment and school $c$ prefers student $s$ to one or more of its admitted students. For later reference, we will refer such a pair as a blocking pair. An assignment is stable if no such pair exists. The notion of stability is central in the college admissions literature. Putting control aside for the time being, this mathematical property is equivalent to the following appealing property in the context of school choice, where schools do not have preferences but instead they have priorities: An assignment is fair if there is no unmatched student-school pair $(s, c)$ where student $s$ prefers school $c$ to her assignment and she has higher priority than some other student who is assigned a seat at school $c$. Therefore, a stable matching in the context of college admissions eliminates justified envy in the context of school choice.\footnote{The observation that connects fairness in a one sided matching problem to stability in a corresponding two sided matching problem has previously been made by Balinski and Sönmez (1999) in the context of Turkish college admissions, where they study a college admissions problem with responsive preferences.} In particular, the Gale-Shapley student optimal algorithm (also known as deferred acceptance algorithm with students proposing) finds a fair assignment that is preferred by every student to any other fair assignment. Moreover, revealing preferences truthfully is a dominant strategy for every student in the preference revelation game in which students submit their preferences over schools first, and then the assignment is determined via the Gale-Shapley student optimal algorithm with submitted preferences.

Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu (2005) have studied a relaxed controlled choice problem in which control is imposed on the maximum number of students from each racial/ethnic group, which we will refer as type specific quotas. Abdulkadiroğlu (2005) has observed the same link between stability and following natural fairness notion in the context of controlled choice in his set up: If there is an unmatched student-school pair $(s, c)$ where student $s$ prefers school $c$ to her assignment and she has higher priority than some other student $s'$ who is assigned a seat at school $c$, then

\begin{equation}
\end{equation}
1. students $s$ and $s'$ belong to different racial/ethnic groups, and

2. the quota for the type of student $s$ is full at school $c$.

This key observation allowed Abdulkadiroğlu (2005) to reduce his relaxed controlled choice problem to a college admissions problem and all results carry over to the relaxed problem.

Although Abdulkadiroğlu and Sönmez (2003)’s and Abdulkadiroğlu (2005)’s approaches to school choice have proved to be essential, they do not solve controlled school choice problems. To demonstrate this, consider a school that can enroll 100 students, and at most 50 of these students can be Caucasian. In this case, a student body of 50 Caucasian students would not violate the maximum quota, yet it is fully segregated. Such an assignment would not be allowed in Minneapolis, White Plains, or St. Louis.

A thorough analysis of fairness and controlled choice requires a substantial generalization of the model. To the best of our knowledge, this is the first paper that formulates the controlled school choice to the full extent and discusses fairness. However extending the model to fully capture controlled choice brings major difficulties.

The first difficulty concerns the definition of blocking pairs, hence the very definition of stability. By law, every student in the United States is entitled to get enrolled at a public school. The stability notions explained above do not incorporate this constraint. In particular, an unmatched student-school pair can block an assignment even if some other students become unassigned after the blocking pair is matched. It is this problem that proves that a controlled school choice problem is not equivalent to a college admissions problem.

Following the laws of a state, an assignment is legally feasible if both (i) every student is enrolled at a public school and (ii) at each school the court-ordered desegregation guidelines are respected. We incorporate these legal constraints in the definition of justified envy, hence in the definition of fairness. The nature of the controlled choice problem imposes that a student-school pair can cause a justified envy (or block) only if matching this pair does neither result in any unassigned student nor violate the controlled choice constraints at any school.

This raises the question of existence of fair and legally feasible assignments. As in previous papers on school choice, stability implies fairness, hence one may be tempted to find a stable assignment. Unfortunately, there
may not be any stable assignment which is legally feasible. Indeed, feasible assignments which are fair may not exist. Due to this impossibility, fairness needs to be weakened in order to respect legal constraints. A natural route is to allow envy only for students of the same type. Then only white students can justifiably envy other white students (but not any black students). It turns out that legally feasible assignments, which are fair for same types, may not exist if we require additionally non-wastefulness. This is a basic condition which says that empty seats should not be wasted if students claim them while the legal constraints can be maintained. A positive result emerges if non-wastefulness is constrained: students can claim empty seats only if the resulting assignment does not cause any envy among students of the same type. In particular, a controlled version of the Gale-Shapley algorithm finds for each controlled school choice problem a legally feasible assignment which is both fair for same types and constrained non-wasteful. The assignment found by the controlled Gale-Shapley algorithm has in addition the important welfare property of weak Pareto-optimality: it is impossible to reassign students to schools such that each student is strictly better off with the reassigned school compared to the school chosen for him by the controlled Gale-Shapley algorithm.

Students’ preferences is the only information which is private. All other components of a controlled school choice problem are commonly known. In order to elicit the true preferences from the students and to base the assignments on true information (and not lies), any school choice program would like to design a mechanism which is incentive compatible. As we show, unfortunately this is impossible while maintaining fairness and the legal constraints. More precisely, there does not exist any feasible mechanism which is incentive compatible, fair for same types and constrained non-wasteful. Furthermore, we demonstrate that giving up constrained non-wastefulness results in a possibility: we identify a feasible mechanism which is incentive compatible and fair for same types. However, that mechanism is very rigid in the sense that each school reserves for each race a fixed number of students of that race. For instance, five seats are reserved for white students and seven seats for black students. Then any controlled school choice problem is segregated according to races and for each race we run Gale-Shapley to assign the seats reserved for that race. Therefore, empty seats may be wasted and the assignment may be highly inefficient.

Depending on the preferences of a controlled school choice program we make several recommendations of how to assign students to the schools. It is
impossible to eliminate envy across different races while respecting controlled school choice constraints. Since these constraints are often legal, many school choice programs need to comply with them and fairness across types has to be abandoned. Incentive compatibility is only guaranteed if the program is ready to accept both highly inefficient assignments of students to schools and the waste of empty seats (although the assignment is based on the true preferences). In real life this cost may be too high and school choice programs may consider giving up incentive compatibility and maintaining some form of efficiency. Our controlled Gale-Shapley algorithm achieves this goal: the output assignment always respects the controlled choice constraints, it is weakly Pareto-optimal, fair for students belonging to the same race, and constrained non-wasteful.

The paper is organized as follows. Section 2 formalizes controlled school choice. Section 3 introduces our desirable criteria, namely fairness for same types and non-wastefulness. Section 4 shows that it is impossible to use results from college admissions problems for controlled school choice. Section 5 shows that there may not exist any feasible assignment which is both fair for same types and non-wasteful. Therefore, we constrain non-wastefulness and show that the controlled Gale-Shapley algorithm always finds a feasible assignment which is both fair for same types and constrained non-wasteful, and which is weakly Pareto-optimal. Section 6 focusses on incentive compatibility and shows that there may not exist any feasible mechanism which is both fair for same types and constrained wasteful. Giving up constrained non-wastefulness allows to divide the school choice problem into several problems, one for each type of students, and for each type the Gale-Shapley algorithm is applied. Section 7 summarizes our recommendations for controlled school choice programs. In the Appendix we allow justified envy across different types and show that the results for fairness across types parallel the corresponding ones for fairness for same types and (constrained) non-wastefulness. Namely, there may not exist any feasible assignment which is fair across types, and there may not exist any feasible mechanism which is both fair across types and incentive compatible.

2 Controlled School Choice

A controlled school choice problem or simply a problem consists of the following:
1. a finite set of students \( S = \{s_1, \ldots, s_n\} \);
2. a finite set of schools \( C = \{c_1, \ldots, c_m\} \);
3. a capacity vector \( q = (q_{c_1}, \ldots, q_{c_m}) \), where \( q_c \) is the capacity of school \( c \in C \);
4. a students’ preference profile \( P_S = (P_{s_1}, \ldots, P_{s_n}) \), where \( P_s \) is the strict preference relation of student \( s \in S \) over \( C \cup \{s\} \) and each school is acceptable under \( P_s \), i.e. \( cP_s s \) for all schools \( c \in C \); \( cP_sc' \) means that student \( s \) strictly prefers school \( c \) to school \( c' \);
5. a schools’ priority profile \( P_C = (P_{c_1}, \ldots, P_{c_m}) \), where \( P_c \) is the strict priority ranking of school \( c \in C \) over \( S \); \( sP_{cs} \) means that student \( s \) has higher priority than student \( s' \) to be enrolled at school \( c \);
6. a type space \( T = \{t_1, \ldots, t_k\} \);
7. a type function \( \tau: S \rightarrow T \), where \( \tau(s) \) is the type of student \( s \);
8. for each school \( c \), two vectors of type specific constraints \( q^T_c = (q^{t_1}_c, \ldots, q^{t_k}_c) \) and \( \overline{q}^T_c = (\overline{q}^{t_1}_c, \ldots, \overline{q}^{t_k}_c) \) such that \( q^t_c \leq q^t_c \leq \overline{q}^t_c \) for all \( t \in T \), and \( \sum_{t \in T} q^t_c \leq q_c \leq \sum_{t \in T} \overline{q}^t_c \).

\( q^t_c \) is the minimal number of slots that school \( c \) must by law allocate to type \( t \) students, called the floor for type \( t \) at school \( c \), whereas \( \overline{q}^t_c \) is the maximal number of slots that school \( c \) is allowed by law to allocate to type \( t \) students, called the ceiling for type \( t \) at school \( c \).

In summary, a school choice problem is given by

\[
(S, C, q, P_S, P_C, T, \tau, (q^T_c, \overline{q}^T_c)_{c \in C}).
\]

Controlled choice constraints deserve further discussion. First, these constraints are imposed by law or the state, and the school choice program has to comply with these constraints. Second, they may be more general. For example, our results would apply if these constraints were given in percentage terms (as in the Minneapolis example above). Third, the type space

\[\text{This constraint is implicitly given in school choice because students are not allowed to reject schools assigned to them.}\]
can be a very rich set. When race is controlled, $T$ is typically composed of 
\{white, black, hispanic, asian\}. The type space and type specific quotas (i.e. 
the model) can further be generalized to divide students into categories of 
several dimensions (as in the Cambridge example above). For example, con-
sider a controlled choice problem where both race and gender are controlled. 
Then $T$ can be constructed as \{white, black, hispanic, asian\}×\{female, male\} 
and $\tau(s) = (\tau^r(s), \tau^g(s)) \in$ \{white, black, hispanic, asian\}×\{female, male\} 
denotes student $s$’s race and gender. Type specific racial quotas may be in-
dependent of gender, and gender quotas may be independent of racial back-
ground. For example, when counting for black students, we do not consider 
their gender; and when counting for female students, we do not consider their 
racial background. Our results apply to this generalization as well. For ex-
positional convenience we will assume that the type space is one-dimensional 
(like race). Then the set of types induces a natural partition of the set of 
students: $(S_t)_{t \in T}$ where $S_t = \{s \in S : \tau(s) = t\}$ is the set of all students of 
type $t$.

An assignment $\mu$ is a function from the set $C \cup S$ to the set of all subsets 
of $C \cup S$ such that

i. $|\mu(s)| = 1$ for every student $s$, and\footnote{Because each student is assigned to exactly one school or no school, we will omit set brackets and write $\mu(s) = c$ instead of $\mu(s) = \{c\}$ and $\mu(s) = s$ instead of $\mu(s) = \{s\}$.} $\mu(s) = s$ if $\mu(s) \notin C$;

ii. $|\mu(c)| \leq q_c$ and $\mu(c) \subseteq S$ for every school $c$;

iii. $\mu(s) = c$ if and only if $s \in \mu(c)$.

Student $s$ is unassigned if $\mu(s) = s$; otherwise $\mu(s)$ denotes the school 
that student $s$ is assigned; $\mu(c)$ denotes the set of students that are assigned 
school $c$; and $\mu^t(c)$ denotes the students of type $t$ that are assigned to school 
c, i.e. $\mu^t(c) = \mu(c) \cap S_t$.

A set of students $S' \subseteq S$ respects (capacity and controlled choice) 
constraints at school $c$ if $|S'| \leq q_c$ and for every type $t \in T$, $\tilde{q}^t_c \leq 
|\{s \in S' : \tau(s) = t\}| \leq \tilde{q}^t_c$. An assignment $\mu$ respects constraints if for 
every school $c$, $\mu(c)$ respects constraints at $c$, i.e. for every type $t$ we have

$$
\tilde{q}^t_c \leq |\mu^t(c)| \leq \tilde{q}^t_c.
$$
Remark 1 If the controlled school choice constraints are given in percentage terms, then an assignment $\mu$ respects constraints if for every school $c$ and every type $t$ we have

$$q^t_c |\mu(c)| \leq |\mu^t(c)| \leq \overline{q}^t_c |\mu(c)|.$$  

This means, for example, that at least 33 per cent of the admitted students are white ($\overline{q}^w_c = 0.33$) and at most 66 per cent of the admitted students are white ($\overline{q}^w_c = 0.66$). Percentage terms do not cause any difficulties and all of our results apply to controlled school choice in percentage terms. In keeping our analysis simple, we assume that constraints are given in quotas. This is always true for school choice problems where the number of students is about the same as the number of empty seats available at the schools.

As outlined before the law of many states in the United States requires students to be assigned to schools such that (i) at each school the constraints are respected and (ii) each student is enrolled at a public school. An assignment $\mu$ is (legally) feasible if $\mu$ respects constraints and every student is assigned a school.\(^8\)

Obviously a controlled school choice problem does not have a feasible solution if there are not enough students of a certain type to fill the minimal number of slots required by law for that type at all schools. Therefore, we will assume that the number of students of any type is bigger than the sum of the floors for that type at all schools, i.e. for each $t \in T$, $|S_t| \geq \sum_{c \in C} q^t_c$. Similarly, in order not to leave any student unassigned we need to have enough slots for each type of students, that is $|S_t| \leq \sum_{c \in C} \overline{q}^t_c$.\(^9\)

From now on we will assume that the legal constraints at schools are such that a legally feasible assignment exists. Otherwise the laws are not compatible with each other and they need to be modified. We will not consider this issue here.

\(^8\)All our results remain unchanged if we drop requirement (ii).

\(^9\)Note that these constraints are not sufficient for the existence of a feasible assignment. For example, consider the problem consisting of three schools and three students. Each student has a different type. The capacities are all equal to 1, the floors are all equal to zero, and the ceilings are given by $\overline{q}^t_{c_1} = \overline{q}^t_{c_2} = \overline{q}^t_{c_3} = 1$, $\overline{q}^t_{c_1} = \overline{q}^t_{c_2} = 0$ and $\overline{q}^t_{c_3} = 1$, and $\overline{q}^t_{c_1} = \overline{q}^t_{c_2} = 0$ and $\overline{q}^t_{c_3} = 1$. There does not exist a feasible assignment because student $s_1$ or student $s_2$ has to be left unassigned if the constraints at school $c_1$ are respected.
3 Fairness and Non-Wastefulness

What are desirable properties of feasible assignments in controlled school choice problems? The following notions are the natural adaptations of their counterparts in standard two-sided matching (without type constraints).

The first requirement is that whenever a student prefers an empty slot to the school assigned to him, the legal constraints are violated when assigning the empty slot to this student while keeping all other assignments unchanged.\(^{10}\)

We say that student \(s\) justifiably claims an empty slot at school \(c\) under the feasible assignment \(\mu\) if

\[
\begin{align*}
(nw1) \quad & cP_s \mu(s) \text{ and } |\mu(c)| < q_c, \\
(nw2) \quad & \frac{\tau(s)}{\mu(s)} < |\mu^{\tau(s)}(\mu(s))|, \text{ and} \\
(nw3) \quad & |\mu^{\tau(s)}(c)| < \tau^{\tau(s)}.
\end{align*}
\]

Here (nw1) means student \(s\) prefers an empty slot at school \(c\) to the school assigned to him; (nw2) means that the floor of student \(s\)'s type is not binding at school \(\mu(s)\); and (nw3) means that the ceiling of student \(s\)'s type is not binding at school \(c\). Hence, under (nw1)-(nw3) student \(s\) can be assigned an empty slot at the better school \(c\) without changing the assignments of the other students and violating the constraints at any school. A feasible assignment \(\mu\) is non-wasteful if no student justifiably claims an empty slot at any school.

A well studied requirement of the literature is fairness or no-envy (Foley, 1967)\(^{11}\). In school choice student \(s\) envies student \(s'\) when \(s\) prefers the school at which \(s'\) is enrolled, say school \(c\), to her school. However, the nature of controlled school choice imposes the following (legal) constraints: Envy is justified only when

(i) student \(s\) has higher priority to be enrolled at school \(c\) than student \(s'\),

(ii) student \(s\) can be enrolled at school \(c\) without violating controlled choice constraints by removing \(s'\) from \(c\), and

\(^{10}\)This requirement is in the spirit of the property “non-wastefulness” introduced by Balinski and Sönmez (1999).

\(^{11}\)See for example Tadenuma and Thomson (1991), for an excellent survey, also see Thomson (forthcoming), Thomson (2000) and Young (1995).
(iii) student $s'$ can be enrolled at another school without violating constraints by removing $s'$ from $c$ in favor of $s$.

Throughout the main text we will require that envy is justified only if both the envying student and the envied student are of the same type. If this is the case, then (ii) and (iii) are always true since then the envying student and the envied student can simply exchange schools. We formulate our notion of fairness more precisely below.

We say that student $s$ justifiably envies student $s'$ at school $c$ under the feasible assignment $\mu$ if

\begin{align*}
(f1) \quad & \mu(s') = c, \ cP_s\mu(s) \text{ and } sP_c s', \\
(f2) \quad & \tau(s) = \tau(s').
\end{align*}

In (f1), student $s'$ is enrolled at school $c$ and both student $s$ prefers school $c$ to his assigned school $\mu(s)$ and student $s$ has higher priority to be enrolled at school $c$ than student $s'$. By (f2), student $s$ and student $s'$ are of the same type. Then we obtain a feasible assignment when students $s$ and $s'$ exchange their slots, i.e. choose $\mu'$ as follows: $\mu'(s) = \mu(s')$, $\mu'(s') = \mu(s)$, and $\mu'(\hat{s}) = \mu(\hat{s})$ for all $\hat{s} \in S \setminus \{s, s'\}$. The assignment $\mu'$ is feasible because $s$ and $s'$ are of the same type and $\mu$ was feasible.

A feasible assignment $\mu$ is fair for same types if no student justifiably envies any student who is of the same type.

We discuss also a stronger notion of fairness where envy is allowed across types, i.e. the envying student and the envied student can be of different types. Such an envy is justified only if student $s$ can be enrolled at school $c$ and student $s'$ at another school while keeping all the other assignments unchanged and satisfying the controlled school choice constraints at all schools. A feasible assignment $\mu$ is fair across types if no student justifiably envies any student. Independently of his own type, a student is allowed to envy any student. The results for this stronger condition parallel the results for fairness for same types and non-wastefulness. The detailed discussion of fairness across types can be found in the Appendix.

4 (No) Connection with College Admissions

Previous papers on school choice (or “student placement”) successfully associated any problem with a college admissions problem and applied well-known
results from this literature. In any of these papers the school choice problem can be reduced to a college admissions problem in which (i) the priority ordering of students at a school reflects that school’s preferences over individual students, (ii) a set of students that do not respect the type specific quotas of a school is not acceptable for that school, (iii) a school’s preferences over acceptable sets of students is \textit{responsive} to the priority ordering of students at that school. Then fairness in the controlled choice problem corresponds to stability in the corresponding college admissions problem. We will show that this approach is not possible here because controlled school choice imposes legal constraints which are absent in the standard two-sided matching.

Every controlled school choice problem \((S,C,q,P_S,P_C,T,\tau,(q^T_c,q^T_c)_{c\in C})\) corresponds to a college admissions problem \((S,C,q,P_S,P_C,\hat{P}_C,T,\tau,(q^T_c,q^T_c)_{c\in C})\), where \(\hat{P}_C\) is the list of colleges’ preferences over sets of students, whereas \(P_C\) is the list of colleges’ preferences over individual students. Parallel to Abdulki- 
diroğlu (2005) we will impose for each college \(c\) a “responsiveness” condition on \(\hat{P}_c\) subject to respecting constraints at \(c\).

A set of students is \textbf{acceptable for college} \(c\) if and only if it respects capacity and controlled choice constraints at college \(c\); furthermore, \(c\)’s preferences over acceptable sets of students are responsive to \(P_c\). That is, for every \(s, s' \in S\) and \(S' \subseteq S\{s, s'\}\), if \(S' \cup \{s\}\) and \(S' \cup \{s'\}\) are both acceptable for \(c\), then \(S' \cup \{s\} \hat{P}_c S' \cup \{s'\}\) if and only if \(sP_cs'\). In addition, for all sets \(S', S'' \subseteq S\), if both \(S'\) and \(S''\) are unacceptable for college \(c\), then \(\emptyset \hat{P}_c S', \emptyset \hat{P}_c S''\) and \(S'I_c S''\). In other words, any unacceptable set of students is ranked below the empty set and any two unacceptable sets are indifferent.

A matching is an assignment. We use assignment for school choice and matching for college admissions. A college-student pair \((c, s)\) \textbf{blocks a matching} \(\mu\) if \(cP_s \mu(s)\) and

\begin{enumerate}
\item either \(\mu(c) \cup \{s\}\) respects constraints at \(c\), or equivalently \(\mu(c) \cup \{s\} \hat{P}_c \mu(c)\);
\item or there exists \(s' \in \mu(c)\) such that both \(\tau(s) = \tau(s')\) and \(sP_s s'\), or equivalently both \(\tau(s) = \tau(s')\) and \((\mu(c) \{s'\}) \cup \{s\} \hat{P}_c \mu(c)\) (by responsiveness).
\end{enumerate}

A matching \(\mu\) is \textbf{stable} if (i) for every \(c\), \(\mu(c)\) is acceptable for \(c\) and (ii) it is not blocked by any college-student pair. Here (s1) corresponds to the property “non-wastefulness” introduced in Balinski and Sönmez (1999) and
(s2) corresponds to the requirement that no student envies another student who is of the same type.

Controlled school choice is fundamentally different from college admissions. In the definition of justified claim and justified envy, the initial assignment $\mu$ is assumed to be feasible. Then student $s$ can be assigned (the possibly empty) slot at his more preferred school and the (possibly) envied student $s'$ can be assigned a slot at another school while respecting controlled choice constraints at all schools. In contrast, while checking for a blocking pair in a matching, in (s1) and (s2) we only check whether the new set of students respects constraints at college $c$; we do not check whether $s'$ is matched with another college; or whether removing $s$ from $\mu(s)$ violates constraints at $\mu(s)$. Therefore, a fair and non-wasteful assignment is not necessarily stable in the corresponding college admissions problem.

Conversely, if a matching is stable in the college admissions problem and it is feasible in the corresponding school choice problem, then it will be fair for same types and non-wasteful in the corresponding school choice problem. However, the feasibility of the matching in the school choice problem is not implied by its stability in the college admissions problem, since stability requires neither that every student is matched with a college nor that the controlled choice constraints are respected at all schools. It is possible that there is no feasible assignment which is stable in the corresponding college admissions problem. The following example illustrates these points.

**Example 1.** There are three white students $w_1$, $w_2$, $w_3$; two black students $b_1$, $b_2$; and two colleges $c_1$ and $c_2$ each with capacity four. State laws require each college to admit at least one student of each type. Each college’s preference over acceptable sets of students is responsive to the following ranking of students: $w_1 P_c w_2 P_c w_3 P_c b_1 P_c b_2$. All students prefer college $c_1$ to college $c_2$. The only stable matching is the following: $\mu(c_1) = \{w_1, w_2, w_3, b_1\}$ and $\mu(c_2) = \emptyset$, i.e. $b_2$ is not matched with any college.\textsuperscript{12} Clearly, $\mu$ is not feasible in the corresponding school choice problem because (i) the minimum quotas for white and black students are not satisfied at school $c_2$ and (ii) student $b_2$ is assigned no school although the law entitles him a slot at a public school. Since $\mu$ is the only matching which is stable, any feasible assignment of the corresponding school choice problem is unstable in the college admissions problem.

\textsuperscript{12}Note that this is also the matching which Abdulkadiroğlu and Sönmez (2003)’s top trading cycles mechanism with type specific quotas finds. Therefore, Example 1 also shows that it is impossible to use this mechanism for controlled school choice.
problem. Furthermore, the unique feasible assignment which is both fair for same types and non-wasteful in the corresponding school choice problem is the following: \( \mu'(c_1) = \{w_1, w_2, b_1\} \) and \( \mu'(c_2) = \{w_3, b_2\} \).

## 5 Existence of Fair Assignments

As described before it is impossible to apply results from college admissions problems to controlled school choice. Since stable matchings always exist in college admissions problems, our first result makes this even clearer: the legal constraints, fairness and non-wastefulness may result in an incompatibility.

**Theorem 1:** The set of feasible assignments which are both fair for same types and non-wasteful may be empty in a controlled school choice problem.

**Proof:** The proof is by means of an example. Consider the following problem consisting of three schools \( \{c_1, c_2, c_3\} \) and two students \( \{s_1, s_2\} \). Each school has a capacity of two (\( q_c = 2 \) for all schools \( c \)). All students are of the same type \( t \). The ceiling of type \( t \) is equal to two at all schools (\( q^t_c = 2 \) for all schools \( c \)). School \( c_2 \) has a floor for type \( t \) of \( q^t_{c_2} = 1 \). All other floors are equal to zero. The schools’ priorities are given by \( s_2P_{c_1}s_1, s_2P_{c_2}s_1 \) and \( s_1P_{c_1}s_2 \). The students’ preferences are given by \( c_1P_{s_1}c_2P_{s_1}c_3P_{s_1}s_1 \) and \( c_3P_{s_2}c_1P_{s_2}c_2P_{s_2}s_2 \). This information is summarized in Table 1.

<table>
<thead>
<tr>
<th>( P )</th>
<th>( P_{c_1} )</th>
<th>( P_{c_2} )</th>
<th>( P_{c_3} )</th>
<th>( P_{s_1} )</th>
<th>( P_{s_2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_2 )</td>
<td>( s_2 )</td>
<td>( s_1 )</td>
<td>( c_1 )</td>
<td>( c_3 )</td>
<td></td>
</tr>
<tr>
<td>( s_1 )</td>
<td>( s_1 )</td>
<td>( s_2 )</td>
<td>( c_3 )</td>
<td>( c_1 )</td>
<td></td>
</tr>
</tbody>
</table>

Next we determine the set of assignments which are feasible for this problem. Feasibility requires that student \( s_1 \) or student \( s_2 \) is assigned school \( c_2 \) and all students are enrolled at a school. Therefore,

\[
\mu_1 = \left( \begin{array}{ccc} c_1 & c_2 & c_3 \\ s_1 & s_2 & \emptyset \end{array} \right) \quad s_2 \text{ envies } s_1 \quad \mu_2 = \left( \begin{array}{ccc} c_1 & c_2 & c_3 \\ s_2 & s_1 & \emptyset \end{array} \right),
\]

17
\[ s_1 \text{ claims } c_1 \uparrow \quad \Downarrow \quad s_2 \text{ claims } c_3 \]

\[ \mu_4 = \left( \begin{array}{ccc} c_1 & c_2 & c_3 \\ \emptyset & s_2 & s_1 \end{array} \right) \quad \leftrightarrow \quad s_1 \text{ envies } s_2 \quad \mu_3 = \left( \begin{array}{ccc} c_1 & c_2 & c_3 \\ \emptyset & s_1 & s_2 \end{array} \right) \]

are the only assignments which are feasible. Now (as indicated above)

(i) \( \mu_1 \) is not fair for same types because \( s_2 \) justifiably envies \( s_1 \) at \( c_1 \),

(ii) \( \mu_2 \) is wasteful because \( s_2 \) justifiably claims an empty slot at \( c_3 \),

(iii) \( \mu_3 \) is not fair for same types because \( s_1 \) justifiably envies \( s_2 \) at \( c_3 \), and

(iv) \( \mu_4 \) is wasteful because \( s_1 \) justifiably claims an empty slot at \( c_1 \).

Hence there is no feasible assignment which is both fair for same types and non-wasteful. \( \square \)

We discuss the robustness of Theorem 1 subject to relaxing constraints. We know from Abdulkadiroğlu (2005) that existence of feasible assignments which are fair for same types and non-wasteful is reestablished if we set all floors equal to zero. If we relax the ceilings, then we may also need to increase the number of seats which are available at a school. In the example used to prove Theorem 1 the ceilings and the capacities are such that each school’s capacity is equal to the total number of students and the ceiling for each type at each school is equal to the total number of students of that type. Hence, Theorem 1 is robust subject to relaxing ceilings and capacities.

Clearly Theorem 1 is a negative result. We will see later that the answer is affirmative to both (i) the existence of feasible assignments which are fair for same types and (ii) the existence of feasible and non-wasteful assignments. Hence, in controlled school choice problem we may retain fairness for same types or non-wastefulness while giving up the other requirement.

Giving up completely the other requirement is not satisfactory for a controlled school choice program. Since in real-life controlled school choice problems typically the total number of slots available is about the same as the number of students, potential violations of non-wastefulness arise less likely than potential envy among students who are of the same type. Therefore, our primary focus should be on fairness for same types and we propose the following: fairness for same types should be always satisfied and among those assignments we chose one which is constrained non-wasteful meaning that if a student justifiably claims an empty slot at a school, then after assigning
him this empty slot the new assignment is no longer fair for same types. Then there will be another student justifiably envying this student at the new school.

We say that a feasible assignment $\mu$ is constrained non-wasteful if:

student $s$ justifiably claims an empty slot at school $c$ under $\mu$

$\Rightarrow$ the assignment $\mu'$ (where $\mu'(s) = c$ and $\mu'(s') = \mu(s')$ for all $s' \in S \setminus \{s\}$) is not fair for same types.

If the feasible assignment $\mu$ is fair for same types and constrained non-wasteful, then the above definition is equivalent to whenever student $s$ justifiably claims an empty slot at school $c$ under $\mu$, then some other student $s'$ justifiably envies student $s$ at school $c$ under the assignment $\mu'$ (where $\mu'$ is defined as above).

The idea of feasible assignments which are both fair for same types and constrained non-wasteful is similar to the one of “bargaining sets”: if student $s$ has an objection to $\mu$ because $s$ claims an empty slot at $c$, then there will be counterobjection once $s$ is assigned to $c$ since some other student will then justifiably envy $s$ at $c$. Roughly speaking, an outcome belongs to the “bargaining set” if and only if for any objection to the outcome there exists a counterobjection.

**THEOREM 2:** The set of feasible assignments which are both fair for same types and constrained non-wasteful is non-empty in a controlled school choice problem.

In showing Theorem 2 we propose a controlled version of the Gale-Shapley algorithm. Recall that in the classical algorithm of Gale and Shapley (1962) students are put tentatively on waiting lists and at any step the students, who do not belong to any waiting list, simultaneously propose to schools to which they did not propose yet. Each school updates its waiting list by accepting the most preferred students from the new proposals and the students who were previously on its waiting list. The other students are rejected. If each student either belongs to a waiting list or has proposed to all schools, then the algorithm ends and the schools are assigned according to the waiting lists.

Our controlled version will have two important differences. First, proposals cannot be simultaneous. When several students propose simultaneously, it may be infeasible to put them on the waiting lists. In Example 1 all white
students propose to school $c_1$ and admitting all of them at school $c_1$ makes it impossible to assign at least one white student to school $c_2$.\(^\text{13}\) In our controlled Gale-Shapley algorithm proposals are sequential (say according to when the applications were received): similar to McVitie and Wilson (1970) at each step one student, who does not belong to any waiting list, proposes to the most preferred school to which he did not propose yet.

Second, when putting a student on a waiting list we need to be sure that all tentative assignments are feasible. In other words, we check whether there is some feasible assignment such that all students are assigned the school to which’s waiting list they belong. In the standard Gale-Shapley algorithm we check only whether the set of most preferred students from the new proposals and the students on the waiting list respects constraints at that school.

**The Controlled Gale-Shapley Algorithm:**

**Start:** Fix an order of the students, in which they are allowed to make proposals to schools, say $s_1 - s_2 - \cdots - s_n$. We will always define a tentative assignment $\nu$ recording the current waiting lists at all schools. The tentative assignment is such that it is possible to allocate the unassigned students to schools such that the resulting assignment is feasible. Let $\mathcal{F}$ denote the set of all feasible assignments and $\nu_0$ be the empty assignment, i.e. $\nu_0(s) = s$ for all $s \in S$. Let $P_S$ be a controlled school choice problem.

1. Let student $s_1$ apply to the school which is ranked first under $P_{s_1}$, say $c_1$. If there is some $\mu \in \mathcal{F}$ such that $\mu(s_1) = c_1$, then set $\nu_1(s_1) = c_1$ and $\nu_1(s) = \nu_0(s) = s$ for all $s \in S \setminus \{s_1\}$; otherwise $s_1$ is rejected by school $c_1$ and we set $\nu_1 = \nu_0$.

$k$. If there is some student $s$ such that $\nu_{k-1}(s) = s$ ($s$ is unassigned), then student $s$ did not yet apply to all the schools which are acceptable to him. Let $s$ be the student with minimal index among those students. Let $c$ be the school which is most preferred under $P_s$ among the schools to which $s$ did not apply yet.

\(^{13}\)Here one may consider rejecting student $w_3$ since $w_3$ has the lowest priority among the white students. Generally (as in the example used to prove Theorem 1), however each white student could propose to a different school and we would not know which students to put on waiting lists.
(i) If there is \( \mu \in \mathcal{F} \) such that \( \mu(s) = c \) and \( \mu(s') = \nu_{k-1}(s') \) for all students \( s' \) such that \( \nu_{k-1}(s') \neq s' \), then student \( s \) justifiably claims an empty slot at school \( c \) under \( \nu_{k-1} \). Then we set \( \nu_k(s) = c \) and \( \nu_k(s') = \nu_{k-1}(s') \) for all \( s' \in S \setminus \{s\} \);  

(ii) If (i) is not true but there is a student \( s' \) of the same type of \( s \) and student \( s \) justifiably envies student \( s' \) at school \( c \) under \( \nu_{k-1} \), then let \( s' \) be the student which has the lowest priority under \( P_c \) among all the students of type \( \tau(s) \) who are tentatively admitted at school \( c \) under \( \nu_{k-1} \). Then we set \( \nu_k(s) = c \), \( \nu_k(s') = s' \), and \( \nu_k(s'') = \nu_{k-1}(s'') \) for all \( s'' \in S \setminus \{s, s'\} \), i.e. school \( c \) rejects \( s' \) and puts \( s \) on its waiting list; and  

(iii) Otherwise (if (i) and (ii) are not true) we set \( \nu_k = \nu_{k-1} \) and student \( s \) is rejected by school \( c \).  

End: The algorithm ends at a Step \( x \) where \( \nu_x(s) \neq s \) for all \( s \in S \). Then the tentative assignments become final and \( \nu_x \) is the controlled Gale-Shapley assignment for profile \( P_S \).  

The assignment found by the controlled Gale-Shapley algorithm may be wasteful because in the example used to prove Theorem 1 the algorithm finds \( \mu_2 \) and student \( s_2 \) justifiably claims an empty slot at school \( c_3 \) under \( \mu_2 \).  

**THEOREM 3:** For any controlled school choice problem the controlled Gale-Shapley algorithm yields a feasible assignment which is both fair for same types and constrained non-wasteful.  

**Proof:** Let \( P_S \) be a controlled school choice problem and \( \mu \) be the assignment that the controlled Gale-Shapley algorithm finds for \( P_S \). We show that (a) \( \mu \) is feasible, (b) \( \mu \) is fair for same types, and (c) \( \mu \) is constrained non-wasteful.  

For (a) it suffices to show at Step \( k \), any student, who is unassigned under \( \nu_{k-1} \), did not yet propose to all schools on his preference. Suppose that \( \nu_{k-1}(s) = s \) and student \( s \) proposed to all schools before.  

Let student \( s \) have been on a waiting list of a school, say school \( c \), until Step \( h \). Then at Step \( h \) another student \( s' \) proposed to \( c \) and school \( c \) rejected \( s \). But then there were other schools \( c' \) which could have given \( s' \) an empty slot keeping all the other matches of \( \nu_h \) unchanged. But \( s \) did not apply
to any of those empty slots (because otherwise he would have received that slot). Therefore, this is impossible.

If student $s$ was never on a waiting list, then let $h$ be the step where student $s$ applied to his most preferred school. But then there were no $\mu' \in \mathcal{F}$ such that $\mu'(s') = \nu_{h-1}(s')$ for all $s' \in S\setminus\{s\}$ such that $\nu_{h-1}(s') \neq s'$. But then $\nu_{h-1}$ is an impossible waiting list at Step $h - 1$, which contradicts the definition of the controlled Gale-Shapley algorithm.

For (b), suppose that $\mu$ is not fair for same types. Then there is a student $s$ who justifiably envies student $s'$ at school $c$ under $\mu$ and both students $s$ and $s'$ are of the same type. Let $s'$ have lowest priority in $\mu(c)$ among the students who are of type $\tau(s)$. Since $cP_s\mu(s)$, student $s$ applied to school $c$ at some step, say Step $k$.

If $\nu_k(s) = c$, then by $\mu(s) \neq c$, student $s$ was later rejected by school $c$ because some student of type $\tau(s)$ applied to school $c$ and had higher priority than $s$ under $P_c$. Now it is impossible that student $s'$ was put on school $c$’s waiting list later because $s'$ must have had higher priority than $s$ and we have $sP_cs'$.

If $\nu_k(s) \neq c$, then (i) was not possible at Step $k$, i.e. $s$ could not justifiably claim an empty slot at school $c$ under $\nu_{k-1}$. Since (ii) was neither possible, all students of type $\tau(s)$ in $\nu_{k-1}(c)$ had higher priority than $s$. Now it is again impossible that student $s'$ was put on school $c$’s waiting list later because $s'$ must have had higher priority than $s$ and we have $sP_cs'$.

It may be that student $s'$ later justifiably claimed an empty slot at school $c$. This is also impossible because given a waiting list $\nu_x$, for each school $c$ and each type, the students of that type admitted at the school only increases, i.e. it is not possible that $s'$ claims an empty slot later whereas $s$ could not do that earlier.

For (c), suppose that $\mu$ is not constrained non-wasteful. Then a student $s$ justifiably claims an empty slot at school $c$ under $\mu$ and $\mu'$ (where $\mu'(s) = c$ and $\mu'(s') = \mu(s')$ for all $s' \in S\setminus\{s\}$) is fair for same types. Since $s$ justifiably claims an empty slot at school $c$, we have $cP_s\mu(s)$ and $s$ must have proposed to $c$, say at Step $k$, before proposing to $\mu(s)$. The following is true in the controlled Gale-Shapley algorithm: once a student is admitted on a waiting list, then the student can only be removed from the waiting list if another student of the same type is admitted. Therefore, for all types $t$ and all schools
Now by the feasibility of $\mu$ and $s$’s justified claim of an empty slot at $c$ under $\mu$, at Step $k$ there was a feasible assignment $\hat{\mu}$ such that $\hat{\mu}(s) = c$ and $\hat{\mu}(\hat{s}) = \nu_{k-1}(\hat{s})$ for all $\hat{s}$ such that $\nu_{k-1}(\hat{s}) \neq \hat{s}$. Hence, $\nu_k(s) = c$ and $s$ was put on the waiting list of $c$ at Step $k$. Since $\mu(s) \neq c$, at a later step, say Step $k'$, school $c$ rejected student $s$ and admitted a student $s'$. Then student $s'$ must be of the same type as $s$ and at Step $k'$ (i) was not true, i.e. student $s'$ could not justifiably claim an empty slot at school $c$ at Step $k'$. But then by the same property (1) for Step $k'$ no student of type $\tau(s)$ can justifiably claim an empty slot at school $c$ under $\mu$, a contradiction to $s$’s justified claim of an empty slot at $c$ under $\mu$. □

In the controlled Gale-Shapley algorithm students with smaller indices are allowed to propose first (and students may be indexed according to when their applications were received by the controlled school choice program). However, it is easy to verify that the order, in which students are allowed to propose, is irrelevant for the conclusion of Theorem 3. Therefore, at each step alternatively we may choose randomly a student from the students who do not belong to any waiting list. This randomization of the controlled Gale-Shapley algorithm ensures that the algorithm becomes anonymous.\(^{14}\) Of course, in contrast to McVitie and Wilson’s sequential version of the Gale-Shapley algorithm, the controlled Gale-Shapley algorithm may yield different outcomes for different orders. For instance, in the example used to prove Theorem 1, the controlled Gale-Shapley algorithm finds $\mu_2$ when student $s_1$ proposes in Step 1 and it finds $\mu_4$ when student $s_2$ proposes in Step 1.

It turns out that the controlled Gale-Shapley algorithm has another desirable feature: the output assignment is always weakly Pareto-optimal in the sense that there exists no feasible assignment which all students strictly prefer to the output assignment, i.e. if $\mu$ is the assignment found by the controlled Gale-Shapley algorithm for the controlled school choice problem $P_S$, then there exists no feasible assignment $\bar{\mu}$ such that $\bar{\mu}(s)P_s\mu(s)$ for all students $s$. If this important welfare property is not satisfied by an assignment procedure, then one may seriously criticize the use of that procedure.

\(^{14}\)Then using Roth and Rothblum (1999) and Ehlers (2002) it can be shown that in a low information environment it is a weakly dominant strategy for each student to submit his true ranking.
because all students unanimously may strictly prefer another assignment (or another procedure).

**THEOREM 4:** For any controlled school choice problem the controlled Gale-Shapley algorithm yields a feasible assignment which is weakly Pareto-optimal.

**Proof:** Let $P_S$ be a controlled school choice problem and $\mu$ be the assignment that the controlled Gale-Shapley algorithm finds for $P_S$.

Suppose that $\mu$ is not weakly Pareto-optimal. Then there exists another feasible assignment $\bar{\mu}$ such that $\bar{\mu}(s) \succ_P \mu(s)$ for all students $s$. We derive a contradiction as follows: first we show that for any type $t$, the school, at which the last type-$t$ student is admitted, admits no type-$t$ student under $\bar{\mu}$; second we show that under both $\mu$ and $\bar{\mu}$ each school is assigned the same number of students; and third we show that there is a cyclical exchange among all types of the schools at which the last type-$t$ students are admitted in the controlled Gale-Shapley algorithm.

Let $t$ be a type and let the controlled Gale-Shapley algorithm admit the students of type $t$ at their seats specified by $\mu$ in the order $i_1, i_2, \ldots, i_l$. This means that student $i_1$ is the first student of type $t$ who gets assigned to $\mu(s)$ in the controlled Gale-Shapley algorithm and student $i_l$ is the last student of type $t$ who gets assigned to $\mu(s)$ in the controlled Gale-Shapley algorithm. Because $\bar{\mu}(s) \succ_P \mu(s)$ for any student $s$, each student $s$ applies to $\bar{\mu}(s)$ before applying to $\mu(s)$. Since $i_l$ is the last type-$t$ student to be admitted, $i_l$ must justifiably claim an empty slot at school $\mu(i_l)$ in the Step $k$ where $i_l$ proposes to $\mu(i_l)$. But then no student $s$ of type $t$ proposed to $\mu(i_l)$ before proposing to $\mu(s)$ because such a student could have claimed an empty slot at school $\mu(i_l)$ (since $i_l$ was able to claim an empty slot at $\mu(i_l)$ at the later Step $k$). Hence, by $\bar{\mu}(s) \succ_P \mu(s)$ for all $s \in S_t$, no student of type $t$ is assigned to $\mu(i_l)$ under $\bar{\mu}$ and we have $\bar{\mu}^t(\mu(i_l)) = \emptyset$.

Since under both $\bar{\mu}$ and $\mu$ all students are assigned a school, in showing $|\bar{\mu}(c)| = |\mu(c)|$ for all schools $c$ it suffices to show $|\bar{\mu}(c)| \leq |\mu(c)|$ for all schools $c$. Suppose that this is not the case, i.e. for some school $c$ we have $|\bar{\mu}(c)| > |\mu(c)|$. Then for some type $t$ we have $|\bar{\mu}^t(c)| > |\mu^t(c)|$. Then $\bar{\mu}^t(c) \neq \emptyset$. Let $i \in \bar{\mu}^t(c)$. Since $c \succ_P \mu(i)$, student $i$ proposed to $c$ before proposing to $\mu(i)$. Since $\mu(c) < q_c$ and $|\bar{\mu}^t(c)| > |\mu^t(c)|$, student $i$ claimed an empty slot at school $c$ when proposing to it. Now this claim must have been justified since the last type-$t$ student to be admitted, student $i_l$, justifiably
claimed an empty slot at $\mu(i)$ and no student of type $t$ is assigned to $\mu(i)$ under $\bar{\mu}$. Thus, student $i$ must have been assigned an empty slot at $c$ when he proposed to $c$ in the controlled Gale-Shapley algorithm. Since our choice $i \in \bar{\mu}(c)$ was arbitrary, school $c$ must admit at least $|\mu'(c)| + 1$ students of type $t$ in the controlled Gale-Shapley algorithm, a contradiction. Hence, we have shown $|\bar{\mu}(c)| = |\mu(c)|$ for all schools $c$.

For each type $t$, let $i_t$ denote the last type-$t$ student to be admitted at $\mu(s)$ in the controlled Gale-Shapley algorithm and let $c' = \mu(i_t)$. Since $|\bar{\mu}(c')| = |\mu(c')|$ and $\bar{\mu}(c') = \emptyset$, there exists at least one type $t'$ such that $|\bar{\mu}'(c')| > |\mu'(c')|$ or equivalently some students of type $t'$ would like to claim the slot of $i_t'$ at school $c'$. For the moment, let us treat types as agents and say that type $t$ is endowed with an empty slot at $c'$ and type $t'$ would like to claim a slot at $c'$. Now, similarly as above, type $t'$ is also endowed with an empty slot at $c'$ and some type $t''$ would like to claim that slot. Because the set of types is finite and each type is endowed with exactly one empty slot, there must exist at least one cyclical exchange from $\mu$ to $\bar{\mu}$: there are types $t_1, \ldots, t_m$ such that $t_1$ claims the slot $c_2'$, $t_2$ claims the slot $c_3'$, $\ldots$, and $t_m$ claims the slot $c_1'$. Now choose the type $t$ such that type $t$ is part of a cyclical exchange and among the types, which are part of a cyclical exchange, type $t$ is the first type to admit all students in the controlled Gale-Shapley algorithm. This means that $i_t'$ is admitted at $\mu(c')$ before any other type $t'$, which is part of a cyclical exchange, admits $i_t'$ at $c'$'. Because $t$ is part of a cyclical exchange, type $t$ claims the “endowment” of another type, say type $t'$. Because $i_t'$ is the last type-$t$ student to be admitted, all type-$t$ students, who would like to claim the slot at $c'$, proposed to $c'$ before. Because type $t$ is the first type to admit all students among all types which are part of a cyclical exchange, at this step both the slot at $c''$ was empty and this cyclical exchange was feasible when the type-$t$ students proposed to $c'$. But then at this step this type-$t$ student justifiably claims an empty slot at school $c'$ and the controlled Gale-Shapley algorithm would have assigned this type-$t$ student to school $c'$, which is a contradiction. □

Remark 2 An immediate consequence of Theorem 4 is that it is impossible to make all white students strictly better off by reassigning their seats and seats left empty among white students while keeping all other students enrolled at their schools. More precisely, for any type $t$ the assignment $\mu_t$ is weakly Pareto-optimal in the sense that it is impossible to make all students
of type $t$ better off by reassigning (in a feasible way) their seats specified by $\mu$ and the seats left empty by $\mu$, i.e. there is no feasible assignment $\bar{\mu}$ such that $\bar{\mu}(s) P_s \mu(s)$ for all $s \in S_t$ and $\bar{\mu}(s') = \mu(s')$ for all $s' \in S \setminus S_t$. This is easily seen by applying Theorem 4 to the problem reduced for type-$t$ students where any school $c$ has $q_c + |\mu^t(c)| - |\mu(c)|$ empty seats available and $S_t$ is the set of students. Applying the controlled Gale-Shapley algorithm to $P_{S_t}$ yields $\mu^t$ and $\mu^t$ needs to be weakly Pareto-optimal by Theorem 4, the desired conclusion.

**Remark 3** Another immediate consequence of Theorem 4 is that it is impossible to make all white students weakly better off by fairly reassigning their seats among white students while keeping all other students enrolled at their schools and all empty slots empty. More precisely, if $\mu$ is the output of the controlled Gale-Shapley algorithm, then for any type $t$ the assignment $\mu^t$ is “best” from the type-$t$ students’ point of view in the following sense: there is no other feasible assignment $\bar{\mu}$ such that (i) $|\bar{\mu}^t(c)| = |\mu^t(c)|$ for all schools $c$, (ii) $\bar{\mu}$ is fair for same types, and (iii) $\bar{\mu}$ Pareto dominates for type-$t$ students the assignment $\mu^t$.

**Remark 4** In the college admissions problem we know that the Gale-Shapley algorithm calculates the stable matching which is “overall best” from the students’ point of view: there is no other stable matching which Pareto dominates the matching calculated by the Gale-Shapley algorithm. Unfortunately this is another feature which does not carry over from college admissions to controlled school choice. In school choice the order, in which students propose, is intrinsic to the outcome of the controlled Gale-Shapley algorithm. For instance, consider a problem consisting of two white students $\{w_1, w_2\}$, two black students $\{b_1, b_2\}$, and four schools $\{c_1, c_2, c_3, c_4\}$. Each school has a capacity of one ($q_c = 1$ for all schools $c$). Any ceiling of any type $t$ (white or black) is equal to one and any floor of any type $t$ (black or white) is equal to zero. The students’ preferences and the colleges’ priorities are given below

---

15This follows from the fact that when allocating the seats of $\mu^t$ to type-$t$ students, $\mu^t$ is the output of the controlled Gale-Shapley algorithm restricted to type-$t$ students. But the controlled Gale-Shapley algorithm restricted to type-$t$ students and the seats of $\mu^t$ is identical with McVitie and Wilson (1970)’s version of the Gale-Shapley algorithm. Then $\mu^t$ is the output of the Gale-Shapley algorithm restricted to type-$t$ students and the seats of $\mu^t$ and we know that for type-$t$ students $\mu^t$ is most preferred among all assignments which are fair for same types.
Suppose that students apply in the order $w_1 - b_1 - w_2 - b_2$ in the controlled Gale-Shapley algorithm. Then student $w_1$ applies first to school $c_1$ and therefore, school $c_1$ will fill its slot with a white student. In the second step student $b_1$ applies to school $c_2$ and therefore, school $c_2$ will fill its slot with a black student. Now it is easy to verify that

$$\mu = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 \\ w_2 & b_2 & w_1 & b_1 \end{pmatrix}$$

is the output of the controlled Gale-Shapley algorithm for the order $w_1 - b_1 - w_2 - b_2$. Obviously all students weakly prefer the feasible assignment (obtained from $\mu$ when students $w_2$ and $b_2$ exchange their seats)

$$\bar{\mu} = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 \\ b_2 & w_2 & w_1 & b_1 \end{pmatrix},$$

which is fair for same types and constrained non-wasteful.

## 6 Incentive Compatibility

Apart from students’ preferences all components of a controlled school choice problem are exogenously determined (like the capacities of the schools) or given by law (like the priority profile and the controlled choice constraints). The only information which is private are students’ preferences over schools. They need to be stated by the students to the school choice program. Since students must be assigned schools for any possible reported profile, the program has to be based on a mechanism selecting an assignment for each problem. In a controlled school choice program the mechanism should respect the legal constraints imposed by the state. A mechanism is **(legally) feasible** if it selects a feasible assignment for any reported profile.
Any program would like to elicit the true preferences from students. If students would misreport, then the assignment chosen by the program is based on false preferences and may be highly unfair for the true preferences.

Avoiding this problem means constructing a mechanism where no student has ever an incentive to misrepresent his true preference for any preferences reported by the other agents. Any mechanism which makes truthful revelation of preferences a dominant strategy for each student is called (dominant strategy) incentive compatible. A feasible mechanism is fair for same types if it selects for any controlled school choice problem a feasible assignment which is fair for same types. Analogously we define non-wastefulness and constrained non-wastefulness, respectively, for a mechanism.

In contrast to the school choice problems studied in previous papers it is impossible to construct a mechanism which is incentive compatible, fair for same types and constrained non-wasteful while respecting the diversity constraints given by law. Therefore, it is impossible to choose for each profile an order in which students propose in the controlled Gale-Shapley algorithm such that this mechanism becomes incentive compatible.

THEOREM 5: In controlled school choice there is no feasible mechanism which is incentive compatible, fair for same types and constrained non-wasteful.

Proof: The proof is by means of an example. Consider the following problem consisting of three schools \( \{c_1, c_2, c_3\} \) and two students \( \{s_1, s_2\} \). Each school has a capacity of two (\( q_c = 2 \) for all schools \( c \)). The type space consists of a single type \( t \), i.e. both students are of the same type \( t \). The ceiling for type \( t \) is equal to two for each school (\( \overline{q}_t = 2 \) for all schools \( c \)). School \( c_1 \) has a floor for type \( t \) of \( q^t_{c_1} = 1 \) and both other schools have a floor of 0 for type \( t \). Schools \( c_1 \) and \( c_2 \) give higher priority to student \( s_2 \) whereas school \( c_3 \) gives higher priority student \( s_1 \). The students’ preferences are given by \( c_2 P_{s_1} c_1 P_{s_1} c_3 P_{s_1} s_1 \) and \( c_3 P_{s_2} c_1 P_{s_2} c_2 P_{s_2} s_2 \). This information is summarized in
Next we determine the set of feasible assignments. Feasibility requires that one of
the students is assigned school $c_1$ and each student is assigned a school.
Then it is straightforward to verify that

$$
\mu_1 = \begin{pmatrix}
  c_1 & c_2 & c_3 \\
  s_1 & \emptyset & s_2
\end{pmatrix},
\mu_2 = \begin{pmatrix}
  c_1 & c_2 & c_3 \\
  s_1 & s_2 & \emptyset
\end{pmatrix},
\mu_3 = \begin{pmatrix}
  c_1 & c_2 & c_3 \\
  s_2 & \emptyset & s_1
\end{pmatrix},
\mu_4 = \begin{pmatrix}
  c_1 & c_2 & c_3 \\
  s_2 & s_1 & \emptyset
\end{pmatrix},
\mu_5 = \begin{pmatrix}
  c_1 & c_2 & c_3 \\
  \{s_1, s_2\} & \emptyset & \emptyset
\end{pmatrix}
$$

is the set of all feasible assignments.

It is easy to check that $\mu_1$ and $\mu_4$ are the only feasible assignments which are both fair for same types and non-wasteful for this controlled school choice problem. Furthermore, under $P_S$,

(i) $\mu_2$ and $\mu_5$ are not constrained non-wasteful since $s_2$ justifiably claims an empty slot at $c_3$ under both $\mu_2$ and $\mu_5$ and $\mu_1$ is fair for same types, and

(ii) $\mu_3$ is not constrained non-wasteful since $s_1$ justifiably claims an empty slot at $c_2$ under $\mu_3$ and $\mu_4$ is fair for same types.

Any feasible mechanism which is both fair for same types and constrained non-wasteful must select either the assignment $\mu_1$ or the assignment $\mu_4$. We will show that in each case there is a student who profitably manipulates the mechanism.

*Case 1:* The mechanism selects $\mu_1$. 

---

Table 3.

<table>
<thead>
<tr>
<th>$P$ :</th>
<th>$P_{c_1}$</th>
<th>$P_{c_2}$</th>
<th>$P_{c_3}$</th>
<th>$P_{s_1}$</th>
<th>$P_{s_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_2$</td>
<td>$s_2$</td>
<td>$s_1$</td>
<td>$c_2$</td>
<td>$c_3$</td>
<td></td>
</tr>
<tr>
<td>$s_1$</td>
<td>$s_1$</td>
<td>$s_2$</td>
<td>$c_1$</td>
<td>$c_1$</td>
<td>$c_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$c_3$</td>
<td>$c_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$s_1$</td>
<td>$s_2$</td>
</tr>
</tbody>
</table>

capacities $q_{c_1} = 2$ $q_{c_2} = 2$ $q_{c_3} = 2$

ceiling for $t$ $q^t_{c_1} = 2$ $q^t_{c_2} = 2$ $q^t_{c_3} = 2$

floor for $t$ $q^t_{c_1} = 1$ $q^t_{c_2} = 0$ $q^t_{c_3} = 0$
Under \( \mu_1 \) student \( s_1 \) is assigned school \( c_1 \). We will show that student \( s_1 \) gains by misreporting his true preference. Suppose that student \( s_1 \) states the (false) preference \( P'_{s_1} \) given by \( c_2 P'_{s_1} c_3 P'_{s_1} c_1 P'_{s_1} s_1 \), and student \( s_2 \) were to report his true preference \( P_{s_2} \). Keeping all other components of the above problem fixed, in the new problem the students’ preferences are \( P'_S = (P'_{s_1}, P_{s_2}) \).

In the new problem under \( \mu_1 \) student \( s_1 \) justifiably envies student \( s_2 \) at school \( c_3 \) since (f1) \( \mu_1(s_1) = c_1, c_3 P'_{s_1} c_1 \) and \( s_1 P_{c_3} s_2 \), and (f2) \( \tau(s_1) = \tau(s_2) \). Furthermore, under \( P'_S \),

(i) \( \mu_1 \) and \( \mu_2 \) are not fair for same types, and

(ii) \( \mu_3 \) and \( \mu_5 \) are not constrained non-wasteful since \( s_1 \) justifiably claims an empty slot at \( c_2 \) under both \( \mu_3 \) and \( \mu_5 \) and \( \mu_1 \) is fair for same types.

Thus, the unique feasible assignment, which is both fair for same types and non-wasteful for the new problem, is \( \mu_4 \). Hence, any feasible mechanism, which is both fair for same types and constrained non-wasteful, must select the assignment \( \mu_4 \) for the new problem. Under \( \mu_4 \) student \( s_1 \) is assigned school \( c_2 \) which is strictly preferred to \( c_1 \) under the true preference \( P_{s_1} \). Thus student \( s_1 \) does better by stating \( P'_{s_1} \) than by stating his true preference \( P_{s_1} \), and the mechanism is not incentive compatible.

Case 2: The mechanism selects \( \mu_4 \).

Under \( \mu_4 \) student \( s_2 \) is assigned school \( c_1 \). Similarly as in Case 1 we will show that student \( s_2 \) gains by misreporting his preference. Suppose that student \( s_2 \) states the (false) preference \( P'_{s_2} \) given by \( c_3 P'_{s_2} c_2 P'_{s_2} c_1 P'_{s_2} s_2 \), and student \( s_1 \) were to report his true preference \( P_{s_1} \). Keeping all other components of the above problem fixed, in the new problem the students’ preferences are \( P'_S = (P_{s_1}, P'_{s_2}) \).

In the new problem under \( \mu_4 \) student \( s_2 \) justifiably envies student \( s_1 \) at school \( c_2 \) since (f1) \( \mu_4(s_2) = c_1, c_2 P'_{s_2} c_1 \) and \( s_2 P_{c_2} s_1 \), and (f2) \( \tau(s_2) = \tau(s_1) \). Furthermore, under \( P'_S \),

(i) \( \mu_4 \) is not fair for same types,

(ii) \( \mu_2 \) and \( \mu_5 \) are not constrained non-wasteful since \( s_2 \) justifiably claims an empty slot at \( c_3 \) under both \( \mu_2 \) and \( \mu_5 \) and \( \mu_1 \) is fair for same types, and
(iii) $\mu_3$ is not constrained non-wasteful since $s_1$ justifiably claims an empty slot at $c_1$ under $\mu_3$ and $\mu_5$ is fair for same types.

Thus, the unique feasible assignment, which is both fair for same types and non-wasteful for the new problem, is $\mu_1$. Hence, any feasible mechanism, which is both fair for same types and constrained non-wasteful, must select the assignment $\mu_1$ for the new problem. Under $\mu_1$ student $s_2$ is assigned school $c_3$ which is strictly preferred to $c_1$ under the true preference $P_{s_2}$. Thus student $s_2$ does better by stating $P'_{s_2}$ than by stating his true preference $P_{s_2}$, and the mechanism is not incentive compatible. □

**Remark 5** The conclusion of Theorem 5 remains unchanged when constrained non-wastefulness is replaced by non-wastefulness. Of course, by Theorem 1 we know that the set of feasible assignments which are both fair for same types and non-wasteful may be empty. Therefore, it is meaningful to require a mechanism to satisfy simultaneously both fairness for same types and non-wastefulness only if there exist feasible assignments which are both fair for same types and non-wasteful. Hence we say that a feasible mechanism is **fair for same types and non-wasteful** if it selects for any controlled school choice problem a feasible assignment which is fair for same types and non-wasteful whenever such an assignment exists. Now the proof of Theorem 5 remains true when constrained non-wastefulness is replaced by non-wastefulness. Hence in controlled school choice there is no feasible mechanism which is incentive compatible, fair for same types and non-wasteful.

**Remark 6** The non-existence of feasible mechanisms, which are incentive compatible, fair for same types and (constrained) non-wasteful, unambiguously shows that controlled school choice is not equivalent to college admission. In all models of school choice studied so far it was possible to connect the school choice problem to the college admissions problem and show that the Gale-Shapley student optimal algorithm is a mechanism which is non-wasteful, fair, and incentive compatible. This was due to the absence of diversity constraints (the floors) which are present in controlled choice.

In college admissions any mechanism, which is incentive compatible for students, chooses for each problem the extreme of the lattice of stable matchings which students prefer over any other stable matching. In controlled school choice there is not always a unique candidate for a feasible assignment which is fair for same types and (constrained) non-wasteful. This provides additional reason for Theorem 5 and Remark 5, i.e. for the non-existence of
feasible mechanisms which are incentive compatible, fair for same types and (constrained) non-wasteful.

In the example used to prove Theorem 5 we know that

\[ \mu_1 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & \emptyset & s_2 \end{pmatrix} \quad \text{and} \quad \mu_4 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_1 & \emptyset \end{pmatrix} \]

are the only feasible assignments which are both fair for same types and (constrained) non-wasteful for this problem.

Student \( s_1 \) prefers \( c_2 \) to \( c_1 \) under \( P_{s_1} \) and student \( s_2 \) prefers \( c_3 \) to \( c_1 \) under \( P_{s_2} \). Thus \( s_1 \) strictly prefers \( \mu_4 \) to \( \mu_1 \) under \( P_{s_1} \) and \( s_2 \) strictly prefers \( \mu_1 \) to \( \mu_4 \) under \( P_{s_2} \). Hence students’ preferences are opposed over the (only) two feasible assignments, which are both fair for same types and (constrained) non-wasteful, and there is no feasible, fair for same types and (constrained) non-wasteful assignment which both students prefer to any other feasible assignment which is both fair for same types and non-wasteful.

When computing the “minimum” \( \wedge \) of \( \mu_1 \) and \( \mu_4 \) (by assigning each student to the school which he least prefers from \( \mu_1 \) and \( \mu_4 \)) we obtain the assignment

\[ \mu_1 \wedge \mu_4 = \begin{pmatrix} c_1 & c_2 & c_3 \\ \{s_1, s_2\} & \emptyset & \emptyset \end{pmatrix} \]

which is feasible but not (constrained) non-wasteful.

Remark 7 Theorem 5 implies that for any order of the students the controlled Gale-Shapley algorithm is not incentive compatible. Due to this fact students may misrepresent their preferences over schools. Now if the students play a Nash equilibrium (NE), what are the properties of the outcome (or the assignment) of any NE? It is easy to see that the outcome of any NE must be constrained non-wasteful.\(^{16}\) Unfortunately, the outcome of a NE may not be fair for same types according to students’ true preferences.

For instance, consider a problem consisting of three students \( \{s_1, s_2, s_3\} \) (all of the same type) and three schools \( \{c_1, c_2, c_3\} \). Each school has a capacity of one \( (q_c = 1 \text{ for all schools } c) \). Any ceiling of any type \( t \) is equal to one and

\(^{16}\)Otherwise a student would justifiably claim an empty slot and after assigning him this empty slot the resulting assignment is fair for same types. Then this student profits from changing his preference such that he proposes to this school before proposing to the school to which he is assigned to.
any floor of any type \( t \) is equal to zero. The students’ preferences and the colleges’ priorities are given below:

\[
\begin{array}{ccccccc}
P : & P_{c_1} & P_{c_2} & P_{c_3} & P_{s_1} & P_{s_2} & P_{s_3} \\
& s_2 & s_1 & s_3 & c_1 & c_2 & c_2 \\
& s_1 & s_3 & s_2 & c_2 & c_1 & c_3 \\
& s_3 & s_2 & s_1 & c_3 & c_3 & c_1 \\
\end{array}
\]

Now if student \( s_3 \) reports \( \bar{P}_{s_3} : c_3c_1c_2 \), then independently of the order, in which students propose, the controlled Gale-Shapley algorithm chooses for \( \bar{P}_S = (P_{s_1}, P_{s_2}, \bar{P}_{s_3}) \) the assignment

\[
\bar{\mu} = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_2 & s_3 \end{pmatrix}.
\]

Obviously \( \bar{\mu} \) is not fair for same types for the true profile since student \( s_3 \) justifiably envies student \( s_2 \) at school \( c_2 \) (and all students are of the same type. It is easy to check that \( \bar{P}_S \) is a NE in the controlled Gale-Shapley algorithm and that \( \bar{P}_S \) is even robust to deviations of any group of students, i.e. \( \bar{P}_S \) is a strong Nash equilibrium. However, this feature is not peculiar to controlled school choice since the above problem is a college admissions problem and we know that in college admissions the outcome of a NE may not be stable according to the true preferences.\(^{17}\)

Any controlled school choice program must give up non-wastefulness, fairness for same types or incentive compatibility. It will wonder whether an existence result reemerges if we give up exactly one of our two basic requirements, namely non-wastefulness or fairness for same types.

Since in real life often the number of available seats is approximately the same as the number of students, potential justified claims of empty seats occur less frequently than potential justified envy. Hence, a school choice program may be ready to give up non-wastefulness while retaining fairness for same types and incentive compatibility. We will demonstrate that this weakening results in existence.\(^{18}\)

\(^{17}\)In school choice problems without control and legal constraints, Ergin and Sönmez (2006) consider revelation games induced by the Boston school choice mechanism and the Gale-Shapley algorithm.

\(^{18}\)Giving up fairness for same types also results in existence. A serial dictatorship (which is used frequently for the allocation of indivisible objects) is a feasible mechanism which is
Example 3. A feasible mechanism which is both fair for same types and incentive compatible.

Fix a feasible assignment, say $\mu$. We relate any controlled school choice problem with a college admissions problem in the following way: break any school $c$ into $k$ schools $\{c(t_1), \ldots, c(t_k)\}$ where $|T| = k$ and $c(t)$ is the part of school $c$ wanting to fill slots with students of type $t$. The capacity of school $c(t)$ is $q_{c(t)} = |\mu^t(c)|$ and the preference of $c(t)$ ranks only students of type $t$ acceptable, in the same order as $P_c$. Note that some slots are wasted at school $c$ if $|\mu(c)| < q_c$. Any student replaces on his preference school $c$ by $|T|$ copies of $c$ in the order $c(t_1), c(t_2), \ldots, c(t_k)$. Then determine the student optimal matching of this related problem. Because (i) all students rank all schools as acceptable, (ii) for any type $t$ there are exactly $\sum_{c \in C} q_{c(t)} = \sum_{c \in C} |\mu^t(c)| = |S_t|$ slots available and (iii) any school $c(t)$ ranks acceptable exactly all students of type $t$, the student optimal matching $\bar{\mu}$ of the related problem satisfies for all types $t$ and all schools $c$, $\bar{\mu}^t(c(t)) \subseteq S_t$ and $|\bar{\mu}(c(t))| = q_{c(t)} = |\mu^t(c)|$.

Thus the feasibility of $\mu$ implies that the student optimal matching of the related problem is a feasible assignment of the controlled school choice problem. We know that the Gale-Shapley student optimal mechanism is incentive compatible. Furthermore the stability of the student optimal matching in the related problem implies that there is no student envying justifiably another student of the same type. Thus the “related” Gale-Shapley mechanism is a feasible mechanism which is both fair for same types and incentive compatible. The mechanism is non-wasteful only if the initial assignment $\mu$ filled all available slots at each school. Furthermore the mechanism is fair (across types) only if all students are of the same type. When all students are of the same type, the floor at a school may be represent the number of students which is necessary not to shut down that school.

Observe that the above mechanism is “rigid”: in Example 3 for each type $t$, the slots, which will be filled with type-$t$ students, are exogenously given by the feasible assignment $\mu$. This inflexibility was the price for in-
centive compatibility of this mechanism. In general this price also includes giving up weak Pareto-optimality because due to the inflexibility all students may be strictly better off with another feasible assignment compared to the assignment chosen by the mechanism in Example 3.

7 Recommendation to School Choice Programs

Without controlled choice and legal constraints, the Gale-Shapley algorithm eliminates any justified envy and makes truthful revelation of preferences a dominant strategy for students (Abdulkadiroğlu and Sönmez, 2003). Once controlled choice constraints are imposed it may be impossible to eliminate any justified envy. The legal constraints allow to eliminate justified envy only among students of the same type (and not of different types). Any state in the United States needs to decide whether controlled choice and legal constraints are more important or whether elimination of any justified envy is more important. In university admissions it is likely that fairness is regarded more important and in those contexts the Gale-Shapley algorithm assigns students to schools in a satisfactory manner. In school choice it is unlikely that controlled choice and legal constraints are ignored completely and envy can be eliminated only among students of the same race.

Once the controlled choice and the legal constraints are accepted, any school choice program needs to decide whether incentive compatibility is more important or whether weak Pareto-optimality or/and constrained non-wastefulness is more important. If the program insists on incentive compatibility, the incentive compatible mechanism we propose basically segregates the problem into several problems, one for each race, and applies the Gale-Shapley algorithm to each problem separately. Making truth telling a dominant strategy brings many serious flaws with it (even though the assignment is based on true preferences). Any incentive compatible mechanism may implement assignments which are highly inefficient and highly wasteful (of empty seats). Due to these many school choice programs may prefer a mechanism which is weakly Pareto-optimal and constrained non-wasteful: this is achieved in practice by the controlled Gale-Shapley algorithm. Furthermore, in low information environments this mechanism is immune to manipulation.

Controlled choice comes with a price. Any program has to disregard at least one desirable property when following the state’s laws. Our results raise the question whether some state laws need to be modified.
We formulate fairness across types precisely below and show that Theorem 1 and Theorem 5 remain unchanged when fairness across types replaces fairness for same types and (constrained) non-wastefulness.

We say that student $s$ **justifiably envies student $s'$ at school $c$ under the feasible assignment** $\mu$ if there exists another feasible assignment $\mu'$ such that

(f1) $\mu(s') = c$, $cP_s\mu(s)$ and $sP_c s'$,

(f2) $\mu'(s) = c$, $\mu'(s') \neq c$, and $\mu'(\hat{s}) = \mu(\hat{s})$ for all $\hat{s} \in S \setminus \{s, s'\}$.

Because $\mu'$ is feasible, (f2) simply says that $(\mu(c) \setminus \{s'\}) \cup \{s\}$ respects the controlled choice constraints at school $c$ and student $s'$ can be enrolled at school $c' = \mu'(s')$ such that $(\mu(c') \setminus \{s\}) \cup \{s'\}$ respects the controlled choice constraints at $c'$; in other words assigning $s$ a slot at $c$, $s'$ a slot at $c'$, and keeping all the other assignments intact does not violate any controlled choice constraint at any school.

A feasible assignment $\mu$ is **fair across types** if no student justifiably envies any student. Independently of his own type, a student is allowed to envy any student.

Since fairness for same types is a weaker requirement than fairness across types, Theorem 1 also shows that there may not exist any feasible assignment which is both fair across types and non-wasteful in a controlled school choice problem. Unfortunately fairness across types alone may be enough for this non-existence result. This can be seen by modifying the example, which is used in the proof of Theorem 1, by introducing a third (dummy) student, who is ranked at the bottom of all priority rankings and whose type is different than $t$.

**THEOREM 1’**: The set of feasible assignments which are fair across types may be empty in a controlled school choice problem.

**Proof**: The proof is by means of an example. The basic idea is similar to the one used in proving Theorem 1. Consider the following problem consisting of three schools $\{c_1, c_2, c_3\}$ and three students $\{s_1, s_2, s_3\}$. Each school has a capacity of one ($q_c = 1$ for all schools $c$). The type space consists of two types $t_1$ and $t_2$. Students $s_1$ and $s_2$ are of type $t_1$ whereas student $s_3$ is of type $t_2$. 

36
For all types the ceiling is equal to one at all schools ($q^t_c = 1$ for all types $t$ and all schools $c$). School $c_2$ has a floor for type $t_1$ of $q^{t_1}_{c_2} = 1$. All other floors are equal to zero. The schools’ priorities are given by $s_2 P_{c_1} s_1 P_{c_3} s_3$, $s_2 P_{c_2} s_1 P_{c_3} s_3$, and $s_1 P_{c_3} s_2 P_{c_3} s_3$. The students’ preferences are given by $c_1 P_{s_1} c_3 P_{s_3} s_1$, $c_3 P_{s_2} c_1 P_{s_2} s_2$ and $c_2 P_{s_3} c_3 P_{s_3} c_1 P_{s_3} s_3$. This information is summarized in Table 2.

<table>
<thead>
<tr>
<th>$P :$</th>
<th>$P_{c_1}$</th>
<th>$P_{c_2}$</th>
<th>$P_{c_3}$</th>
<th>$P_{s_1}$</th>
<th>$P_{s_2}$</th>
<th>$P_{s_3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_2$</td>
<td>$s_2$</td>
<td>$s_1$</td>
<td>$c_1$</td>
<td>$c_3$</td>
<td>$c_2$</td>
<td></td>
</tr>
<tr>
<td>$s_1$</td>
<td>$s_1$</td>
<td>$s_2$</td>
<td>$c_3$</td>
<td>$c_1$</td>
<td>$c_3$</td>
<td></td>
</tr>
<tr>
<td>$s_3$</td>
<td>$s_3$</td>
<td>$s_3$</td>
<td>$c_2$</td>
<td>$c_2$</td>
<td>$c_1$</td>
<td></td>
</tr>
</tbody>
</table>

capacities $q_{c_1} = 1$ $q_{c_2} = 1$ $q_{c_3} = 1$
floor for $t_1$ $q^{t_1}_{c_1} = 0$ $q^{t_1}_{c_2} = 1$ $q^{t_1}_{c_3} = 0$
floor for $t_2$ $q^{t_2}_{c_1} = 1$ $q^{t_2}_{c_2} = 1$ $q^{t_2}_{c_3} = 1$
floor for $t_2$ $q^{t_2}_{c_1} = 0$ $q^{t_2}_{c_2} = 0$ $q^{t_2}_{c_3} = 0$

Next we determine the set of assignments which are both feasible and fair across types for this problem. Feasibility requires that student $s_1$ or student $s_2$ is assigned school $c_2$ and all students are enrolled at a school. Therefore,

$$
\mu_1 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_2 & s_3 \end{pmatrix} \quad s_2 \text{ envies } s_1 \\
\mu_2 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_1 & s_3 \end{pmatrix},
$$

$$
\mu_3 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_3 & s_1 & s_2 \end{pmatrix},
$$

are the only assignments which are feasible. Now (as indicated above)

(i) $\mu_1$ is not fair across types because $s_2$ justifiably envies $s_1$ at $c_1$,

(ii) $\mu_2$ is not fair across types because $s_2$ justifiably envies $s_3$ at $c_3$,

(iii) $\mu_3$ is not fair across types because $s_1$ justifiably envies $s_2$ at $c_3$,

(iv) $\mu_4$ is not fair across types because $s_1$ justifiably envies $s_3$ at $c_1$.  

37
Hence there is no assignment which is both feasible and fair across types. □

We discuss the robustness of Theorem 1’ subject to relaxing type specific constraints. If we relax the ceilings and the capacities such that each school’s capacity is equal to the total number of students and the ceiling for each type at each school is equal to the total number of students of that type, then the logic of the proof of Theorem 1’ does not remain true. Indeed, if the ceilings are relaxed (as described above), then there exists always an assignment which is both feasible and fair across types: break any school $c \neq c_1$ into $|T|$ schools \{$c(t_1), \ldots, c(t_k)$\} where $c(t)$ is the part of school $c$ wanting to fill $q^t_c$ positions with students of type $t$ (and the preference of $c(t)$ ranks only students of type $t$ as acceptable, in the same order as $P_c$), and break school $c_1$ into $|T|$ schools \{$c_1(t_1), \ldots, c_1(t_k)$\} where $c_1(t)$ is the part of school $c_1$ with wanting to fill $|S_t| - \sum_{c \neq c_1} q^t_c$ positions with students of type $t$ (and the preference of $c(t)$ ranks only students of type $t$ as acceptable, in the same order as $P_c$). Each school $c$ is replaced by the $|T|$ copies of $c$ in the order $c(t_k), c(t_{k-1}), \ldots, c(t_1)$ on any student’s preference relation. The related problem has floor zero and falls into the class of problems studied by Abdulkadiroğlu (2005). Thus his results apply. We claim that the assignment determined by the Gale-Shapley student optimal algorithm is feasible and fair across types. Feasibility simply follows from the fact that for each type there are exactly the same number of positions available as there are students of that type. In showing fairness across types, suppose that a student $s$ justifiably envies student $s'$ at school $c$. Then students $s$ and $s'$ must have different types (since otherwise the Gale-Shapley student optimal assignment is not stable). Furthermore, either student $s$ is not assigned to school $c_1$ or student $s'$ is not assigned to school $c_1$. If student $s$ is not assigned to school $c_1$, then $s$ cannot be removed from his assigned school without violating constraints (since $\tau(s) \neq \tau(s')$). If student $s'$ is not assigned to school $c_1$, then $s'$ cannot be removed from his assigned school without violating constraints (since $\tau(s') \neq \tau(s)$). Hence, in both cases envy is not justified and the Gale-Shapley student optimal assignment is fair across types for the relaxed problem.

Similarly as before, a feasible mechanism is **fair across types** if it selects an assignment which is both feasible and fair across types for any controlled school choice problems having a non-empty set of feasible assignments which

---

19Recall that $S_t = \{s \in S : \tau(s) = t\}$. 

38
are fair across types. Parallel to Theorem 5 and Remark 5, fairness across types is incompatible with feasibility and incentive compatibility.

**THEOREM 5’**: In controlled school choice there is no feasible mechanism which is both incentive compatible and fair across types.

*Proof*: The proof is by means of an example. The idea of the proof is similar to the one used in proving Theorem 4. Consider the following problem consisting of three schools \( \{c_1, c_2, c_3\} \) and three students \( \{s_1, s_2, s_3\} \). Each school has a capacity of one \( (q_c = 1 \text{ for all schools } c) \). The type space consists of two types \( t_1 \) and \( t_2 \). Students \( s_1 \) and \( s_2 \) are of type \( t_1 \) whereas student \( s_3 \) is of type \( t_2 \). For all types the ceiling is equal to one at all schools \( (\bar{q}_t^e = 1 \text{ for all types } t \text{ and all schools } c) \). School \( c_1 \) has a floor for type \( t_1 \) of \( q_{t_1}^{f_1} = 1 \). All other floors are equal to zero. The schools’ priorities are given by \( s_2 P_{c_1} s_1, s_2 P_{c_2} s_1 P_{c_3} s_3 \) and \( s_1 P_{c_1} s_2 P_{c_2} s_3 \). The students’ preferences are given by \( c_2 P_{s_1} c_1 P_{s_2} s_3, c_3 P_{s_2} c_1 P_{s_3} c_2 P_{s_2} s_2 \) and \( c_2 P_{s_3} c_3 P_{s_3} c_1 P_{s_3} s_3 \). This information is summarized in Table 4.

<table>
<thead>
<tr>
<th>( P )</th>
<th>( P_{c_1} )</th>
<th>( P_{c_2} )</th>
<th>( P_{c_3} )</th>
<th>( P_{s_1} )</th>
<th>( P_{s_2} )</th>
<th>( P_{s_3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_2 )</td>
<td>( s_2 )</td>
<td>( s_1 )</td>
<td>( c_2 )</td>
<td>( c_3 )</td>
<td>( c_2 )</td>
<td></td>
</tr>
<tr>
<td>( s_1 )</td>
<td>( s_1 )</td>
<td>( s_2 )</td>
<td>( c_1 )</td>
<td>( c_1 )</td>
<td>( c_3 )</td>
<td></td>
</tr>
<tr>
<td>( s_3 )</td>
<td>( s_3 )</td>
<td>( s_3 )</td>
<td>( c_3 )</td>
<td>( c_2 )</td>
<td>( c_1 )</td>
<td></td>
</tr>
</tbody>
</table>

Next we determine the set of assignments which are both feasible and fair across types for this problem. Feasibility requires that student \( s_1 \) or student \( s_2 \) is assigned school \( c_1 \) and all students are enrolled at a school. If student \( s_1 \) is assigned school \( c_1 \), then \( s_2 \) needs to be assigned school \( c_3 \) since otherwise \( s_3 \) is assigned school \( c_3 \), \( s_2 \) school \( c_2 \), and \( s_2 \) justifiably envies \( s_3 \) at \( c_3 \). Similarly, if student \( s_2 \) is assigned school \( c_1 \), then \( s_1 \) needs to be assigned school \( c_2 \) since otherwise \( s_3 \) is assigned school \( c_2 \), \( s_1 \) school \( c_3 \), and \( s_1 \) justifiably envies \( s_3 \) at

39
Now it is straightforward to verify that
\[
\mu = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_3 & s_2 \end{pmatrix} \quad \text{and} \quad \bar{\mu} = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_1 & s_3 \end{pmatrix}
\]
are the only assignments which are both feasible and fair across types for this problem.

Any mechanism which is both feasible and fair across types must select either the assignment $\mu$ or the assignment $\bar{\mu}$. We will show that in each case there is a student who profitably manipulates the mechanism.

**Case 1:** The mechanism selects $\mu$.

Under $\mu$ student $s_1$ is assigned school $c_1$. We will show that student $s_1$ gains by misreporting his true preference. Suppose that student $s_1$ states the (false) preference $P'_s$ given by $c_2 P'_{s_1} c_3 P'_{s_1} c_1 P'_{s_1} s_1$, and all other students were to state their true preferences. Keeping all other components of the above problem fixed, in the new problem the students' preferences are $P'_S = (P'_{s_1}, P_{s_2}, P_{s_3})$.

In the new problem under $\mu$ student $s_1$ justifiably envies student $s_2$ at school $c_3$ through the feasible assignment
\[
\mu' = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_3 & s_1 \end{pmatrix}
\]
since $\mu(s_1) = c_1, c_3 P'_{s_1} c_1$ and $s_1 P_{s_3} s_2$. Now it is straightforward to verify that the unique feasible and fair across types assignment of the new problem is $\bar{\mu}$. Thus any mechanism, which is both feasible and fair across types, must select the assignment $\bar{\mu}$ for the new problem. Under $\bar{\mu}$ student $s_1$ is assigned school $c_2$ which is strictly preferred to $c_1$ under the true preference $P_{s_1}$. Thus student $s_1$ is better off by stating $P'_{s_1}$ than by stating his true preference $P_{s_1}$, and the mechanism is not incentive compatible.

**Case 2:** The mechanism selects $\bar{\mu}$.

Under $\bar{\mu}$ student $s_2$ is assigned school $c_1$. Similarly as in Case 1 we will show that student $s_2$ gains by misreporting his preference. Suppose that student $s_2$ states the (false) preference $P'_s$ given by $c_3 P'_{s_2} c_2 P'_{s_2} c_1 P'_{s_2} s_2$, and all other students were to state their true preferences. Keeping all other components of the above problem fixed, in the new problem the students' preferences are $P'_S = (P_{s_1}, P'_{s_2}, P_{s_3})$. 

40
In the new problem under $\bar{\mu}$ student $s_2$ justifiably envies student $s_1$ at school $c_2$ through the feasible assignment

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

since $\bar{\mu}(s_2) = c_1, c_2 P_{s_2} c_1$ and $s_2 P_{c_2} s_1$. Now it is straightforward to verify that $\mu$ is the unique feasible assignment which is fair across types for the new problem. Thus any mechanism, which is both feasible and fair across types, must select the assignment $\mu$ for the new problem. Under $\mu$ student $s_2$ is assigned school $c_3$ which is strictly preferred to $c_1$ under the true preference $P_{s_2}$. Thus student $s_2$ does better by stating $P'_{s_2}$ than by stating his true preference $P_{s_2}$, and the mechanism is not incentive compatible. \hfill \Box

Remark 8 Similar to Remark 6, in controlled school choice there is not always a unique candidate for a feasible assignment which is fair across types. This provides again additional reason for Theorems 5’.

In the example used to prove Theorem 5’ let the controlled school choice problem be given by Table 4 except for school $c_1$’s capacity constraints and ceilings: let $q_{c_1} = 2$ and $\bar{q}_{t_{c_1}}^1 = \bar{q}_{t_{c_1}}^2 = 2$. Then it is straightforward to verify that

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

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

are assignments which are both feasible and fair across types for this problem. Since student $s_1$ prefers $c_2$ to $c_1$ under $P_{s_1}$ and student $s_2$ prefers $c_3$ to $c_1$ under $P_{s_2}$, students’ preferences are opposed over $\mu$ and $\bar{\mu}$. Obviously, there is no feasible and fair assignment which students $s_1$ and $s_2$ prefer to $\mu$ and $\bar{\mu}$: if there were such an assignment, then neither $s_1$ nor $s_2$ is assigned $c_1$ and constraints at school $c_1$ are violated since the floor for type $t_1$ at school $c_1$ is equal to one.

When computing the “minimum” $\wedge$ of $\mu$ and $\bar{\mu}$ (by assigning each student to the school which he least prefers from $\mu$ and $\bar{\mu}$) we obtain the assignment

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

which is feasible but not fair across types since student $s_2$ justifiably envies student $s_3$ at school $c_3$. 

41
References


