THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

# Discussion Paper Series
## UQ School of Economics

Paper Name: An Efficient, Computationally Tractable School Choice Mechanism
Date: February 2024

Andrew McLennan, Shino Takayama and Yuki Tamura

Discussion
Paper No. 668

# An Efficient, Computationally Tractable School Choice Mechanism

Andrew McLennan* Shino Takayama† Yuki Tamura‡

February 9, 2024

**Abstract**

We show that the application of the Generalized Constrained Probabilistic Serial mechanism of Balbuzanov (2022) (which generalizes the Probabilistic Serial mechanism of Bogomolnaia and Moulin (2001)) to school choice has attractive properties. The mechanism is intuitively simple, assigning to each student, at each moment in the unit interval, probability of receiving a seat in her favorite school among those that are available then. It is $sd$-efficient and effectively strategy proof. We provide an algorithm, based on a generalization of Hall's marriage theorem, for computing the mechanism, which has been implemented, and seems likely to have reasonable running times even for the world's largest school choice problems.

**Keywords:** School Choice, Object Allocation, Efficiency, Fairness, Strategy Proofness, Probabilistic Serial Mechanism, Hall's Marriage Theorem.

In a seminal paper Abdulkadiroğlu and Sönmez (2003) propose the application, to school choice, of two mechanisms based on matching theory. The *student proposes deferred acceptance* (DA) mechanism was originally proposed by Gale and Shapley (1962), and it has been widely adopted for school choice and similar problems around the world. The *top trading cycles* (TTC) mechanism was originated by David Gale, as described by Shapley and Scarf (1974), and although it has some superior theoretical properties, it has found less practical acceptance. This paper argues that the *generalized constrained probabilistic serial* (GCPS) mechanism of Balbuzanov (2022), which is a generalization of the probabilistic serial (PS) mechanism of Bogomolnaia and Moulin (2001) (henceforth BM) is also viable and attractive as a school choice mechanism.

*School of Economics, University of Queensland, a.mclennan@economics.uq.edu.au

†School of Economics, University of Queensland, s.takayama1@uq.edu.au

‡Center for Behavioral Institutional Design, NYU Abu Dhabi; yuki.tamura@nyu.edu

In practice (e.g. Pathak (2017)) transparency and straightforward incentives are required in order for school choice mechanisms to be accepted by parents. The *Boston* (or *immediate acceptance*) mechanism begins by assigning as many students to their favorite (according to the submitted rankings) schools as possible. It then assigns as many of the remaining students to their second favorite schools as possible, and it continues similarly, in the $k^{\text{th}}$ round assigning as many of the remaining students to their $k^{\text{th}}$ favorite schools as possible. Since it is possible that a student can (for example) greatly increase her chance of being accepted at her second favorite school if she ranks it as her favorite, the Boston mechanism is not strategy proof, and in fact it is strategically tricky, with high stakes. Nevertheless it continues to be widely used because it is conceptually simple.

The GCPS mechanism has a similar conceptual simplicity: during the unit interval of time, each student is assigned probability of receiving a seat in her favorite school, among those she is eligible for, at unit speed, until some capacity constraint is encountered, at which point each student who can no longer be assigned probability in her favorite school switches to receiving probability in the favorite school among those that still have some available capacity. The process continues similarly, with each student at each time receiving probability of receiving a seat in her favorite school among those that are still available, so that at time $1$ each student has a probability distribution over schools. As we explain in Online Appendix A, it is possible to compute a random deterministic assignment with a distribution that realizes these probabilities.

Strictly speaking, the GCPS mechanism is not strategy proof, but we will argue that it is effectively strategy proof in an easily understood sense. In order to manipulate by submitting a false ranking, there must first be a period of time during which the student is receiving probability of a school that is worse than the one she would have been consuming if she had told the truth, followed by a period of time during which she is receiving probability of a school that is better than the one she would have been consuming, but if many students are competing for desirable schools, the extension of the latter period due to the manipulation is necessarily brief.

Both DA[1] and TTC[2] require that the schools have strict rankings of the students

[1]At the outset in DA each student applies to her favorite school. Each school with more applicants than its capacity rejects the lowest priority applicants beyond the number it can serve. In each subsequent round each student who was rejected in the preceding round applies to her favorite school among those that have not rejected her, and each school retains the highest priority applicants, up to its capacity, among those who have applied in all rounds, and rejects all others. The process continues in the same manner until there is a round with no rejections.

[2]In TTC each student points to her favorite school and each school points to its highest priority student. The resulting directed graph has at least one cycle, each student in a cycle is assigned to the school she points to, and she is removed from the mechanism, along with the seat she claimed in her school. This process is then repeated with the remaining students and seats, and it continues in this

that are called *priorities*. The assignments produced by DA are ex post efficient[3] if the schools' priorities actually reflect society's values. However, if the priorities are generated randomly, simply in order to fulfill the requirements of the mechanism, then assignments can be inefficient[4]. In a study of New York City data Abdulkadiroğlu et al. (2009) found that the inefficiencies arising in this way are quantitatively significant.

The outcomes produced by TTC are ex post efficient relative to the students' preferences. Specifically, the students leaving the mechanism in the first round are receiving their favorite schools, the students leaving the mechanism in the second round are receiving their favorite schools among those that remain, so their assignments cannot be improved without disturbing the assignments from the first round, and so forth.

The assignment probabilities produced by GCPS are $sd$-*efficient*: there are no other assignment probabilities that give each student a probability distribution over schools that first order stochastically dominates the one given by the GCPS assignment, with strict domination for some students. Consequently any probability distribution over deterministic assignments that realizes the GCPS probabilities assigns positive probability only to ex post efficient assignments. An important observation of BM is that, in the context of object allocation, random priority[5] may produce assignment probabilities that are not $sd$-efficient, even though it produces ex post efficient assignments.

In practice almost all school choice mechanisms limit the number of schools that a student can rank, and in very large districts such restrictions seem unavoidable. We focus on a version of the GCPS mechanism that finishes in a single round (other possibilities are described below) and does not assign any student to a school she did not rank. Specifically, we assume that each student is assigned a *safe school* which is guaranteed to accept the student if she is not admitted to a school she prefers to it, and which she may be required to attend if other schools do not admit her.

We assume that for each school, the number of students for whom that school is the safe school is not greater than the school's capacity. The GCPS mechanism requires that there is a probabilistic assignment that assigns all students to schools they are eligible for and that does not exceed any school's capacity, and our assumption insures that the assignment of each student to her safe school is such an assignment. Other

manner until all students have been assigned. When different schools have different priorities, the role of priorities in TTC is hard to grasp. (See Leshno and Lo (2020).)

[3]More precisely, DA produces assignments that are *stable*: there is no student-school pair such that the student prefers that school to the one she has been matched with and the student has a higher priority at that school than some other student that has been assigned to that school.

[4]In the simplest instance, if Bob prefers Carol School to Alice School, while Ted prefers Alice School, the mechanism may nevertheless assign Bob to Alice School and Ted to Carol School if the schools' priorities "prefer" that outcome.

[5]In random priority for object allocation the agents are ordered randomly, the first agent chooses her favorite object, the second agent chooses her favorite of the remaining objects, and so forth.

methods, such as requiring each student to rank the three closest schools, may be able to insure the existence of a feasible allocation while also providing a lower bound on the student's outcome, and an upper bound on what the student can insist on by not ranking other alternatives. Our focus on safe schools is primarily for the sake of simplicity and concreteness.

Mechanisms that would be strategy proof without restrictions on the number of schools that can be ranked become manipulable when such restrictions are imposed. Haeringer and Klijn (2009) study the Nash equilibria of matching based mechanisms with such limitations. Calsamiglia et al. (2010) is an experimental study of the effects of constraining the number of schools that can be ranked, for DA and TTC; a main finding is that constraints have a large negative effect on manipulability, and reduce efficiency and stability while increasing segregation.

In the GCPS mechanism with safe schools each student submits only a ranking of those schools she weakly prefers to her safe school. Safe schools are also possible with DA: if each student has the highest possible priority at her safe school, and each school has enough capacity for all the students for which it is the safe school, then a student will never be rejected by her safe school, and only needs to rank schools she weakly prefers to it. Both for GCPS and DA, strategy proofness is largely restored by having safe schools if most students prefer at most a small number of schools to the safe school. Of course having safe schools that students are likely to find desirable is consistent with the main goal of school choice, which is to assign students to schools they would like to attend. Some systems (e.g., the state of Victoria in Australia) have *neighborhood priority* in which each student's safe school is the one whose district contains her residence.

In the New York City High School Match as of 2006 (Pathak, 2006) each student submitted a ranking of up to 12 schools. Of the roughly 100,000 participants, over 8,000 were unmatched after the main round, in the sense that they were not offered a seat by any school they ranked. These students submitted new rank ordered lists for the supplementary round, in which schools with unfilled capacity participated. Students who did not receive a seat in the supplementary round were assigned administratively. We do not know the particular considerations that motivated this design. (One possibility is that neighborhood priority would have impeded a goal of school desegregation since there was a high degree of de facto residential segregation.)

The important point for us is that GCPS can also be employed in multiround systems: each student's safe school in the first round is participation in the second round, for each student in the second round the safe school is participation in the third round or administrative assignment, and so forth. In the remainder we assume a single round because this setting is simple, but rich enough to encompass the relevant technical issues.

One of the advantages of the GCPS mechanism is that the schools' priorities need not be strict. At one extreme, the GCPS mechanism makes it possible to have *dichotomous priorities*: each school gives equal consideration to all students who are qualified, perhaps by virtue of gender for single sex schools or test score cutoffs for selective schools. Welfare analysis is simplified and clarified because one may consider only the priorities that express societal values, whereas DA and TTC may mix such priorities with arbitrary tie breaking.

However, the priorities in school choice mechanisms often express societal values. For example, in China (Wang and Zhou, 2020) the student's score on a standardized exam is taken to be her priority, presumably reflecting a policy objective of providing the most highly demanded resources, and the widest range of options, to the most talented students. Priorities may be affected by gender, minority status, and residential location. In effect, DA computes a vector of priority cutoffs for the schools, and each student is matched with the most preferred school among those whose threshhold she exceeds. (This is also the case for TTC if all schools have the same priorities.)

GCPS can be used to attain similar outcomes. When each school has finitely many priority classes, an ideal outcome for any particular school is that it either has excess capacity, and accepts all students who apply, or it has a cutoff priority class, students with lower priority are not admitted, students with higher priority have no probability of being required to attend a school they like less, and the number of seats assigned to students in the cutoff class allows the school's capacity to be exactly utilized. In Section 5 we show that there is a setting of the parameters of the GCPS mechanism that attains this ideal simultaneously at all schools.

Two technical innovations underlie the computational feasibility of the GCPS mechanism. To facilitate the discussion we quickly review some basic results (without proofs) and terminology.

A *polytope* $Q$ may be defined to be the convex hull of a finite set of points, or as an intersection of finitely many closed half spaces that happens to be bounded. To avoid technical detail our discussion in this paragraph assumes that $Q$ is full dimensional, in the sense that its affine hull is the entire Euclidean space of which it is a subset. Among the finite systems of weak linear inequalities that may be used to define $Q$, there is a unique (up to rescaling of inequalities by multiplication by positive scalars) such system that is minimal, and that is contained in any other such system. Its elements are the *facet inequalities* of $Q$. For each facet inequality the corresponding *facet* is the subset of $Q$ on which the facet inequality holds with equality. A subset of $Q$ is a *face* if it is $Q$ itself, the null set, or the intersection of some set of facets. A polytope $Q$ is the convex hull of a finite set of points, and among the finite sets whose convex hulls are $Q$, there is a unique such set that is minimal in the sense that it is contained in any other such set, whose elements are the *vertices* of $Q$. The vertices of $Q$ may also be described as its extreme points, where an *extreme point* of $Q$ is a point that cannot be expressed as a convex combination of other points of $Q$.

In Balbuzanov (2022) the set of feasible allocations is a given polytope $Q$ in the nonnegative orthant of the space of matrices of assignment probabilities. (Echenique et al. (2021) follow this approach in their study of pseudo-market equilibria with constraints.) Let $R$ be the intersection of the nonnegative orthant with the sum of $Q$ and the nonpositive orthant. That is, a point in the nonnegative orthant is in $R$ if and only if it lies below some point of $Q$.

The GCPS allocation process is a piecewise linear function $p\colon [0,1] \to R$. It begins with $p(0)$ equal to the origin and increases each student's probability of receiving her favorite school, among those she is allowed to consume, until one of the facet inequalities of $R$ is encountered. A key result (Balbuzanov's Proposition 1) is that the facet inequalities of $R$ (other than the nonnegativity conditions) require that weighted sums of probabilities, with nonnegative weights, not exceed certain quantities. When the process encounters one or more facet inequalities, each student's set of allowed objects is updated by disallowing further consumption of probabilities that would result in one of these facet inequalities being violated. The process then continues, with each student increasing the probability of receiving her favorite allowed school until additional facet inequalities of $R$ are encountered, and again the students' sets of allowed schools are updated. (For the problems we study each student's set of allowed schools is always nonempty.) Eventually the process arrives at a point $p(1) \in Q$ that is, by definition, the GCPS allocation.

A computational implementation of the GCPS mechanism must have a way of detecting when the allocation process encounters a facet of $R$. Our first main result is a generalization of Hall's marriage theorem that gives a set of inequalities, in closed form, that contains the facet inequalities of $R$. Our second main innovation is a fast algorithm for computing the GCPS allocation. In addition to computing $p$, it also computes a piecewise linear path $\overline{p}\colon [0,1] \to Q$ such that $p(t) \leq \overline{p}(t)$ for all $t$. Having chosen a trajectory for $\overline{p}$, $p$ continues along the trajectory given by the students' favorite allowed schools, and $\overline{p}$ continues along the chosen trajectory, until the time at which some resource constraint is encountered by $p$, or continuing further would result either in $\overline{p}$ leaving $Q$ or $p$ no longer lying below $\overline{p}$. At that time a combinatoric calculation of bounded complexity gives either a new trajectory for $\overline{p}$ or a facet inequality of $R$ that is satisfied by $p$ and $\overline{p}$ at that time.

We briefly describe the structure of the remainder. The next section reviews related literature. Section 2 states and proves our generalization of Hall's marriage theorem. During the allocation process there can be a *critical pair* consisting of a set $J$ of agents and a set $P$ of objects such that the agents in $J$ must be assigned all of the remaining capacity of the objects in $P$. Section 3 studies such pairs. Section 4 describes the algorithm for computing the GCPS allocation. Section 5 states the aforementioned result concerning simultaneous satisfaction of priorities at all schools, and describes an iterative adjustment procedure in which the GCPS allocation is computed repeatedly, hoping to approximate this outcome.

Section 6 shows that GCPS allocations are $sd$-efficient, and also efficient in relation to other orderings of the set of probability measures on objects derived from an ordinal preference that correspond to the limits of extreme risk loving and extreme risk averse cardinal preferences. Section 7 considers the fairness properties of GCPS allocations. Section 8 argues that although the GCPS mechanism is not fully strategy proof, it is very difficult to manipulate in its application to school choice, and we present two theoretical results in this direction. Section 9 provides some concluding remarks.

Online Appendix A describes a special case of an algorithm of Budish et al. (2013) that passes from a matrix of assignment probabilities to a random deterministic assignment whose distribution realizes the given probabilities. Online Appendix B gives a brief informal description of the software package *GCPS Schools*, which implements the algorithms described in Section 4 and Online Appendix A. *GCPS Schools* also contains an application that generates sample school choice problems. Online Appendix D contains the proofs of the results of Section 8. Other proofs are in Online Appendix C.

# 1 Background and Related Literature

The literature on school choice is now vast; Abdulkadiroğlu and Andersson (2022) is a recent survey. In this section we survey some of the literature that is most closely related our work.

In response to the inefficiencies observed by Abdulkadiroğlu et al. (2009), a rather extensive literature (Erdil and Ergin, 2008, Kesten, 2010, Tang and Yu, 2014, Kesten and Ünver, 2015) studies how the outcome of DA might be adjusted ex post, possibly by substituting new priorities. In fact there are theoretical barriers to improving efficiency by manipulating the breaking of ties in the schools' rankings. Gale and Shapley (1962) show that DA yields the best outcome for each student that can be achieved in any allocation without justified envy for the given priorities. Improving on results of Kesten (2006) and Erdil and Ergin (2008), Theorem 1 of Abdulkadiroğlu et al. (2009) asserts that, for any member of a large class of tie breaking rules, there is no mechanism that is both strategy proof for that tie breaking rule and gives outcomes that weakly Pareto dominate those produced by DA.

Although it is not widely used, TTC continues to be a topic of research. Some papers (Morrill, 2015, Hakimov and Kesten, 2018, Grigoryan, 2023) proposed modified versions of the mechanism. Leshno and Lo (2020) analyze it in terms of cutoffs for the schools.

The GCPS mechanism may be applied to domains other than school choice. For example, motivated by matching of medical residents with hospitals in Japan and similar problems, Kamada and Kojima (2015, 2017) study mechanisms in which regional caps on the number of residencies are implemented by imposing caps on the number of residencies at individual hospitals in the region. This can lead to a hospital rejecting applicants as a result of the hospital's cap even though other hospitals in the region have unfilled vacancies. They propose a more flexible version of DA in which some hospitals are allowed to exceed their caps if the total number of doctors matched to the region is below the region's cap. Similar effects can be achieved by running the GCPS mechanism repeatedly while adjusting the caps of individual hospitals.

One way to implement affirmative action objectives has been suggested by Abdulkadiroğlu and Sönmez (2003). For example, a school may be divided into three subschools, one with 30% of the seats that is reserved for minority students, one with 30% of the seats that is reserved for majority students, and one with 40% of the seats that accepts all students. "Hard" upper and lower bounds for the percentages of students of different types are extensively used in practice, but Kojima (2012) and Hafalir et al. (2013) point out that they lead to conflicts with other objectives, and Ehlers et al. (2014) suggest implementing affirmative action goals using soft bounds. Such an approach can be implemented, at least informally, by running the GCPS algorithm multiple times while adjusting the parameters to better reconcile competing objectives.

We now describe the PS mechanism of BM and subsequent generalizations. BM study the problem of assigning a different object from a finite set to each of finitely many agents, based on their reported strict ordinal preferences. BM provide an intuitive description of the PS mechanism in which each object is regarded as a perfectly divisible cake of unit size. At each moment in the unit interval of time each agent consumes, at unit speed, probability of her favorite cake, among those that have not yet been fully consumed. Provided that there are at least as many objects as agents, at time 1 each agent has a probability distribution over the objects, and for each object the sum of the assignment probabilities is not greater than one. Among the most important theoretical results are that the PS mechanism is $sd$-efficient but not strategy proof.

Extensions of BM's cake eating procedure have been proposed in (at least) six other papers. Using the method of network flows (see Section 2) Katta and Sethuraman (2006) extend the PS mechanism to profiles of preferences with indifferences. Their mechanism has both the PS mechanism for strict preferences and the mechanism proposed by Bogomolnaia and Moulin (2004) for matching problems with dichotomous preferences as special cases. Bogomolnaia (2015) provides a welfarist characterization of it.

Kojima (2009) studies perhaps the simplest extension of BM in which agents receive multiple objects. Each agent receives $r \geq 2$ objects, and the number of objects is $r$ times the number of agents. The mechanism is shown to be $sd$-efficient and envy-free, but not weakly strategy proof, as we explain in more detail in Section 8.

Yilmaz (2010) studies house allocation problems with existing tenants, which are object allocation problems in which some objects have owners who can insist on not receiving a worse object. He proposes the special case of the mechanism studied here for that problem, and in particular he recognizes the relationship between Hall's marriage theorem, its generalization by Gale (1957), and the set of feasible allocations. His algorithm is generalized by Athanassoglou and Sethuraman (2011) to problems in which agents have fractional endowments. Yilmaz (2009) uses the methods of Katta and Sethuraman (2006) to extend the mechanism to the domain of preferences with indifferences.

Budish et al. (2013) study problems in which there are constraints that require that certain sums of probabilities are bounded, either below, in which case the constraint is a *floor constraint*, or above, in which case it is a *ceiling constraint*. For a problem with only ceiling constraints in which there is a "null object" (e.g., being unemployed, unhoused, or unschooled) that is available in infinite supply, and which is not involved in any constraint, they propose a *generalized probabilistic serial* (GPS) mechanism. As in BM, at each moment in $[0, 1]$ each agent increases her probability of her favorite available object. When a ceiling constraint binds with equality, the sets of available objects are revised by disallowing further consumption of probabilities that would violate a constraint. Since the null object is always available, each agent's set of available objects is always nonempty. Thus at time 1 each agent has total probability one, and the GPS assignment is defined as the probability shares that have been consumed by each agent at time 1.

As we have already described, Balbuzanov (2022) generalizes the Budish et al. (2013) mechanism by allowing the set of feasible allocations to be an arbitrary polytope $Q$ in the nonnegative orthant of the space of matrices of assignment probabilities.

## 2 A Generalized Hall's Marriage Theorem

In this section we introduce the formal framework, state and prove the generalization of Hall's theorem, and provide useful characterizations of $Q$ and $R$.

A *communal endowment economy* (CEE) is a quintuple $E = (I, O, r, q, g)$ in which $I$ is a nonempty finite set of *agents*, $O$ is a nonempty finite set of *objects*, $r \in \mathbb{R}^I_+$, $q \in \mathbb{R}^O_+$, and $g \in \mathbb{R}^{I \times O}_+$. We say that $r_i$ is $i$'s *requirement*, $q_o$ is the *quota* of $o$, and $g_{io}$ is *$i$'s $o$-max*. We say that $E$ is *integral* if $r \in \mathbb{Z}^I_+$, $q \in \mathbb{Z}^O_+$, and $g \in \mathbb{Z}^{I \times O}_+$. In comparison with most models of random assignment, the matrix $g$ is the main novelty, and we will see that it may represent several things and be used in various ways.

Several types of CEE occur in our discussion. A *Gale supply-demand CEE* is a CEE $E$ such that $g_{io} \in \{0, r_i\}$ for all $i \in I$ and $o \in O$. We say that $E$ is a *school choice CEE* if $r_i = 1$ for all $i$, and we write $E = (I, O, 1, q, g)$ to indicate that this is the case. In a school choice CEE elements of $I$ are *students*, elements of $O$ are *schools*, and each student must receive a seat in some school. In an integral school choice CEE each school has an integral number of seats, and for each student $i$ and school $o$, $g_{io} = 1$ if $i$ is eligible to attend $o$ and weakly prefers it to her safe school, and otherwise $g_{io} = 0$. (In Section 5 we will describe a method that uses fractional $g_{io}$ to adjust the GCPS allocation when the schools have nondichotomous priorities.)

A *Hall marriage problem* is a CEE such that for all $i$ and $o$, $r_i = 1$, $q_o = 1$, and $g_{io} \in \{0, 1\}$. In this case elements of $I$ are *boys* and elements of $O$ are *girls*. Intuitively a Hall marriage problem is a bipartite graph with an edge connecting boy $i$ to girl $o$ if $i$ and $o$ are compatible.

An *allocation* for $I$ and $O$ is a matrix $p \in \mathbb{R}^{I \times O}_+$. Such a $p$ is *integral* if $p \in \mathbb{Z}^{I \times O}_+$. In general, for any matrix $p \in \mathbb{R}^{I \times O}$ and any $i \in I$ and $o \in O$, $p_i = (p_{io})_{o \in O}$ and $p_o = (p_{io})_{i \in I}$ are the corresponding row and column. A *partial allocation* for $E$ is an allocation $p$ such that $\sum_o p_o \le r$, $\sum_i p_i \le q$, and $p_{io} \le g_{io}$ for all $i$ and $o$. A *feasible allocation* is a partial allocation $m$ such that $\sum_o m_o = r$. A partial allocation $p$ is *possible* if there is a feasible allocation $m$ such that $p \le m$. Let $Q$ be the set of feasible allocations, and let $R$ be the set of possible partial allocations.

For $J \subset I$ and $P \subset O$ let $J^c = I \setminus J$ and $P^c = O \setminus P$ be the complements. We say that $E$ satisfies the *generalized marriage condition* (GMC) if, for every $J \subset I$ and $P \subset O$,

$$\sum_{i \in J} r_i \le \sum_{i \in J} \sum_{o \in P^c} g_{io} + \sum_{o \in P} q_o.$$

We will refer to this relation as the *GMC inequality* for $(J, P)$. Note that the GMC inequality for $(\{i\}, \emptyset)$ is $r_i \le \sum_o g_{io}$, and the GMC inequality for $(I, O)$ is $\sum_i r_i \le$

$\sum_o q_o$. The GMC is obviously necessary for the existence of a feasible allocation. Our first main result is:

**Theorem 1.** The CEE $E$ has a feasible allocation if and only if it satisfies the GMC.

The Gale (1957) supply-demand theorem[6] is the special case of this for a Gale supply-demand CEE.

If $E$ is a Hall marriage problem, the set of *neighbors* of boy $i$ is $N_g(i) = \{\, o \in O : g_{io} = 1 \,\}$, and for $J \subset I$ we set $N_g(J) = \bigcup_{i \in J} N_g(i)$. We say that $E$ satisfies the *marriage condition* if $|J| \leq |N_g(J)|$ for all $J \subset I$. The GMC inequality for $J$ and $P = N_g(J)$ gives this inequality. Conversely, for a given $J \subset I$, the contribution of $o \in N_g(J)$ to the right hand side of the GMC inequality is minimized if $o \in P$, and the contribution of $o \in N_g(J)^c$ is minimized if $o \in P^c$, so $|J| \leq |N_g(J)|$ for all $J$ implies that the GMC is satisfied. Therefore Theorem 1 implies that $E$ has a feasible allocation if and only if the marriage condition is satisfied.

For a Hall marriage problem an integral feasible allocation is called a *matching*. (Each of the boys has a different partner.) Hall's marriage theorem asserts that a Hall marriage problem has a matching if and only if it satisfies the marriage condition. To pass from a feasible allocation to a matching one can repeatedly adjust the allocation along paths of fractional allocations that alternate between boys and girls, and either form a loop or pass from one incompletely allocated girl to another. A more precise and general version of this argument is given in Online Appendix A.

For $i \in I$ let
$$\alpha_i = \{\, o \in O : g_{io} > 0 \,\}$$
be the set of objects that are *possible* for $i$, and for $P \subset O$ let $J_P = \{\, i \in I : \alpha_i \subset P \,\}$ be the set of agents who cannot be allocated objects outside of $P$. If $E$ is an integral school choice CEE, then for any $P \subset O$, $J_P$ minimizes the difference between the right hand side and the left hand side of the GMC inequality. Therefore $E$ satisfies the GMC if and only if, for each $P \subset O$,
$$|J_P| \leq \sum_{o \in P} q_o.$$

Our proof of Theorem 1 is a simple application of the method of network flows. (Ahuja et al. (1993) provides a general introduction and overview.) Let $(N, A)$ be a directed graph ($N$ is a finite set of *nodes* and $A \subset N \times N$ is a set of *arcs*) with distinct distinguished nodes $s$ and $t$, called the *source* and *sink* respectively. For the sake of

[6] Although this result is attributed to Gale (1957) by Yilmaz (2010), and perhaps others, this exact formulation does not appear in Gale's paper. The paper does consider slightly more complicated problems, and it is easy to see that this result can be obtained from Gale's methods in the same manner.

simplicity and clarity of intuition (the formal analysis can be more general) we assume that $(n, s), (t, n), (n, n) \notin A$ for all $n \in N$, and that $(n', n) \notin A$ whenever $(n, n') \in A$.

A *flow* is a function $f \colon N \times N \to \mathbb{R}$ such that:

(a) for all $n$ and $n'$, if $(n, n') \notin A$, then $f(n, n') \leq 0$ ;

(b) for all $n$ and $n'$, $f(n, n') = -f(n', n)$;

(c) $\sum_{n' \in N} f(n', n) = 0$ for all $n \in N \setminus \{s, t\}$.

If neither $(n, n')$ nor $(n', n)$ is in $A$, then (a) and (b) imply that $f(n, n') = 0$. In conjunction with the other requirements, (c) can be understood as saying that for each $n$ other than $s$ and $t$, the total flow into $n$ is equal to the total flow out. Note that (a) and (b) imply that $f(s, n), f(n, t) \geq 0$ for all $n \in N$. Applying (b), then (c), gives

$$0 = \sum_{n \in N} \sum_{n' \in N} f(n', n) = \sum_{n' \in N} f(n', s) + \sum_{n' \in N} f(n', t),$$

so we may define *value* of $f$ to be

$$|f| = \sum_{n \in N} f(s, n) = \sum_{n \in N} f(n, t).$$

A *capacity* is a function $c \colon N \times N \to [0, \infty]$ such that $c(n, n') = 0$ whenever $(n, n') \notin A$. A flow $f$ is *bounded* by a capacity $c$ if $f(n, n') \leq c(n, n')$ for all $(n, n')$. A *maximum flow* for $c$ is a flow $f$ that is maximal for $|f|$ among the flows bounded by $c$.

A *cut* is a set $S \subset N$ such that $s \in S$ and $t \in S^c$ where $S^c = N \setminus S$ is the complement. For a capacity $c$, the *capacity* of $S$ is

$$c(S) = \sum_{(n,n') \in S \times S^c} c(n, n').$$

The value $|f|$ is the net flow from $S$ to $S^c$, hence the total flow from $S$ to $S^c$ minus the total flow from $S^c$ to $S$, so $|f| \leq c(S)$ when $f$ is bounded by $c$, and thus the maximum value of flows bounded by $c$ is not greater than the minimum capacity of a cut for $c$. The max-flow min-cut theorem (Ford and Fulkerson, 1956) asserts that these two quantities are equal.

Let $f$ be a flow bounded by $c$, and let $S$ be a cut. If $\sum_{(n,n') \in S \times S^c} f(n, n') = c(S)$, then $f$ is a maximal flow for $c$, $S$ is minimum capacity cut for $c$, and since $f \leq c$ we have $f(n, n') = c(n, n')$ for all $(n, n') \in S \times S^c$. Conversely, if $f$ is a maximal flow for $c$ and $S$ is minimum capacity cut for $c$, then $c(S) = |f| = \sum_{(n,n') \in S \times S^c} f(n, n')$. It is well known (Ford and Fulkerson, 1956, Shapley, 1961, Ore, 1962) that the set of minimal cuts is a lattice in the sense that if $S_1$ and $S_2$ are minimal cuts, then so are $S_1 \cup S_2$ and $S_1 \cap S_2$[7].

For the given CEE $E$ we define a particular directed graph $(N_E, A_E)$ in which $N_E = \{s\} \cup I \cup O \cup \{t\}$ and

$$A_E = \{\, a_i : i \in I \,\} \cup \{\, a_{io} : i \in I, o \in O \,\} \cup \{\, a_o : o \in O \,\}$$

where, for $i \in I$ and $o \in O$, $a_i = (s, i)$, $a_{io} = (i, o)$, and $a_o = (o, t)$. Let $c_E$ be the capacity in which $c_E(a_i) = r_i$, $c_E(a_{io}) = g_{io}$, and $c_E(a_o) = q_o$. If $p$ is an allocation, there is a unique flow $f_p$ such that $f_p(a_{io}) = p_{io}$ for all $i$ and $o$ that has $f_p(a_i) = \sum_o p_{io}$ for all $i$ and $f_p(a_o) = \sum_i p_{io}$ for all $o$. Evidently $p$ is a partial allocation if and only if $f_p$ is bounded by $c_E$, and it is a feasible allocation if and only if, in addition, $f_p(a_i) = r_i$ for all $i$, which is the case if and only if $|f_p| = \sum_i r_i$. Conversely, if $f$ is a flow bounded by $c_E$ with $|f| = \sum_i r_i$ and thus $f(a_i) = r_i$ for all $i$, then setting $m_{io} = f(a_{io})$ gives a feasible allocation $m$.

Thus there is a feasible allocation if and only if the maximum value of a flow bounded by $c_E$ is $\sum_i r_i$. The max flow-min cut theorem implies that this is the case if and only if the minimum capacity of a cut for $c_E$ is $\sum_i r_i$. Since $c_E(\{s\}) = \sum_i r_i$, there is a feasible allocation if and only $c_E(S) \geq \sum_i r_i$ for all cuts $S$.

For $J \subset I$ and $P \subset O$ let $S_{(J,P)} = \{s\} \cup J \cup P$. This is a cut, and if $S$ is a cut, then $S = S_{(J,P)}$ where $J = S \cap I$ and $P = S \cap O$. An arc can go from a node in $S_{(J,P)}$ to a node in $S^c_{(J,P)}$ by going from $s$ to a node in $J^c$, by going from a node in $J$ to a node in $P^c$, and by going from a node in $P$ to $t$, so

$$c_E(S_{(J,P)}) = \sum_{i \in J^c} r_i + \sum_{i \in J} \sum_{o \in P^c} g_{io} + \sum_{o \in P} q_o.$$

Thus there is a feasible allocation if and only if $\sum_i r_i \leq c_E(S_{(J,P)})$ for all $J \subset I$ and $P \subset O$, and subtracting $\sum_{i \in J^c} r_i$ from both sides reveals that this inequality is equivalent to the GMC inequality for $J$ and $P$.

Hall's marriage theorem, the Gale supply-demand theorem, and the max-flow min-cut theorem are three members of a large and important class of results in combinatorial matching theory that are equivalent in the informal sense that relatively simple arguments (described in detail by Reichmeider (1978, 1985)) allow one to pass from any member of the class to any other. As yet another member of this class, Theorem 1 does not provide distinctly novel mathematical information. Its primary significance here,

[7] If $n \in S_1 \cup S_2$ and $n' \in (S_1 \cup S_2)^c = S_1^c \cap S_2^c$ (or $n \in S_1 \cap S_2$ and $n' \in S_1^c \cup S_2^c$) then either $n \in S_1$ and $n' \in S_1^c$ or $n \in S_2$ and $n' \in S_2^c$, and in either case $f(n, n') = c(n, n')$.

and perhaps more generally, is that the test it provides is in closed form.

The next result gives a finite collection of inequalities, in closed form, that contains the facet inequalities of $R$. If $p$ is a partial allocation, let

$$E - p = (I, O, r', q', g')$$

be the derived CEE in which $r'_i = r_i - \sum_o p_{io}$, $q'_o = q_o - \sum_i p_{io}$, and $g'_{io} = g_{io} - p_{io}$. If $p$ is a partial allocation, $m$ is an allocation, and $p \leq m$, then $m$ is a feasible allocation for $E$ if and only if $m - p$ is a feasible allocation for $E - p$. Thus a partial allocation $p$ is possible if and only if $E - p$ has a feasible allocation, which of course is the case if and only if $E - p$ satisfies the GMC. Substituting the definitions above into the GMC inequality for $E - p$ and $(J, P)$, then simplifying, gives

$$\sum_{i \in J^c} \sum_{o \in P} p_{io} \leq \sum_{o \in P} q_o + \sum_{i \in J} \sum_{o \in P^c} g_{io} - \sum_{i \in J} r_i. \tag{1}$$

**Proposition 1.** *$R$ is the set of partial allocations $p$ such that $p_{io} \leq g_{io}$ and $\sum_o p_{io} \leq r_i$ for all $i$ and $o$ and (1) holds for all $J \subset I$ and $P \subset O$.*

# 3 Critical Pairs

In this section we work with a given CEE $E$ that satisfies the GMC. For $J \subset I$ and $P \subset O$ we say that the pair $(J, P)$ is *critical* for $E$ if $(J, P) \neq (\emptyset, \emptyset)$ and it satisfies the GMC inequality for $(J, P)$ with equality:

$$\sum_{i \in J} r_i = \sum_{i \in J} \sum_{o \in P^c} g_{io} + \sum_{o \in P} q_o.$$

We refer to this condition as the *GMC equality* for $(J, P)$. Our goal in this section is to understand the relationship between critical pairs and feasible allocations, and how the various critical pairs for $E$ are related to each other.

Evidently, if $(J, P)$ is critical for $E$, then any feasible allocation $m$ gives the agents in $J$ all of the endowment of objects in $P$ and also as much of the objects in $P^c$ as $g$ allows. Conversely, if $m$ is a feasible allocation such that $\sum_{i \in J} m_{io} = q_o$ for all $o \in P$ and $m_{io} = g_{io}$ for all $i \in J$ and $o \in P^c$, then

$$\sum_{i \in J} r_i = \sum_{i \in J} \sum_{o} m_{io} = \sum_{i \in J} \sum_{o \in P^c} m_{io} + \sum_{o \in P} \sum_{i \in J} m_{io} = \sum_{i \in J} \sum_{o \in P^c} g_{io} + \sum_{o \in P} q_o.$$

**Lemma 1.** *For $J \subset I$ and $P \subset O$ the following are equivalent:*

(a) *$(J, P)$ is critical for $E$;*

(b) There is a feasible allocation $m$ such that $\sum_{i \in J} m_{io} = q_o$ for all $o \in P$ and $m_{io} = g_{io}$ for all $i \in J$ and $o \in P^c$;

(b) For every feasible allocation $m$, $\sum_{i \in J} m_{io} = q_o$ for all $o \in P$ and $m_{io} = g_{io}$ for all $i \in J$ and $o \in P^c$.

In particular, if $(J, \emptyset)$ is critical, then then every feasible allocation $m$ has $m_{io} = g_{io}$ for all $i \in J$ and $o \in O$, and if $(\emptyset, P)$ is critical, then $q_o = 0$ for all $o \in P$. (The latter situation will arise during the allocation process as objects' quotas are exhausted.)

If $(J, P)$ is critical, let

$$E_{(J,P)} = (J, O, r|_J, q', g|_{J \times O}) \quad \text{and} \quad E^{(J,P)} = (J^c, P^c, r|_{J^c}, q'', g|_{J^c \times P^c})$$

where $q'_o = q_o$ if $o \in P$, $q'_o = \sum_{i \in J} g_{io}$ if $o \in P^c$, and $q'' \colon P^c \to \mathbb{R}_+$ is the function $q''_o = q_o - \sum_{i \in J} g_{io}$. Any feasible allocation for $E$ is the sum of a feasible allocation for $E_{(J,P)}$ and a feasible allocation for $E^{(J,P)}$, so $E_{(J,P)}$ and $E^{(J,P)}$ satisfy the GMC. Conversely, any sum of a feasible allocation for $E_{(J,P)}$ and a feasible allocation for $E^{(J,P)}$ is a feasible allocation for $E$. Thus a critical pair splits the given allocation problem into two smaller problems of the same type. This is very important because it allows our algorithm to be recursive.

We say that $E$ is *critical* if $(I, O)$ itself is a critical pair, which is the case if and only if $\sum_i r_i = \sum_o q_o$, so that any feasible allocation consumes all of the available resources. If $(J, P)$ is a critical pair for $E$, then $E_{(J,P)}$ is critical, and $E^{(J,P)}$ is critical if and only if $E$ is critical.

We say that $E$ is *simple* if there are no critical pairs $(J, P)$ with $(J, P) \neq (I, O)$. A critical pair $(J, P)$ is *minimal* if there is no critical pair $(J', P')$ with $J' \subset J$, $P' \subset P$, and $(J', P') \neq (J, P)$. The next result (whose proof follows easily from the discussion above and is therefore left as an exercise) implies that if $(J, P)$ is a minimal critical pair for $E$, then $E_{(J,P)}$ is simple.

**Lemma 2.** If $(J, P)$ is critical for $E$, $J' \subset J$, and $P' \subset P$, then $(J', P')$ is critical for $E$ if and only if it is critical for $E_{(J,P)}$.

A pair $(J, P)$ is critical if and only if $\sum_i r_i = c_E(S_{(J,P)})$, i.e., $S_{(J,P)}$ is a minimal cut for $c_E$. For any pairs $(J, P)$ and $(J', P')$ the definition of $S_{(J,P)}$ easily implies that $S_{(J,P)} \cup S_{(J',P')} = S_{(J \cup J', P \cup P')}$ and $S_{(J,P)} \cap S_{(J',P')} = S_{(J \cap J', P \cap P')}$. Since the set of minimal cuts for $c_E$ is a lattice we have:

**Proposition 2.** The set of critical pairs for $E$ is a lattice in the sense that if $(J, P)$ and $(J', P')$ are critical pairs, then so are $(J \cup J', P \cup P')$ and $(J \cap J', P \cap P')$.

If $(J, P)$ is a critical pair, then any feasible allocation $m$ has $m_{io} = 0$ for all $i \in J^c$ and $o \in P$, and in this sense $g_{io} > 0$ is illusory. We say that $E$ is *tight* if $g_{io} = 0$ for all critical pairs $(J, P)$ and all $i \in J^c$ and $o \in P$.

If $(J, P)$ is a critical pair for $E$, the $(J, P)$*-tightening of* $E$ is $E' = (I, O, q, r, g')$ where $g'_{io} = 0$ if $i \in J^c$ and $o \in P$, and otherwise $g'_{io} = g_{io}$. Since $E$ satisfies the GMC, it has a feasible allocation $m$, which necessarily has $m_{io} = 0$ for all $i \in J^c$ and $o \in P$, so it is a feasible allocation for $E'$, and consequently $E'$ satisfies the GMC. A *tightening sequence* for $E$ is a sequence $(J_1, P_1), \ldots, (J_\ell, P_\ell)$ for which there is a sequence $E_0 = E, E_1, \ldots, E_\ell$ of CEE's such that for each $j = 1, \ldots, \ell$, $(J_j, P_j)$ is a critical pair for $E_{j-1}$ and $E_j$ is the $(J_j, P_j)$-tightening of $E_{j-1}$. By induction each $E_j$ satisfies the GMC.

The following result is obvious:

**Lemma 3.** If $E = (I, O, r, q, g)$ satisfies the GMC, $(J, P)$ is a critical pair for $E$, $g' \leq g$, $E' = (I, O, r, q, g')$, and $E'$ satisfies the GMC, then $(J, P)$ is a critical pair for $E'$.

In view of the last result, if $(J_1, P_1), \ldots, (J_\ell, P_\ell)$ and $(J'_1, P'_1), \ldots, (J'_{\ell'}, P'_{\ell'})$ are tightening sequences, then so is $(J_1, P_1), \ldots, (J_\ell, P_\ell), (J'_1, P'_1), \ldots, (J'_{\ell'}, P'_{\ell'})$. Therefore starting with $E$ and repeatedly tightening with respect to critical pairs, including pairs that become critical as a result of the tightening, until no further tightening is possible, leads to a tight CEE that is independent of the order of tightening, that we call the *tightening of* $E$.

# 4 The Allocation Procedure

We now describe how the GCPS mechanism can be computed. We work with a fixed CEE $E = (I, O, r, q, g)$ that satisfies the GMC and a profile $\succ = (\succ_i)_{i \in I}$ of strict preferences[8] over $O$. (The symbols $\prec_i$, $\succeq_i$, and $\preceq_i$ have their usual meanings.) Let $T = \max_i r_i$. The allocation procedure is a piecewise linear function $p \colon [0, T] \to R$ with $p(0) = 0$, $p(t) \in R \setminus Q$ for all $t < T$, and $p(T) \in Q$. The *GCPS allocation* is

$$GCPS(E, \succ) = p(T).$$

At each moment the trajectory of $p$ increases, at unit speed, each agent's assignment of her favorite object, among those that are still available to her, while leaving other allocations fixed. This direction is adjusted when an agent attains her requirement, when an agent $i$'s assignment of an object $o$ reaches $g_{io}$, and when $p$ arrives at one of

[8] The mechanism actually depends only on the preferences of each agent $i$ over her set $\alpha_i$ of possible objects, but it is simplest to treat each $\succ_i$ as a complete ordering of $O$.

the facets of $R$. If $t^*$ is the first time such that $p(t^*)$ is in a facet of $R$, so that for some pair $(J, P)$ the GMC equality holds, then the residual CEE $E - p(t^*)$ is not simple and $(J, P)$ is a minimal critical pair for it. The GCPS allocation has a recursive definition: for $t \in (t^*, T]$, $p(t)$ is, by definition, the sum of $p(t^*)$ and the results of applying the allocation procedure to $(E - p(t^*))_{(J,P)}$ and $(E - p(t^*))^{(J,P)}$ on the interval $[t^*, T]$.

The main computational challenge is to detect when $p$ arrives at a facet of $R$.

One possible implementation first passes to the description of $Q$ as a convex hull of vertices. The vertices of $R$ are all the points obtained from vertices of $Q$ by changing some of the components to zero, and one may then pass from this set of vertices to the description of $R$ as an intersection of finitely many half spaces. The computational problem of passing from the description of a polytope as a convex hull of vertices to its description as an intersection of half spaces, and the reverse computation, are well studied, and efficient softwares for these tasks are available. (See Section 3 of Balbuzanov (2022).) However, even if the number of bounding inequalities of $Q$ and the number of bounding inequalities of $R$ are small, large data structures can arise at intermediate stages of the computation. For example, for the problem of assigning $n$ objects to $n$ agents the numbers of facet inequalities of $Q$ and $R$ are constant multiples of $n$, but $Q$ has $n!$ vertices.

Theorem 1 improves on this by showing that the facet inequalities of $R$ are a subset of the set of GMC inequalities. For a given $P \subset O$ it is easy to find the $J \subset I$ that minimizes the difference between the two sides of the GMC for $(J, P)$. Thus there is a computational burden that is roughly proportional to the number $2^{|O|}$ of subsets of $O$. An algorithm using this approach has been implemented, and works reasonably well for moderate (roughly $|O| \leq 50$) numbers of schools.

The procedure we describe now is much more efficient, especially for large problems. While computing $p$, we also compute an auxilary piecewise linear function $\overline{p}\colon [0, T] \to Q$ such that $p(t) \leq \overline{p}(t)$ for all $t$. We assume that $\overline{p}(0)$ is given. Possibly $\overline{p}(0)$ is the assignment of safe schools, or it may be the output of an algorithm that computes a maximal flow for the network $(N_E, A_E)$.

The combined function $(p, \overline{p})$ is piecewise linear, and $[0, T]$ is a finite union of intervals $[t_0, t_1], [t_1, t_2], \ldots, [t_{K-1}, t_K]$, where $t_0 = 0$ and $t_K = T$, such that on each interval $[t_k, t_{k+1}]$ the derivative of $(p, \overline{p})$ is constant. Suppose that we have already computed $p(t_k)$ and $\overline{p}(t_k)$. For each agent $i$ we compute the set $\alpha_i(t_k)$ of objects that are still possible for $i$, and we determine her $\succ_i$-favorite element $e_i^k$. Let $\theta^k \in \mathbb{Z}^{I \times O}$ be the matrix such that $\theta_{io}^k = 1$ if $o = e_i^k$, and otherwise $\theta_{io}^k = 0$.

There are now two possibilities. The first is that there is some $t' > t_k$ such that $p(t_k) + \theta^k(t - t_k) \in R$ for all $t \in [t_k, t']$. In this case we will find a $\theta \in \mathbb{Z}^{I \times O}$ such that for some $t' > t_k$ and all $t \in [t_k, t']$, $\overline{p}(t_k) + \theta(t - t_k) \in Q$ and

$$p(t_k) + \theta^k(t - t_k) \leq \overline{p}(t_k) + \theta(t - t_k). \tag{$*$}$$

Now $t_{k+1}$ is the first time after $t_k$ such that one of the following holds: a) $t_{k+1} = r_i$ for some $i$; b) $p_{ie_i^k}(t_{k+1}) = g_{ie_i^k}$ for some $i$; c) $\overline{p}(t_k) + \theta(t - t_k) \notin Q$ for $t > t_{k+1}$; d) $(*)$ does not hold for $t > t_{k+1}$. For $t \in [t_k, t_{k+1}]$ we have determined that $p(t) = p(t_k) + \theta^k(t - t_k)$, and we set $\overline{p}(t) = \overline{p}(t_k) + \theta(t - t_k)$. Having determined $p(t_{k+1})$ and $\overline{p}(t_{k+1})$, we repeat the process.

The second possibility is that it is not possible to continue $p$, as described above, without leaving $R$, because $p(t_k)$ satisfies the GMC equality of some pair $(J, P)$. In this case we will find such a pair, then descend recursively to the computation of the GCPS allocations of $(E - p(t_k))_{(J,P)}$ and $(E - p(t_k))^{(J,P)}$. We now describe an algorithm that determines which of these possibilities holds, finding a satisfactory $\theta$ in the first case and a critical $(J, P)$ in the second case.

Suppose that there is a $\theta \in \mathbb{Z}^{I \times O}$ such that for there exists a $t' > t_k$ such that for all $t \in [t_k, t']$, $(*)$ holds and $\overline{p}(t_k) + \theta(t - t_k) \in Q$. Together these conditions imply that $p(t_k) + \theta^k(t - t_k) \in R$, so the first possibility above holds, and we can use $\theta$ to define the continuation of $\overline{p}$. The algorithm may be thought of as a search for such a $\theta$.

For a given $\theta$, a $t' > 0$ as above exists if and only if $\theta$ satisfies the following conditions:

(a) For each $i$ and $o$:

  (i) If $o \notin \alpha_i(t_k)$, then $\theta_{io} = 0$.

  (ii) If $\overline{p}_{io}(t_k) = p_{io}(t_k)$, then $\theta_{io} \geq 0$, and if, in addition, $o = e_i^k$, then $\theta_{io} \geq 1$.

  (iii) If $\overline{p}_{io}(t_k) = g_{io}$, then $\theta_{io} \leq 0$.

(b) For each $i$, $\sum_o \theta_{io} = 0$.

(c) For each $o$, if $\sum_i \overline{p}_{io}(t_k) = q_o$, then $\sum_i \theta_{io} \leq 0$.

Our search for a suitable $\theta$ begins by defining an initial $\theta \in \mathbb{Z}^{I \times O}$ as follows. For each $i$, if $\overline{p}_{ie_i^k}(t_k) > p_{ie_i^k}(t_k)$, then we set $\theta_{io} = 0$ for all $o$. If $\overline{p}_{ie_i^k} = p_{ie_i^k}(t_k)$, then we set $\theta_{ie_i^k} = 1$, we set $\theta_{io_i} = -1$ for some $o_i \neq e_i^k$ such that $\overline{p}_{io_i}(t_k) > p_{io_i}(t_k)$, and we set $\theta_{io} = 0$ for all other $o$. Evidently $\theta$ satisfies (a) and (b).

Let

$$\tilde{P} = \{\, o : \sum_i \overline{p}_{io}(t_k) = q_o \text{ and } \sum_i \theta_{io} > 0 \,\}.$$

If $\sum_{o \in \tilde{P}} \sum_i \theta_{io} = 0$, then (c) holds. Suppose that this is not the case. We now describe a construction that may or may not be possible. When it is possible, it passes from $\theta$ to a $\theta' \in \mathbb{Z}^{I \times O}$ satisfying (a) and (b) such that if $\tilde{P}' = \{\, o : \sum_i \overline{p}_{io}(t_k) = q_{o_h}$ and $\sum_i \theta'_{io} > 0 \,\}$, then $\sum_{o \in \tilde{P}'} \sum_i \theta'_{io} = \sum_{o \in \tilde{P}} \sum_i \theta_{io} - 1$. Repeating this construction will eventually produce an element of $\mathbb{Z}^{I \times O}$ satisfying (a)–(c) unless, at some point, the construction becomes impossible.

Choose $o_0 \in \tilde{P}$, and let $P_0 = \{o_0\}$. We define sets $J_1, P_1, J_2, P_2, \ldots$ inductively. If $P_{h-1}$ is given, let $J_h = \bigcup_{o \in P_{h-1}} J_h(o)$ where

$$J_h(o) = \{\, i : o \in \alpha_i \text{ and if } \overline{p}_{io}(t_k) = p_{io}(t_k), \text{ then } \theta_{io} > 0 \text{ and } \theta_{io} > 1 \text{ if } o = e_i^{\succ} \,\}.$$

If $J_h$ is given, let $P_h = \bigcup_{i \in J_h} P_h(i)$ where

$$P_h(i) = \{\, o \in \alpha_i : \theta_{io} < 0 \text{ if } \overline{p}_{io}(t_k) = g_{io} \,\}.$$

Suppose that for some $h$ there is an $o_h \in P_h \setminus P_{h-1}$ such that either $\sum_i \overline{p}_{io_h} < q_{o_h}$ or $\sum_i \theta_{io_h} < 0$. We can find a $i_h \in J_h$ such that $o_h \in J_h(i_h)$, then find an $o_{h-1} \in P_{h-1}$ such that $i_h \in J_h(o_{h-1})$, and so forth. Define $\theta'$ by setting

$$\theta'_{i_g o_{g-1}} = \theta_{i_g o_{g-1}} - 1 \quad \text{and} \quad \theta'_{i_g o_g} = \theta_{i_g o_g} + 1$$

for $g = 1, \ldots, h$ and $\theta'_{io} = \theta_{io}$ for all other $(i, o)$.

It is easy to see that $\theta'$ satisfies (a), because $\theta'_{io}$ differs from $\theta_{io}$ only when the difference is permitted, according to (i)–(iii). For each $i$, $\sum_o \theta'_{io} = \sum_o \theta_{io}$, either by construction if $i = i_g$ for some $g$, or because $\theta'_{io} = \theta_{io}$ for all $o$. Since $\theta$ satisfies (b), $\theta'$ also satisfies (b).

For $o \notin \{o_0, \ldots, o_h\}$ we have $\theta'_{io} = \theta_{io}$ for all $i$ and thus $\sum_i \theta'_{io} = \sum_i \theta_{io}$. For each $g = 1, \ldots, h-1$ we have $\sum_i \theta'_{io_g} = \sum_i \theta_{io_g}$ by construction. Clearly $\sum_i \theta'_{io_0} = \sum_i \theta_{io_0} - 1$ and $\sum_i \theta'_{io_h} = \sum_i \theta_{io_h} + 1$. Since either $\sum_i \overline{p}_{io_h} < q_{o_h}$ or $\sum_i \theta'_{io_h} \le 0$, $\theta'$ has all desired properties.

It is impossible to construct $\theta'$ in this manner if there is no $h$ and $o_h \in P_h \setminus P_{h-1}$ such that $\sum_i \overline{p}_{io_h} < q_{o_h}$ or $\sum_i \theta_{io_h} < 0$. Supposing that this is the case, let $J = \bigcup_h J_h$ and $P = \bigcup_h P_h$. We have $\sum_i \overline{p}_{io}(t_k) = q_o$ for all $o \in P$. If $o \in P$ and $i \notin J$, then $\overline{p}_{io}(t_k) = p_{io}(t_k)$. If $i \in J$ and $o \notin P$, then $\overline{p}_{io}(t_k) = g_{io}$ if $o \in \alpha_i \setminus P$, and $\overline{p}_{io}(t_k) = g_{io} = 0$ if $o \notin \alpha_i$. Thus $\overline{p}(t_k) - p(t_k)$ is a feasible allocation for $E - p(t_k)$ that gives all of the resources in $P$ to agents in $J$, and it gives $g_{io} - p_{io}(t_k)$ to $i \in J$ whenever $o \in O \setminus P$, so, by Lemma 1, $(J, P)$ is a critical pair for $E - p(t_k)$.

Summarizing, the algorithm repeatedly extends $p$ and $\overline{p}$ to intervals such as $[t_k, t_{k+1}]$

until $p(t_k)$ satisfies the GMC equality for a pair $(J, P)$, at which point it descends recursively to computation of the GCPS allocations of $(E - p(t_k))_{(J,P)}$ and $(E - p(t_k))^{(J,P)}$. If $p(t_k)$ does not satisfy such a GMC inequality, it finds a $\theta$ satisfying (a)–(c) by repeatedly adjusting a $\theta$ that satisfies (a) and (b) until it also satisfies (c).

The theoretical and practical complexity of the algorithm depends on the following factors:

(a) How many linear segments can the piecewise linear function $(p, \overline{p})$ have, and how many does it typically have?

(b) How many adjustments $\theta \to \theta'$ can be required to achieve a $\theta$ satisfying (a)–(c), and how many are typically required?

(c) How long can the paths $o_o, i_1, o_1, \ldots, i_h, o_h$ be, and how long are they typically?

The algorithm has been implemented for school choice problems in the software package *GCPS Schools*, as described in Appendix B. A computational experiment described there illuminates these issues.

It seems to be quite difficult to get a theoretical upper bound on the number of linear segments of $(p, \overline{p})$, and even to prove that the algorithm has polynomial time worst case complexity. Empirically, the number of linear segments is roughly proportional to the number of students.

The construction of the initial $\theta$ implies that the number of students is an upper bound on the number of adjustments of $\theta$. In practice the number of adjustments seems to vary between 5% and 20% of this number.

The number of schools is a crude upper bound on the length of a path arising in the adjustment of $\theta$, and it is not obvious how one might develop a tighter theoretical bound. In practice paths are quite short, with an average length between one and two.

Overall the implementation of the algorithm seems to work quite well. Extrapolating from currently available data, it seems reasonable to hope that it will be applicable to the largest school choice problems, with running times measured in hours or at worst days.

# 5 Nondichotomous Priorities

We take as given a school choice CEE $E = (I, O, 1, q, g)$ and a profile $\succ = (\succ_i)$ of preferences over $O$. The most basic application of the procedure described in the last section computes the GCPS allocation for dichotomous priorities, so that each $g_{io}$ is either $0$ or $1$.

Many schools systems use more complex priorities that assign points based on grades, test scores, minority status, the student's proximity to the school, and perhaps other factors. DA typically refines these priorities by replacing them with (perhaps randomly generated) strict refinements, and it then implicitly computes an approximate priority cutoff for each school. In this section we consider how the GCPS mechanism might be used to achieve a similar effect.

Azevedo and Leshno (2016) study a model with a continuum of students in which the schools' cutoffs are analogous to prices. Our framework is similar, insofar as our students are, in effect, infinitely divisible. The discussion below, and the proof of Theorem 2 in Appendix C, follows the corresponding material in Appendix A of their paper. Similar ideas appear in Abdulkadiroğlu et al. (2015).

We assume that for each student $i$ and school $o$, the student's priority at $o$ is a positive integer $e_{io}$, where the school "prefers" higher priority students. We take advantage of our framework's ability to allocate fractional quantities of seats by having cutoffs that are not integral. For a positive integer $e$ and a real number $c \geq 0$ let

$$\rho(e,c) = \begin{cases} 1, & c \leq e-1, \\ e-c, & c \in (e-1,e), \\ 0, & c \geq e. \end{cases}$$

If school $o$ has cutoff $c_o$, $\min\{g_{io}, \rho(e_{io}, c_o)\}$ is the maximum allowed consumption of $o$ by $i$.

Individual demand of student $i$, as a function of a vector of cutoffs $c \in \mathbb{R}^O_+$, is generated by having the student consume as much of the favorite school as allowed, then as much of the second favorite school as allowed, and so forth, until either she has one unit of probability or she has consumed as much of each school as she is allowed to consume. For each $i$ the demands $D_{io}(c)$ of $i$ for the various schools $o$ can be defined implicitly by requiring that:

(a) $0 \leq D_{io}(c) \leq \min\{g_{io}, \rho(e_{io}, c_o)\}$;

(b) $\sum_o D_{io}(c) \leq 1$;

(c) $D_{io}(c) = \min\{g_{io}, \rho(e_{io}, c_o)\}$ if either $\sum_o D_{io}(c) < 1$ or there is an $o'$ such that $o' \prec_i o$ and $D_{io'}(c) > 0$.

It is not difficult to show that each $D_{io}$ is a continuous function. For $o \in O$ let $D_o(c) = \sum_i D_{io}(c)$. An obvious but important property of demand, which Azevedo and Leshno call *gross substitutes*, is that when $o \neq o'$, $D_{io}(c)$ and $D_o(c)$ are nondecreasing functions of $c_{o'}$.

For $o \in O$ and $c_{-o} \in \mathbb{R}_+^{O \setminus \{o\}}$ let

$$I_o(c_{-o}) = \{\, c_o \geq 0 : D_o(c_o, c_{-o}) \leq q_o \text{ and } D_o(c_o, c_{-o}) = q_o \text{ if } c_o > 0 \,\}.$$

Let $\overline{e}$ be an integer that is larger than any $e_{io}$. For each $i$, $D_{io}(\cdot, c_{-o})$ is continuous and nonincreasing, and $D_{io}(\overline{e}, c_{-o}) = 0$. Therefore $D_o(\cdot, c_{-o})$ is continuous and nonincreasing, and $D_o(\overline{e}, c_{-o}) = 0$. Thus $I_o(c_{-o})$ is a nonempty closed subinterval of $[0, \overline{e}]$, either by the intermediate value theorem or because $D_o(0, c_{-o}) < q_o$ and $I_o(c_{-o}) = \{0\}$.

Let $T \colon [0, \overline{e}]^O \to [0, \overline{e}]^O$ be the function with component functions

$$T_o(c) = \operatorname*{arg\,min}_{c'_o \in I_o(c_{-o})} |c'_o - c_o|.$$

The set

$$\mathcal{F} = \{\, c \in [0, \overline{e}]^O : \text{for each } o,\ D_o(c) \leq q_o \text{ and } D_o(c) = q_o \text{ if } c_o > 0 \,\}$$

of fixed points of $T$ is the set of market clearing cutoffs.

**Theorem 2.** $\mathcal{F}$ is a complete lattice[9], and $D_{io}(c) = D_{io}(c')$ for all $i$, $o$, and $c, c' \in \mathcal{F}$.

For each $i$ and $o$ let $d^*_{io}$ be the common value of $D_{io}(c)$ for $c \in \mathcal{F}$, and let $d^*$ be the matrix with these entries.

In general $d^*$ need not be a feasible allocation. For example, if student $i$'s acceptable schools all have high demand, it can happen that for each of them the only possible values of the cutoff are above the student's priority at the school, in which case $d^*_{io} = 0$ for all $o$. As we mentioned at the outset, this happens also with DA. Possible responses to this are to use either DA or GCPS in a multiround system, or to assign such students administratively. However, the priorities are, in principle, expressions of society's values that are intended to influence the outcome in a socially desirable direction. A failure to provide a seat to a student may be regarded as a failure of the priorities to correctly express society's values, which suggests that they should be reconsidered.

If it is feasible, $d^*$ is indeed an ideal fulfillment of the priorities, providing each school only to those for whom it has the greatest social value, and avoiding the inefficiencies that can arise when DA is applied to strict priorities that refine the coarse

[9]For $x, y \in \mathbb{R}^O$, $x \vee y$ and $x \wedge y$ are the vectors with components $(x \vee y)_o = \max\{x_o, y_o\}$ and $(x \wedge y)_o = \min\{x_o, y_o\}$. A set $L \subset \mathbb{R}^O$ is a *lattice* if, for all $x, y \in L$, $x \vee y$ and $x \wedge y$ are elements of $L$, and it is a *complete lattice* if, in addition, for every $S \subset L$, $L$ contains a least upper bound $\bigvee S$ and a greatest lower bound $\bigwedge S$. Since $\bigvee \emptyset$ and $\bigwedge \emptyset$ are elements of $L$, a complete lattice is nonempty.

priorities that actually express social values. For example, if the coarse priorities of Ann and Bob for $o_1$ and $o_2$ are the schools' cutoff priorities, Ann prefers $o_1$, and Bob prefers $o_2$, DA may nevertheless assign Ann to $o_2$ and Bob to $o_1$. If $c$ is a vector of cutoffs such that $D_{io}(c) = d^*_{io}$ for all $i$ and $o$, then, by Theorem 3, $d^*$ is $sd$-efficient for the school choice CEE $E' = (I, O, 1, q, g')$ where $g'_{io} = \min\{g_{io}, \rho(e_{io}, c_o)\}$, so it cannot happen that both $d^*_{\text{Ann},o_2} > 0$ and $d^*_{\text{Bob},o_1} > 0$.

We now describe a computational procedure that attempts to compute $d^*$. To begin with we define and characterize a useful function. For a positive integer $e$, a number $c \in [e-1, e]$, a student $i$, and $p_i \in [0,1]^O$ such that $\sum_o p_{io} \le 1$ let

$$\tau_{io}(e, c, p_i) = \begin{cases} 0, & e_{io} < e, \\ e - c, & e_{io} = e \text{ and } p_{io} \ge e - c, \\ \min\{e - c, \sum_{o' \preceq_i o} p_{io'}\}, & e_{io} = e \text{ and } p_{io} \le e - c, \\ \sum_{o' \preceq_i o} p_{io'}, & e_{io} > e. \end{cases}$$

The main idea will be to choose $e$ and $c$ such that $\sum_i \tau_{io}(e, c, p_i) = q_o$.

We claim that $\tau_{io}(e, e-1, p_i) = \tau_{io}(e-1, e-1, p_i)$. This is obviously the case when $e_{io} > e$ and when $e_{io} < e-1$, and it is easy to see that this holds when $e_{io} = e-1$. When $e_{io} = e$ we have $\tau_{io}(e-1, e-1, p_i) = \sum_{o' \preceq_i o} p_{io'}$ and $\tau_{io}(e, e-1, p_i) = \min\{1, \sum_{o' \preceq_i o} p_{io'}\} = \sum_{o' \preceq_i o} p_{io'}$. Thus we may regard $\tau_{io}$ as a function of $c \ge 0$ and $p_i$. Clearly $\tau_{io}(\cdot, p_i)$ is continuous on each interval $[e-1, e]$, so it is continuous, and it is obviously nonincreasing.

We begin with a school choice CEE $E^0 = (I, O, 1, q, g^0)$ that has $g^0_{io} \in [0,1]$ for all $i$ and $o$. We assume that it satisfies the GMC, and that all the derived CEE's below also satisfy the GMC. Let $p^1$ be the partial allocation resulting from applying the GCPS algorithm to $E^0$. For each $o$, since $\sum_i \tau_{io}(0, p^1_i) = \sum_i \sum_{o' \preceq_i o} p^1_{io'}$ and $\sum_i \tau_{io}(\bar{e}, p^1_i) = 0$, we may define $c^1_o$ to be either $0$ if $\sum_i \sum_{o' \preceq_i o} p^1_{io'} < q_o$ or the smallest number such that $\sum_i \tau_{io}(c^1_o, p^1_i) = q_o$. For each $i$ and $o$ let $g^1_{io} = \min\{g^0_{io}, \rho(e_{io}, c^1_o)\}$. Let $E^1 = (I, O, 1, q, g^1)$.

We continue this construction inductively. Suppose that $E^k = (I, O, 1, q, g^k)$ is given. Let $p^{k+1}$ be the partial allocation resulting from applying the GCPS algorithm to $E^k$. For each $o$ define $c^{k+1}_o$ to be either $0$ if $\sum_i \sum_{o' \preceq_i o} p^{k+1}_{io'} < q_o$ or the smallest number such that $\sum_i \tau_{io}(c^{k+1}_o, p^{k+1}_i) = 0$. For each $i$ and $o$ let $g^{k+1}_{io} = \min\{g_{io}, \rho(e_{io}, c^{k+1}_o)\}$. Let $E^{k+1} = (I, O, 1, q, g^{k+1})$.

Due to the complexity of the GCPS mechanism, proving that $p^k \to d^*$ would certainly be hard, at best. At the same time it is hard to imagine that the direct effect of setting $c_o$ to equilibrate supply and demand of $o$ can be overwhelmed by the indirect effects coming from adjusting the $c_{o'}$ for $o' \neq o$. (The analysis in the first part of Online Appendix D gives a detailed picture of these indirect effects.) We are optimistic that this method will be a practically reliable scheme for computing $d^*$, but this remains to be seen. The issue of whether there is an efficient algorithm for computing $d^*$ is currently unresolved.

Policy issues related to priorities are extremely complex. On the one hand it seems desirable to provide the best resources to those who can extract the greatest benefit, so it makes sense to give high priority to students with good grades or test scores, but doing so could exacerbate inherited inequality. Giving students points for proximity to a school may facilitate allowing each student to attend a nearby school, but preventing a student from attending a distant school if she is willing to incur the travel costs has a paternalistic aspect. The literature on peer effects in education (Epple and Romano, 2011, Sacerdote, 2011) is extensive, finding significant effects with causal pathways that are not yet well understood. In particular, Burke and Sass (2013) find that low achieving students derive substantial benefits from having high achieving peers, and Vigdor and Nechyba (2007) find that classroom heterogeneity can lead to higher test scores. One could easily list additional issues.

Balancing various concerns in practice requires information concerning what would actually happen under various policies. Because our mechanism allows priorities to be coarse rather than strict, we widen the range of alternatives that can be considered, and we more clearly distinguish between priorities that reflect actual societal values and those that merely fulfill the requirements of a mechanism. Our computational methods potentially allow easy computation of counterfactual outcomes resulting from applying various alternatives to historical data.

# 6 Efficiency

In this section we work with a CEE $E = (I, O, r, q, g)$ that satisfies the GMC and a fixed profile of preferences $\succ$. Our objective is to show that the GCPS mechanism applied to $E$ and $\succ$ yields an allocation that is efficient in strong senses. However, we should first of all mention that mechanisms that are *ordinal* (that is, based on the agents' reports of ordinal preferences) and nondictatorial often allow allocations that are inefficient relative to cardinal utility functions consistent with the ordinal preferences (Featherstone and Niederle, 2008, Miralles, 2009, Abdulkadiroğlu et al., 2011, Troyan, 2012, Abdulkadiroğlu et al., 2015).

For $i \in I$, an *allocation for* $i$ is a vector $m_i \in \mathbb{R}^O_+$ such that $m_i \leq g_i$ and $\sum_o m_{io} = r_i$. The *stochastic dominance relation* $sd(\succ_i)$ on allocations for $i$ derived from $\succ_i$ is defined by $m'_i \, sd(\succ_i) \, m_i$ if $\sum_{p \succeq_i o} m'_{ip} \geq \sum_{p \succeq_i o} m_{ip}$ for all $o \in O$. Usually in applications of this concept $m_i$ is a probability distribution on $O$, but the concept makes perfect sense in our more general context, and standard arguments generalize straightforwardly to show that $m'_i \, sd(\succ_i) \, m_i$ if and only if $\sum_o m'_{io} u_i(o) \geq \sum_o m_{io} u_i(o)$ for any cardinal utility function $u_i : O \to \mathbb{R}$ such that for all $o, o' \in O$, $u_i(o) \geq u_i(o')$ if and only if $o \succeq_i o'$.

Two other well-studied extensions of a given preference to preferences over lotteries relate to lexicographic preferences (Cho, 2016; Schulman and Vazirani, 2015; Cho and Doğan, 2016; Saban and Sethuraman, 2014; Cho, 2018). The first extension, which is called the *downward lexicographic* extension ($dl$-extension) compares two $i$-allocations as follows. One of the $i$-allocations is preferred if it assigns a higher amount of the most preferred object than the other. If the two $i$-allocations assign the same amount of the most preferred object, then the one that is preferred is the one that assigns the greater amount of the second most preferred object. If the two amounts are equal again, then the $i$-allocation that assigns a greater amount of the third most preferred object is preferred, and so on. The second extension, which is called the *upward lexicographic* extension ($ul$-extension) is a dual of the $dl$-extension. It lexicographically minimizes amounts of less preferred objects, starting from the least preferred object[10]. The $dl$- and $ul$-extensions yield preferences that represent the limits of standard vNM utility functions with extreme risk loving and extreme risk aversion, respectively.

For $e \in \{sd, dl, ul\}$, a feasible allocation $m'$ $e$-*dominates* another feasible allocation $m$ if $m'_i \, e(\succ_i) \, m_i$ for all $i$ and there is some $i$ such that $m'_i \neq m_i$. A feasible allocation $m$ is $e$-*efficient* if there is no feasible allocation that $e$-dominates it.

**Theorem 3.** For $e \in \{sd, dl, ul\}$, the GCPS allocation for $E$ and $\succ$ is $e$-*efficient*.

# 7 Fairness for School Choice

We now briefly consider fairness properties of the GCPS mechanism applied to a school choice CEE $E = (I, O, 1, q, g)$ that satisfies the GMC and a profile of preferences $\succ$. It is obvious from the definition that the mechanism satisfies anonymity (the outcomes do not depend on the ordering of the agents, or their "names") and equal treatment of equals (the GCPS gives the same allocations to $i$ and $j$ if $r_i = r_j$, $g_i = g_j$, and $\succ_i = \succ_j$).

The other fairness property considered by BM is envy-freeness. They show that the PS mechanism is envy-free in the strong sense that if $m_i$ and $m_j$ are the allocations of

[10]Formally, the *downward lexicographic relation* $dl(\succ_i)$ on allocations for $i$ derived from $\succ_i$ is defined by specifying that $m'_i \, dl(\succ_i) \, m_i$ if either $m'_i = m_i$ or there is an $o \in O$ such that $\sum_{p \succeq_i o'} m'_{ip} = \sum_{p \succeq_i o'} m_{ip}$ for all $o' \in O$ such that $o' \succ_i o$ and $\sum_{p \succeq_i o} m'_{ip} > \sum_{p \succeq_i o} m_{ip}$. The *upward lexicographic relation* $ul(\succ_i)$ is defined by specifying that $m'_i \, ul(\succ_i) \, m_i$ if either $m'_i = m_i$ or there is an $o \in O$ such that $\sum_{o' \succeq_i p} m'_{ip} = \sum_{o' \succeq_i p} m_{ip}$ for all $o' \in O$ such that $o \succ_i o'$ and $\sum_{o \succeq_i p} m'_{ip} < \sum_{o \succeq_i p} m_{ip}$.

the PS mechanism for $\succ$, then $m_i \, sd(\succ_i) \, m_j$. It is not reasonable to expect this if the two agents have different opportunities, and in recognition of this Abdulkadiroğlu and Sönmez (2003) introduced a notion of no justified envy. This concept takes on different meanings depending on the setting. (For a recent discussion see Romm et al. (2020).) We follow Yilmaz (2010) in the context of school choice. If $E$ is a school choice CEE with $g \in \{0,1\}^{I \times O}$, we say that $m \in Q$ *has no justified envy* if, for all $i, j \in I$, if $\alpha_i \subset \alpha_j$ and $o_i \succ_i o_j$ for all $o_i \in \alpha_i$ and $o_j \in \alpha_j \setminus \alpha_i$, then $m_i \, sd(\succ_i) \, m_j$. Intuitively, $i$'s envy of $j$ is not justified if $i$ is not eligible to attend a desirable element of $\alpha_j$, or if $j$ can demand a seat in a desirable element of $\alpha_i$ because less desirable elements of $\alpha_i$ are not in $\alpha_j$.

**Proposition 3.** If $E$ is a school choice CEE with $g \in \{0,1\}^{I \times O}$, then $GCPS(\succ)$ has no justified envy.

# 8 Strategy Proofness for School Choice

With the exception of some of the material towards the end of Section 4, and the result in the last section, up to this point our results have concerned the general GCPS mechanism. In this section we examine the extent to which the GCPS mechanism, applied specifically to school choice with safe schools, is resistant to manipulation. We fix an integral school choice CEE $E = (I, O, 1, q, g)$ that satisfies the GMC and a profile $\succ$ of preferences over $O$. For each student $j$ the safe school is the $\succ_j$-worst element of $\alpha_j$. We also fix a particular student $i \in I$ whose possible deviations from truthful reporting we will study.

BM (p. 310) show that there is no probabilistic allocation mechanism for object allocation that is ex post efficient, strategy proof, and envy free. Theorem 4 of Yilmaz (2010) states that for house allocation problems with existing tenants, there is no mechanism that is individually rational, strategy proof, and has no justified envy. Their settings are special cases of ours, so we cannot hope that the GCPS mechanism is fully strategy proof: there necessarily exist situations in which, for some vNM utility function consistent with the true ordinal preferences, expected utility can be increased by reporting a false preference. Nevertheless we will argue that for school choice the failures of strategy proofness of the GCPS mechanism are minor when each school has many students, and do not significantly impair its usefulness.

There are three different ways a student might try to manipulate: a) reporting that some of the schools that are actually worse than the safe school are better than it; b) reporting that some of the schools that are actually better than the safe school are worse; c) reordering of the schools that are better than the safe school. A manipulation attempt of type a) will be called an *augmentation*; following Roth and Rothblum (1999), a manipulation attempt of type b) will be called a *truncation*; a manipulation attempt of type c) will be called a *reordering*. We discuss these in turn.

Manipulation by augmentation is unambiguously unsuccessful:

**Theorem 4.** Let $\alpha'_i = \alpha_i \cup \{o^*\}$, where $o^*$ is an element of $O \setminus \alpha_i$, and let $\succ'_i$ be a preference over $O$ that has $\alpha'_i$ as the set of schools weakly preferred to the safe school, and that agrees with $\succ_i$ on $\alpha_i$. Let $\succ' = (\succ'_i, \succ_{-i})$. Then

$$GCPS_i(\succ)\, sd(\succ_i)\, GCPS_i(\succ').$$

As a matter of logic, this result does not rule out the possibility that other forms of manipulation might become successful, or more successful, if supplemented with an augmentation, but this possibility seems to have slight practical importance.

Yilmaz (2010) (Example 5) presents the following example of an unambiguously successful truncation manipulation for a house allocation problem with existing tenants[11]. There are three homeowners and three houses, with $1$ endowed with $a$, $2$ endowed with $b$, and $3$ endowed with $c$, and preferences $b \succ_1 c \succ_1 a$, $a \succ_2 b$, and $b \succ_3 a \succ_3 c$. In the GCPS process $b$ is exhausted at time $\frac{1}{2}$, which results in $P = \{a\}$ becoming critical, with $J_P = \{2\}$, so the GCPS allocation gives $\frac{1}{2}b + \frac{1}{2}c$ to $1$, $a$ to $2$, and $\frac{1}{2}b + \frac{1}{2}c$ to $3$. If $1$ reports the preference $b \succ'_1 a$ (i.e., $b \succ'_1 a \succ'_1 c$) then $P = \{a, b\}$ is critical at time $0$, and the allocation gives $b$ to $1$, $a$ to $2$, and $c$ to $3$. As Yilmaz points out, this manipulation continues to be possible in any of the problems obtained by replacing each homeowner-current house pair with any number of copies. For example, if there are five copies of agent $1$, three copies of agent $2$, and two copies of agent $3$, and one of the copies of agent $1$ reports $b \succ'_1 a$, then $P = \{a, b\}$ is critical at time $\frac{2}{3}$, and the allocation gives $b$ to the deviator, $a$ to the agents of type $2$, and $\frac{2}{3}b + \frac{1}{3}c$ to nondeviant agents of type $1$ and agents of type $3$.

Azevedo and Budish (2019) introduce a notion of asymptotic strategy proofness in large economies: a mechanism is *strategy proof in the large* if, along any sequence of increasingly large economies in which each agent's belief concerning the types of the other agents is given by i.i.d. draws from a fixed distribution over the set of possible types, the maximal gains from manipulation vanish in the limit. They show that if a mechanism is envy-free, then it is strategy proof in the large. The intuition is that the gain from reporting an incorrect type is the sum of the gain from getting that type's al-

[11]Theorem 3 of Cho (2018) asserts that the PS mechanism is $dl$-strategy proof, which means that manipulation never results in a $dl$-better allocation. This example shows that Cho's result does not extend to house allocation problems with existing tenants.

location, which is nonpositive by envy-freeness, plus the gain from changing the overall allocation, which diminishes as the agent's importance in the economy shrinks. Because the example above is robust with respect to the numbers of the three types, it shows that the GCPS mechanism is not strategy proof in the large, but neverthless the intuition still has at least an informal applicability, which is reflected in Theorem 5 below.

In general, in order for a truncation manipulation to succeed, the manipulation must induce a critical set of schools $P$ at some time during the allocation process that includes the manipulating student's safe school and a school or schools that that student wishes to consume, and that does not include the school or schools that the student falsely reports to be worse than her safe school, so that students who can consume those schools are outside of $J_P$ and thus prevented from consuming the schools the manipulator desires. The manipulator's safe school must be in high enough demand that the manipulator does not end up simply consuming more of that school, and the schools that the manipulator desires must also be in high demand, since otherwise the manipulator can get what she wants simply by ranking her favorite of those schools at the top of her reported preference, in which case the manipulation does not gain anything.

This set of requirements is rather lengthy and specific, but this type of manipulation does not seem to have a knife-edge quality, and one can easily imagine successful truncation manipulations in quite complex settings. In addition, it does not seem that the student needs highly specific information in order to know that the manipulation is likely to succeed or at least do no harm. In particular, if the student believes that her safe school is so popular that there is no chance that a truncation manipulation attempt will result in her consuming more of her safe school, then such an attempt can (roughly) weakly dominate truthful revelation.

To what extent does this type of manipulation impair the usefulness of the GCPS mechanism for school choice? As we explain below, reordering manipulations are quite difficult, so a student whose safe school is not highly popular is (with minor exceptions) incentivized to reveal truthfully, independent of the extent to which others are attempting truncation manipulations. Thus the Nash equilibria of the mechanism are not drastically different from the naive understanding of it obtained by assuming truthful revelation. The outcome of the mechanism when there are successful truncation manipulations is $sd$-efficient for the true preferences. Indeed, truncation manipulations seem to be mainly an annoyance from the point of view of fairness: a student with a highly desired safe school may have an opportunity (possibly at some risk) to amplify her good fortune at others' expense.

In the remainder of this section we consider reordering manipulations. Proposition 1 of BM asserts (among other things) that the PS mechanism is weakly strategy proof:

if reporting a false preference gives an allocation that is weakly $sd$-preferred to the allocation resulting from truthful revelation, then the two allocations are the same. Kojima (2009) shows, by means of the following example, that weak strategy proofness does not extend to the allocation of $r \geq 2$ objects per agent. Let there by two agents 1 and 2 and four objects $a$, $b$, $c$, and $d$, so $r = 2$. Let the true preferences be $a \succ_1 b \succ_1 c \succ_1 d$ and $b \succ_2 c \succ_2 a \succ_2 d$. If the agents report these preferences, then the PS mechanism gives $(1, 0, \frac{1}{2}, \frac{1}{2})$ to agent 1 and $(0, 1, \frac{1}{2}, \frac{1}{2})$ to agent 2. On the other hand, if agent 1 reports $\succ_1'$, where $b \succ_1' a \succ_1' c \succ_1' d$, and agent 2 reports $\succ_2$, then the PS mechanism gives $(1, \frac{1}{2}, 0, \frac{1}{2})$ to agent 1 and $(0, \frac{1}{2}, 1, \frac{1}{2})$ to agent 2. Thus, when agent 1 reports the truth she receives the probability distribution over pairs $\frac{1}{2}(a, c) + \frac{1}{2}(a, d)$, while misrepresenting yields $\frac{1}{2}(a, b) + \frac{1}{2}(a, d)$, which stochastically dominates (in an obvious sense) the allocation resulting from truthful reporting.

A key idea of BM's proof of their Proposition 1 is that once other agents begin eating an object, they continue until that object is exhausted. Consequently, in order to obtain the same amount of her favorite object as in the allocation resulting from truthful revelation, an agent must report that it is her favorite. In order to get the maximal amount of her best among the objects that are still available after her favorite has been fully allocated, she cannot report a preference that ranks it below some other object that is still available, and so forth inductively. In the example above, consumption of a school by other students can cease before the school is fully allocated, which allows student 1 to advantageously defer its consumption.

Since students do not consume multiple seats, one might hope that the GCPS mechanism is weakly strategy proof for school choice, but the following example shows that this is not the case. There are five schools with $q_a = q_b = q_c = q_d = 1$ and $q_e \geq 4$. There are eight students, with preferences $a \succ_1 b \succ_1 c \succ_1 d$, $a \succ_2 e$, $a \succ_3 e$, $b \succ_4 e$, $b \succ_5 e$, $c \succ_6 e$, $d \succ_7 e$, and $c \succ_8 d$. (For each student the lowest ranked school is the safe school.) Up until time $\frac{1}{3}$ each student consumes her favorite school. At time $\frac{1}{3}$ school $a$ is exhausted, and the set $\{b, c, d\}$ also becomes critical, with remaining capacity $\frac{1}{3}b + \frac{1}{3}c + \frac{2}{3}d$ that is just sufficient to serve the needs of students 1 and 8, who cannot attend $e$. If student 1 reports truthfully she receives $\frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}d$ because student 8 consumes what remains of school $c$ between time $\frac{1}{3}$ and time $\frac{2}{3}$. If instead she reports that her preference is $a \succ_1' c \succ_1' b \succ_1' d$, then she and student 8 divide what is left of school $c$ between time $\frac{1}{3}$ and time $\frac{1}{2}$, so she receives $\frac{1}{3}a + \frac{1}{3}b + \frac{1}{6}c + \frac{1}{6}d$. In this example consumption of school $b$ by other students ceases before the school is fully allocated, which allows student 1 to defer its consumption.

We now consider a different example illustrating how strategy proofness can fail. Suppose that $a$ and $b$ are the agent's first and second favorite object, with $q_a = q_b = 1$, and there are $A-1$ other people who have $a$ as their favorite and $B-1$ other people who have $b$ as their favorite, where $1 < A < B$. Further, assume that no one outside the set of agents who have $a$ as their favorite will ever consume any $a$ and no one outside the set of agents who have $b$ as their favorite will ever consume any $b$. If the agent reports the truth she will receive $\frac{1}{A}$ units of $a$ and none of $b$. If she reports that $b$ is her favorite and $a$ is her second favorite, then she will consume $b$ between time 0 and time $\frac{1}{B}$ while $\frac{A-1}{B}$ units of $A$ are being consumed by others, and then she will receive $\frac{1}{A}(1 - \frac{A-1}{B})$ units of $a$, so her total consumption of $a$ and $b$ will be $\frac{1}{A}(1 + \frac{1}{B})$. This can be an improvement if the utility difference between $a$ and the agent's third favorite object is more than $A$ times the utility difference between $a$ and $b$.

This example suggests that, in general, the benefit of manipulatively consuming an inferior object (to change the later availability schedule of other objects) will be small in comparison with the amount of manipulation if there are many agents competing for the objects. In the special case of our general model in which $E$ is integral and $g_{io} = r_i$ (in effect) for all $i$ and $o$, Kojima and Manea (2010) (henceforth KM) establish an exact result along these lines: for a given utility function $u_i$ consistent with a preference $\succ_i$, if $\min_o q_o/r_i$ is sufficiently large, then agent $i$ will not be able to increase the expected utility from the probabilistic serial mechanism by reporting a different preference $\succ_i'$.

Following KM we present a result that shows that for a given vNM utility function, if, for each student, the number of students with the same preferences and opportunities is large, relative to the ratio of the largest utility difference to the smallest utility difference, then truthful reporting is a weakly dominant strategy.

**Theorem 5.** Let $E$ be an integral school choice CEE that satisfies the GMC, and let $\succ = (\succ_i)_{i \in I}$ be a preference profile. Let $\succ_i'$ be an alternative preferences for some $i \in I$, and let $\succ' = (\succ_i', \succ_{-i})$. Let $N_0 = |\{\, j \in I : \alpha_j = \alpha_i \text{ and } \succ_j = \succ_i \,\}|$. Let $u_i \colon A \to \mathbb{R}$ be a cardinal utility function consistent with $\succ_i$, and let

$$d_i = \min_{o \succ_i p} u_i(o) - u_i(p) \quad \text{and} \quad D_i = \max_{o \succ_i p} u_i(o) - u_i(p).$$

If $d_i/D_i \geq \left(\frac{N_0+1}{N_0}\right)^{|O|} - 1$, then $u_i(GCPS(\succ)) \geq u_i(GCPS(\succ'))$.

The proof of this result has two phases. In the first it is shown that the effect of the manipulation by a student $i$ on the overall allocation is bounded by the amount that $i$'s own consumption differs between the eating schedule induced by the true preference and eating schedule induced by the reported preference.

The second phase bounds the benefits of the periods of time during which the student can eat from a school that would not be available if the student reported her true preference. The additional amount of the school that the student consumes during such a period is the amount that is available at the beginning of the period divided by the number of students eating from this school. In the KM setting the set of agents eating an object type is weakly increasing while the object type is available, so having a large number of objects of each type implies that if an object type is fully consumed, then the final rate of consumption is high. In our setting the number of agents eating from a school can decrease when sets of schools become critical, so we need an additional assumption to insure that the number of agents competing for each school is high. The simplest way to insure this is to assume that there are many students with the same opportunities and preferences as the manipulator.

We regard Theorem 5, and especially its proof, as illustrative of the difficulties of reordering manipulation, rather than as a complete explanation of them. For one thing, a key point is that in order to manipulate successfully, the manipulator must believe that during the time when a school is available due to the manipulation, there will be scant competition. While this is certainly not the case under our assumption, there are many other reasons that competition for the school might be expected. It will be evident that noticing an opportunity to manipulate typically requires much more information than a student is likely to possess.

We have seen that manipulation by augmentation is impossible. Manipulation by truncation is sometimes possible, and less frequently entails little risk, but it does little to change the incentives of other students, so (in contrast with the Boston mechanism) it does not lead to Nash equilibria that are drastically different from truthful revelation. Manipulation by reordering has large costs and low rewards when there are many agents for each object, which is typically the case for school choice. On the whole, failures of strategy proofness are minor and do little to impair the practical applicability of the GCPS mechanism to school choice.

# 9 Concluding Remarks

We have provided a school choice mechanism that is a specialization of the GCPS mechanism of Balbuzanov (2022), which is in turn a generalization of the PS mechanism of BM. This mechanism guarantees each student a seat in a school that is at least as desirable as any of the schools she is legally entitled to attend. When there are many students for each school, it is effectively strategy proof. It is $sd$-efficient, which (as BM stress) is a stronger condition than ex post efficiency. In contrast, DA based on randomly generated priorities for the schools is (at least in its most basic form) not even ex post efficient. The GCPS mechanism is implementable: the assignment probabilities it generates can be obtained from a randomization over pure assignments. It satisfies anonymity, equal treatment of equals, and a natural generalization of the envy-freeness condition satisfied by the PS mechanism. Using a novel generalization of Hall's marriage theorem, we have described a computational implementation of this mechanism that seems to be tractable even for quite large school choice problems.

A possibility we intend to explore in subsequent research is that instead of consuming probability of desirable objects, the agents may discard probability of undesirable objects. In the case of $n$ agents and $n$ objects, each agent is endowed with one unit of each object, and at each time during the interval $[0, n-1]$ she discards probability of the least desirable object that she has not fully discarded for which discarding is still allowed. Discarding of an object is disallowed when the agents' total remaining endowment of it is one, but it may also be disallowed for some agents in the event that the process reaches a facet of $R$. The characterization of the PS mechanism given by Bogomolnaia and Heo (2012) implies that the discarding mechanism is certainly different from the probabilistic serial mechanism, but otherwise its properties await investigation. It seems appropriate for problems, perhaps such as chore assignment, in which the agents' main concern is to avoid the objects that are most noxious for them.

A possibility stressed by BM, Cho (2018), and Balbuzanov (2022) (perhaps among others) is that the PS mechanism can be varied by making the eating speeds depend on various things. This seems unmotivated in school choice, but in other domains it may be quite interesting. In particular, in chore assignment some agents may be unqualified to receive certain objects, and one may recognize this by taking away their endowments of such objects at the outset, but this seems unfair insofar it amounts to giving them a head start. Giving such agents slower discarding speeds is one way this issue could be addressed.

Although we have emphasized the school choice application, we expect that the underlying idea of our procedure, the application of the GCPS mechanism to a CEE, is potentially of interest in many other domains, with many variations.

# References


Abdulkadiroğlu, A. and Andersson, T. (2022), School choice. NBER Working Paper 29822.

Abdulkadiroğlu, A., Che, Y.-K. and Yasuda, Y. (2011), 'Resolving conflicting preferences in school choice: The Boston mechanism reconsidered', *Am. Econ. Rev.* **101**, 399–410.

Abdulkadiroğlu, A., Che, Y.-K. and Yasuda, Y. (2015), 'Expanding choice in school choice', *Am. Econ. J-Microecon* **7**, 1–42.

Abdulkadiroğlu, A., Pathak, P. A. and Roth, A. E. (2009), 'Strategy-proofness versus efficiency with indifferences: Redesigning the NYC high school match', *Am. Econ. Rev.* **99**, 1954–1978.

Abdulkadiroğlu, A. and Sönmez, T. (2003), 'School choice: a mechanism design approach', *Am. Econ. Rev.* **93**, 729–747.

Ahuja, R. K., Magnanti, T. L. and Orlin, J. B. (1993), *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall.

Akbarpour, M. and Nikzad, A. (2020), 'Approximate random allocation mechanisms', *Rev. Econ. Stud* **87**, 2473–2510.

Athanassoglou, S. and Sethuraman, J. (2011), 'House allocation with fractional endowments', *Int. J. Game Theory* **40**, 481–513.

Azevedo, E. M. and Budish, E. (2019), 'Strategy-proofness in the large', *Rev. Econ. Stud.* **86**, 81–116.

Azevedo, E. M. and Leshno, J. D. (2016), 'A supply and demand framework for two-sided matching markets', *J. Polit. Econ.* **124**, 1235–1268.

Balbuzanov, I. (2022), 'Constrained random matching', *J. Econ. Theory* **203**. Article 105472.

Bogomolnaia, A. (2015), 'Random assignment: Redefining the serial rule', *J. Econ. Theory* **158**, 308–315.

Bogomolnaia, A. and Heo, E. J. (2012), 'Probabilistic assignment of objects: Characterizing the serial rule', *J. Econ. Theory* **147**(5), 2072–2082.

Bogomolnaia, A. and Moulin, H. (2001), 'A new solution to the random assignment problem', *J. Econ. Theory* **100**(2), 295–328.

Bogomolnaia, A. and Moulin, H. (2004), 'Random matching under dichotomous preferences', *Econometrica* **72**(1), 257–279.

Budish, E., Che, Y.-K., Kojima, F. and Milgrom, P. (2013), 'Designing random allocation mechanisms: Theory and practice', *Amer. Econ. Rev.* **103**, 585–623.

Burke, M. A. and Sass, T. R. (2013), 'Classroom peer effects and student achievement', *J. Labor Econ.* **31**, 51–82.

Calsamiglia, C., Haeringer, G. and Klijn, F. (2010), 'Constrained school choice: An experimental study', *Am. Econ. Rev.* **100**, 1860–1874.

Chen, L., Kyng, R., Liu, Y., Gutenberg, M. and Sachdeva, S. (2022), 'Maximum flow and minimum-cost flow in almost-linear time', *arXiv preprint* .

Cho, W. J. (2016), 'Incentive properties for ordinal mechanisms', *Game. Econ. Behav.* **95**, 168–177.

Cho, W. J. (2018), 'Probabilistic assignment: an extension approach', *Soc. Choice Welfare* **51**(1), 137–162.

Cho, W. J. and Doğan, B. (2016), 'Equivalence of efficiency notions for ordinal assignment problems', *Econ. Lett.* **146**, 8–12.

Echenique, F., Miralles, A. and Zhang, J. (2021), 'Constrained pseudo-market equilibrium', *Am. Econ. Rev.* **111**, 3699–3732.

Ehlers, L., Hafalir, I. E., Yenmez, M. B. and Yildirim, M. A. (2014), 'School choice with controlled choice constraints: Hard bounds versus soft bounds', *J. Econ. Theory* **153**, 648–683.

Epple, D. and Romano, R. E. (2011), Peer effects in education: a survey of theory and evidence, *in* 'Handbook of Social Economics', Vol. 1, Elsevier, pp. 1053–1163.

Erdil, A. and Ergin, H. (2008), 'What's the matter with tie breaking? Improving efficiency in school choice', *Am. Econ. Rev.* **98**, 669–689.

Featherstone, C. R. and Niederle, M. (2008), Ex ante efficiency in school choice mechanisms: An experimental investigation. National Bureau of Economic Research (NBER) Working Paper 14618.

Ford, L. and Fulkerson, D. (1956), 'Maximal flow through a network', *Canadian J. Math.* **8**, 399–404.

Gale, D. (1957), 'A theorem on flows in networks', *Pacific J. Math.* **7**, 1073–1082.

Gale, D. and Shapley, L. (1962), 'College admissions and the stability of marriage', *Am. Math. Mon.* **65**, 9–15.

Goldberg, A. V. and Tarjan, R. E. (1988), 'A new approach to the maximum flow problem', *J. ACM* **35**, 921–940.

Grigoryan, A. (2023), ‘Top trading cycles with reordering: improving match priority in school choice’, *Soc. Choice Welfare* **60**, 383–396.

Haeringer, G. and Klijn, F. (2009), ‘Constrained school choice’, *J. Econ. Theory* **144**, 1921–1947.

Hafalir, I. E., Yenmez, M. B. and Yildirim, M. A. (2013), ‘Effective affirmative action in school choice’, *Theor. Econ.* **8**, 325–363.

Hakimov, R. and Kesten, O. (2018), ‘The equitable top trading cycles mechanism for school choice’, *Int. Econ. Rev.* **59**, 2219–2258.

Kamada, Y. and Kojima, F. (2015), ‘Efficient matching under distributional constraints: Theory and applications’, *Amer. Econ. Rev.* **105**, 67–99.

Kamada, Y. and Kojima, F. (2017), ‘Stability concepts in matching under distributional constraints’, *J. Econ. Theory* **168**, 107–142.

Katta, A.-K. and Sethuraman, J. (2006), ‘A solution to the random assignment problem on the full preference domain’, *J. Econ. Theory* **131**, 231–250.

Kesten, O. (2006), ‘On two competing mechanisms for priority-based allocation problems’, *J. Econ. Theory* **127**, 155–171.

Kesten, O. (2010), ‘School choice with consent’, *Q. J. Econ.* **127**, 1297–1346.

Kesten, O. and Ünver, M. U. (2015), ‘A theory of school-choice lotteries’, *Theor. Econ.* **10**, 543–595.

Kojima, F. (2009), ‘Random assignment of multiple indivisible objects’, *Math. Soc. Sci.* **57**, 134–142.

Kojima, F. (2012), ‘School choice: Impossibilities for affirmative action’, *Games Econ. Behav.* **75**, 685–693.

Kojima, F. and Manea, M. (2010), ‘Incentives in the probabilistic serial mechanism’, *J. Econ. Theory* **145**, 293–314.

Leshno, J. D. and Lo, I. (2020), The cutoff structure of top trading cycles in school choice. Columbia University Working Paper.

Miralles, A. (2009), School choice: The case for the Boston mechanism, *in* ‘Auctions, Market Mechanisms, and Their Applications 2009’, pp. 58–60.

Morrill, T. (2015), 'Two simple variations of top trading cycles', *Theor. Econ.* **60**, 123–140.

Ore, Ø. (1962), *Theory of Graphs*, American Mathematical Society Colloquium Publications, American Mathematical Society.

Pathak, P. (2006), Lotteries in school assignment. Harvard University.

Pathak, P. A. (2017), What really matters in designing school choice mechanisms, *in* 'Advances in Economics and Econometrics', Cambridge University Press, Cambridge, pp. 176–214.

Reichmeider, P. F. (1978), The equivalence of some combinatorial matching theorems. Ph.D. thesis, Adelphi University.

Reichmeider, P. F. (1985), *The Equivalence of Some Combinatorial Matching Theorems*, Polygonal Publishing House, Washington.

Romm, A., Roth, A. E. and Shorrer, R. I. (2020), Stability vs. no justified envy. draft, Stanford University.

Roth, A. E. and Rothblum, U. G. (1999), 'Truncation strategies in matching markets–in search of advice for participants', *Econometrica* **67**, 21–43.

Saban, D. and Sethuraman, J. (2014), 'A note on object allocation under lexicographic preferences', *J. Math. Econ.* **50**, 283–289.

Sacerdote, B. (2011), Peer effects in education: How might they work, how big are they and how much do we know thus far?, *in* 'Handbook of the Economics of Education', Vol. 3, Elsevier, pp. 249–277.

Schulman, L. J. and Vazirani, V. V. (2015), 'Allocation of divisible goods under lexicographic preferences', *arXiv preprint* .

Shapley, L. S. (1961), 'On network flow functions', *Naval Res. Logist. Q.* **8**, 151–158.

Shapley, L. S. and Scarf, H. (1974), 'On cores and indivisibility', **1**, 23–38.

Tang, Q. and Yu, J. (2014), 'A new perspective on Kesten's school with consent idea', *J. Econ. Theory* **154**, 543–561.

Troyan, P. (2012), 'Comparing schoool choice mechanisms by interim and ex-ante welfare', *Games Econ. Behav.* **75**, 936–947.

Vigdor, J. and Nechyba, T. (2007), Peer effects in North Carolina public schools, *in* 'Schools and the Equal Opportunity Problem', The MIT Press, pp. 73–101.

Wang, T. and Zhou, C. (2020), 'High school admission reform in China: a welfare analysis', *Rev. Econ. Design* **24**, 215–269.

Yilmaz, Ö. (2009), 'Random assignment under weak preferences', *Game. Econ. Behav.* **66**, 546–558.

Yilmaz, Ö. (2010), 'The probabilistic serial mechanism with private endowments', *Game. Econ. Behav.* **69**, 475–491.

# For Online Publication

## A Implementability

In this Appendix we consider the problem of passing from a matrix of assignment probabilities to a random deterministic assignment whose distribution realizes the given probabilities, showing that this is possible, and describing an algorithm for this task.

Let $E = (I, O, 1, q, g)$ be an integral school choice CEE that satisfies the GMC, let $Q$ be its set of feasible allocations, and let $m$ be an element of $Q$. Budish et al. (2013) say that such an $m$ is *implementable* if the assignment probabilities are those resulting from some probability distribution over deterministic assignments[12]. Recalling that the vertices of $Q$ are its extreme points, we see that in order for every element of $Q$ to be implementable, each of its vertices must be a deterministic assignment, which is to say that its entries are elements of $\{0, 1\}$. Conversely, since $Q$ is the set of convex combinations of its vertices, if each vertex is a deterministic assignment, then every element of $Q$ is implementable.

As we explain in detail below, Theorem 1 of Budish et el. has the following result as a special case.

**Theorem 6.** Each vertex of $Q$ is integral.

The Birkhoff-von Neumann theorem asserts that if $|I| = |O|$, then the set of bistochastic matrices with entries indexed by $I \times O$ is the convex hull of the set of bistochastic matrices with entries in $\{0, 1\}$. Evidently the Birkhoff-von Neumann theorem is a special case of Theorem 6.

[12]Recently Akbarpour and Nikzad (2020) expanded the scope of this concept by studying a notion of approximate implementation that is appropriate when some constraints need not be satisfied exactly.

We quickly review the related concepts and results of Budish et al. (2013). A *constraint set* is a nonempty subset of $I \times O$, and a *constraint structure* $\mathcal{H}$ is a set of constraint sets. A vector of quotas $\mathbf{q} = (q_S, q^S)_{S \in \mathcal{H}}$ is *integral* if $q_S, q^S \in \mathbb{Z}$ for all $S$. An allocation $m$ is *feasible* under $\mathbf{q}$ if $q_S \leq \sum_{io \in S} m_{io} \leq q^S$ for all $S \in \mathcal{H}$. Let $\mathcal{M}_{\mathbf{q}}$ be the set of feasible allocations for $\mathbf{q}$. If $\mathcal{H}$ contains all singletons, then $\mathcal{M}_{\mathbf{q}}$ is bounded, hence a polytope. The constraint structure $\mathcal{H}$ is *universally implementable* if, whenever $\mathbf{q}$ is integral, each vertex of $\mathcal{M}_{\mathbf{q}}$ is integral. A constraint structure is a *hierarchy* if, for all $S, S' \in \mathcal{H}$, we have $S \subset S'$ or $S' \subset S$ or $S \cap S' = \emptyset$, and $\mathcal{H}$ is a *bihierarchy* if there are hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$ such that $\mathcal{H}_1 \cup \mathcal{H}_2 = \mathcal{H}$ and $\mathcal{H}_1 \cap \mathcal{H}_2 = \emptyset$. Theorem 1 of Budish et al. (2013) (which is also a generalization of the Birkhoff-von Neumann theorem) asserts that if $\mathcal{H}$ is a bihierarchy, then it is universally implementable. (Their Theorem 2 is a partial converse, giving conditions under which if $\mathcal{H}$ is universally implementable, then it is a bihierarchy.)

Let $\mathcal{H}^1 = \{\, \{i\} \times O : i \in I \,\}$, $\mathcal{H}^2 = \{\, \{(i,o)\} : (i,o) \in I \times O \,\}$, and $\mathcal{H}^3 = \{\, I \times \{o\} : o \in O \,\}$, corresponding to the constraints $\sum_o m_{io} = 1$, $0 \leq m_{io} \leq g_{io}$, and $\sum_i m_{io} \leq q_o$ respectively. We can show that $\mathcal{H} = \mathcal{H}^1 \cup \mathcal{H}^2 \cup \mathcal{H}^3$ is a bihierarchy either by setting $\mathcal{H}_1 = \mathcal{H}^1 \cup \mathcal{H}^2$ and $\mathcal{H}_2 = \mathcal{H}^3$ or by setting $\mathcal{H}_1 = \mathcal{H}^1$ and $\mathcal{H}_2 = \mathcal{H}^2 \cup \mathcal{H}^3$, so our Theorem 6 follows from their Theorem 1.

The practical implementation of a random allocation depends not only on the existence of a representation of it as a convex combination of pure allocations, but also on an efficient algorithm for generating a random pure allocation with a probability distribution that averages to the given allocation. To this end we describe the argument in Appendix B of the Online Appendices of Budish et al., which they attribute to Tomomi Matsui and Akihisa Tamura, as it applies to our setting.

We work with the network $(N_E, A_E)$ of Section 2. If $m \in Q$, the *nonintegrality set* of $m$ is $C(m) \cup D(m) \subset A_E$ where $C(m) = \{\, (i,o) \in I \times O : m_{io} \notin \mathbb{Z} \,\}$ and $D(m) = \{\, (o,t) \in O \times \{t\} : \sum_i m_{io} \notin \mathbb{Z} \,\}$. The next result implies that points of $Q$ that are not integral are not extreme points of $Q$, hence not vertices, so Theorem 6 follows.

Recall that the *floor* of a real number $x$ is the largest integer that is not greater than $x$, and the *ceiling* of $x$ is the smallest integer that is not less than $x$. When $x$ is an integer, it is both the floor and ceiling of itself.

**Proposition 4.** If $E$ is integral, $m \in Q$, and the nonintegrality set of $m$ is nonempty, then there are $m^0, m^1 \in Q \setminus \{m\}$ such that $m$ is a convex combination of $m^0$ and $m^1$, and for both $h = 0, 1$:

(a) For each $i$ and $o$, $m^h_{io}$ is between the floor and the ceiling of $m_{io}$.

(b) For each $o$, $\sum_i m^h_{io}$ is between the floor and the ceiling of $\sum_i m_{io}$.

(c) The nonintegrality set of $m^h$ is a proper subset of the nonintegrality set of $m$.

*Proof.* For each $i$, if there is an $o \in O$ such that $(i,o) \in C(m)$, then (since $\sum_o m_{io} = r_i \in \mathbb{Z}$) there are at least two such $o$. For each $o$, if there is exactly one $i$ such that $(i,o) \in C(m)$, then $\sum_i m_{io} \notin \mathbb{Z}$ and consequently $(o,t) \in D(m)$. Since $\sum_o \sum_i m_{io} = \sum_i r_i \in \mathbb{Z}$, there cannot be exactly one $o$ such that $(o,t) \in D(m)$.

An *allowed cycle* is a sequence $n_1, \ldots, n_h$ of $h > 2$ distinct nodes in $I \cup O \cup \{t\}$ such that for all $g = 1, \ldots, h$ either $(n_g, n_{g+1}) \in C(m) \cup D(m)$ or $(n_{g+1}, n_g) \in C(m) \cup D(m)$. (The indices are integers mod $h$.) By hypothesis there are $n_1$ and $n_2$ such that $(n_1, n_2) \in C(m)$. If we have already chosen distinct $n_1, \ldots, n_g$ satisfying the required condition, then there is $n_{g+1} \neq n_{g-1}$ such that either $(n_g, n_{g+1}) \in C(m) \cup D(m)$ or $(n_{g+1}, n_g) \in C(m) \cup D(m)$. Since $N$ is finite, continuing in this fashion leads eventually to $n_{g+1} \in \{n_1, \ldots, n_{g-2}\}$, so this process eventually constructs an allowed cycle.

Let $n_1, \ldots, n_h$ be an allowed cycle. For each $i$ and $o$, if $(i,o) = (n_g, n_{g+1})$ ($(i,o) = (n_{g+1}, n_g)$) for some $g$, then we say that $(i,o)$ is a *forward* (*backward*) *arc*. For $\gamma \in \mathbb{R}$ let $m^\gamma \in \mathbb{R}^{I \times O}$ be the matrix with components

$$m^\gamma_{io} = \begin{cases} m_{io} + \gamma, & (i,o) \text{ is a forward arc,} \\ m_{io} - \gamma, & (i,o) \text{ is a backward arc,} \\ m_{io}, & \text{otherwise.} \end{cases}$$

Let $\alpha$ be the smallest positive number such that one of the following occurs:

(a) $m^\alpha_{io} \in \mathbb{Z}$ for some $(i,o) \in C(m)$.

(b) $\sum_i m^\alpha_{io} \in \mathbb{Z}$ for some $(o,t) \in D(m)$.

Let $\beta$ be the smallest positive number such that $m^{-\beta}$ satisfies one of these conditions. Let $m^0 = m^\alpha$ and $m^1 = m^{-\beta}$, so that $m = \frac{\beta}{\alpha+\beta} m^0 + \frac{\alpha}{\alpha+\beta} m^1$.

For each $i$ and $g$ such that $n_g = i$, $(i, n_{g-1})$ is a backward arc and $(i, n_{g+1})$ is a forward arc, so $\sum_o m^\gamma_{io} = \sum_o m_{io} = r_i$ for all $\gamma$. Since $E$ is integral, it follows that $m^0$ and $m^1$ satisfy all the constraints defining $Q$. It is now easy to see that $m^0$ and $m^1$ satisfy (a)–(c) of the statement. $\square$

To generate a random integral allocation whose expectation is the given $m$ we repeatedly execute the computation described in this argument, passing to $m^0$ with probability $\frac{\beta}{\alpha+\beta}$ and passing to $m^1$ with probability $\frac{\alpha}{\alpha+\beta}$. The number of times this must be done is bounded by the number of elements of the nonintegrality set of $m$, and the running time of each step is bounded by a the maximum size of a cycle, which is also the number of elements of the nonintegrality set at that step. Thus this algorithm's worst case complexity is bounded by a constant times the square of the number of elements of the nonintegrality set of $m$, which is at most $\sum_i |\alpha_i|$ in our intended application.

# B GCPS Schools

For the application to school choice, a version of the algorithm described in Section 4 has been encoded, using the `C` programming language, as an executable `gcps`, in the software package `GCPS Schools`, which can be downloaded[13]. `GCPS Schools` also contains two other executables, `make_ex` and `purify`. The second of these is a straightforward implementation of the algorithm of Budish et al. (2013) described in Online Appendix A, which passes from the output of `gcps` to a random deterministic allocation whose distribution induces the assignment probabilities computed by `gcps`.

The executable `make_ex` generates random school choice problems of the sort that might occur in large school districts. The schools and students are spaced evenly around a circle. Each student's safe school is the school that is closest to her home. Each school has a random *valence*, which is normally distributed, for each student-school pair there is a normally distributed *idiosyncratic shock*, and the student's utility for a seat in the school is the sum of these two quantities minus the distance between her home and the school. The schools that the student is eligible for are those that provide at least as much utility as the safe school, and the student's preference over such schools is the one induced by these utilities.

The computation of the GCPS allocation begins with a feasible allocation $\overline{p}(0) \in Q$. Computing such a point is equivalent to computing a maximal flow for the network $(N_E, A_E)$ with capacity $c_E$. We use the push-relabel algorithm of Goldberg and Tarjan (1988), specialized to $(N_E, A_E)$. (Fifteen algorithms for the maximal flow problem were already known at the time of that paper, but the literature continues to advance, e.g., Chen et al. (2022).)

Table 1 reports the result in which `gcps` was applied to two series of examples produced by `make_ex`. In the first series the number of schools is fixed at 10 while the number of seats per school increases from 10 to 100, while in the second series the

[13]Open the url `https://github.com/Coup3z-pixel/SchoolOfChoice/` in a web browser. Clicking on the file `gcps_schools.tar` opens a page for that file. Clicking the `raw` button on the line for the file downloads the file to your browser. After placing the file in a suitable directory, in a Unix command line terminal at that directory give the command `tar xvf gcps_schools.tar`. In the directory `gcps_schools` created by that command the document *GCPS_Schools_User_Guide.pdf* has further instuctions.

number of seats per school is fixed at 10 while the number of schools increases from 10 to 100. In all examples there are nine students for every ten seats, so in both series the number of students increases from 90 to 900.

| Schools | Seats/sch | Segments | Splits | Pivots | h-sum |
|---|---|---|---|---|---|
| 10 | 10 | 33 | 7 | 372 | 542 |
| 10 | 20 | 68 | 7 | 1596 | 2275 |
| 10 | 30 | 77 | 7 | 3468 | 4888 |
| 10 | 40 | 123 | 7 | 6780 | 10280 |
| 10 | 50 | 129 | 7 | 10224 | 14651 |
| 10 | 60 | 143 | 7 | 12885 | 17221 |
| 10 | 70 | 209 | 8 | 19782 | 28272 |
| 10 | 80 | 256 | 8 | 29553 | 43179 |
| 10 | 90 | 230 | 7 | 35178 | 54779 |
| 10 | 100 | 270 | 7 | 37227 | 58825 |
| | | | | | |
| 10 | 10 | 33 | 7 | 372 | 542 |
| 20 | 10 | 58 | 14 | 1140 | 1689 |
| 30 | 10 | 98 | 20 | 2499 | 4399 |
| 40 | 10 | 159 | 30 | 5594 | 9128 |
| 50 | 10 | 169 | 34 | 9538 | 14471 |
| 60 | 10 | 188 | 40 | 11253 | 16371 |
| 70 | 10 | 200 | 45 | 14260 | 21243 |
| 80 | 10 | 260 | 55 | 19289 | 28084 |
| 90 | 10 | 296 | 61 | 29952 | 38886 |
| 100 | 10 | 286 | 62 | 27033 | 40506 |

Table 1: numbers of events during `gcps` computations

In order to explain Table 1 we review the main features of the algorithm. The path $p$ of the GCPS allocation process, and the path $\overline{p}$ of the feasible allocation such that $p(t) \leq \overline{p}(t)$ for all $t$, are both piecewise linear, so the combined function $(p, \overline{p})$ is also piecewise linear. Having arrived at the values of these functions at a particular point in time, and having determined directions for $p$ and $\overline{p}$, the algorithm computes the amount of time that these directions can be followed before some constraint becomes binding, and it computes the parameters of the residual problem that results from following these directions for that amount of time. The arithmetic burden of one of these computations is proportional to the number of student-school pairs such that the school is possible for the student, i.e, the number of students times the average number of schools that are possible for a student. In both series of examples the number of linear pieces (the variable *Segments* in the table) is approximated, quite roughly, by the number of students divided by three. Thus the overall burden of computing the times at which segments end, and the new values of the parameters at the endpoints, seems to be proportional to the square of the number of students.

After following the trajectories of $p$ and $\overline{p}$ to the end of a segment at time $t$, the algorithm either computes a new trajectory for $\overline{p}$ that allows the trajectory of $p$ to continue, or it finds a critical pair $(P, J)$, and it descends recursively to the derived problems for $(E - p(t))_{(J,P)}$ and for $(E - p(t))^{(J,P)}$. If it finds such a pair, and $P$, $J$, $P^c$, and $J^c$ are all nonempty, then we say that the process *splits*. The combined complexity of the two subproblems is less than the complexity of the problem from which they are derived, so such events do not give rise to complexity concerns. The number of such events (throughout the recursive descent) is not greater than the number of schools minus one. From Table 1 we see that the actual number of such events is roughly two thirds of the number of schools.

In the computation of a critical pair $(P, J)$ or a new direction for $\overline{p}$ that allows $p$ to continue in the same direction, the computation starts with an $|I|\times|O|$ matrix of integers $\theta$. This matrix is modified repeatedly until it reaches a satisfactory state, and we say that each individual modification is a *pivot*. In Section 4 we explained that each pivot decreases a certain quantity by one, the matrix is satisfactory when this quantity is zero, and the initial value of this quantity cannot be greater than the number of students, so the number of pivots cannot exceed the number of students. The fifth column of Table 1 reports the total number of pivots. In the first sequence of experiments the number of pivots per segment increases from roughly 10 to roughly 160, and in the second sequence of experiment this number increases from roughly 10 to roughly 64. If the number of pivots per segment does not grow more rapidly that the number of students, and the number of segments is roughly proportional to the number of students, then the computational burden of computing pivots is roughly proportional to the square of the number of students times the average cost of a pivot.

A pivot that does not result in a critical pair decreases $\theta_{i_1o_0}, \theta_{i_2o_1}, \ldots, \theta_{i_ho_{h-1}}$ by one and increases $\theta_{i_1o_1}, \theta_{i_2o_2}, \ldots, \theta_{i_ho_h}$ by one, for some sequences $o_0, o_1, \ldots, o_h$ of schools and $i_1, \ldots, i_h$ of students. The quantity *h-sum* is the sum, across all pivots, of the integer $h$. In all cases it is between one and two times the number of pivots, so the average cost of a pivot does not seem to grow with the size of the problem.

Thus both the cost of computing endpoints of segments, and the cost of pivoting, seem to be at worst proportional to the square of the number of students. The example with 100 schools and 10 seats and 9 students per school has a running time of 0.7

seconds, which suggests that even for the world's largest school choice problem (New York City, with over 500 schools and 100,000 students) the algorithm should run to completion within a few hours.

As we mentioned at the end of Appendix A, the running time of the BCKM algorithm described there is bounded by a constant times the square of $\sum_i |\alpha_i|$, and thus by a constant times $|I|^2$ for any given bound on $|\alpha_i|$. We have not done systematic experiments on `purify`, but in our experience it runs very quickly, as one would expect if the cycles that it finds in the graph of nonintegral entries of the matrix $m$ are typically short. This intuition, and our experience, strongly suggest that the running time of `purify` will not be a factor that limits the applicability of the GCPS mechanism.

# C Proofs of Results in Sections 5–7

This appendix collects relatively short proofs.

*Proof of Theorem 2.* Gross substitutes implies that the lower and upper bounds of $I_o(c_{-o})$ are nondecreasing functions of each component of $c_{-o}$, so $T_o(c)$ is a nondecreasing function of each of these components, and it is also obviously a nondecreasing function of $c_o$. Thus $T$ is an increasing function: if $c \leq c'$, then $T(c) \leq T(c')$. Tarski's fixed point theorem implies that $\mathcal{F}$ is a complete sublattice of $[0, \overline{e}]^O$.

Fixing $c, c' \in \mathcal{F}$, let $c^+ = c \vee c'$. For an arbitrary $o \in O$ assume without loss that $c_o \leq c'_o$. By gross substitutes we have $D_o(c^+) \geq D_o(c')$. If $c'_o > 0$, then $D_o(c') = q_o \geq D_o(c)$, and if $c'_o = 0$, then $c_o = c'_o$ and (by symmetry) $D_o(c^+) \geq D_o(c)$. Thus $D_o(c^+) \geq \max\{D_o(c), D_o(c')\}$. This holds for all $o$, so

$$\sum_o D_o(c^+) \geq \sum_o \max\{D_o(c), D_o(c')\} \geq \max\Big\{\sum_o D_o(c), \sum_o D_o(c')\Big\}.$$

Since $c^+ \geq c$ we have $1 - \sum_o D_{io}(c^+) \geq 1 - \sum_o D_{io}(c)$ for each $i$, and thus $|I| - \sum_o D_o(c^+) \geq |I| - \sum_o D_o(c')$. By symmetry $|I| - \sum_o D_o(c^+) \geq |I| - \sum_o D_o(c)$, so

$$|I| - \sum_o D_o(c^+) \geq |I| - \max\Big\{\sum_o D_o(c), \sum_o D_o(c')\Big\}.$$

We now conclude that all the weak inequalities above are in fact equalities, and in particular $D_o(c) = D_o(c^+) = D_o(c')$ for all $o$.

Gross substitutes and the definition of demand give $D_{io}(c^+) \leq D_{io}(c^+_o, c_{-o}) \leq D_{io}(c)$ for each $i$ and $o$. For each $o$, since $D_o(c) = D_o(c^+)$, it follows that $D_{io}(c) = D_{io}(c^+)$ for all $i$. By symmetry $D_{io}(c') = D_{io}(c^+) = D_{io}(c)$. □

The following result is essentially due to Cho and Doğan (2016). (Lemma 3 of BM is a precursor.) We provide no proof because it is easy to see that their proof works, essentially without any modification, in our more general setting.

**Lemma 4.** $sd$-efficiency, $dl$-efficiency, and $ul$-efficiency are equivalent.

The following is a special case of the proof of Proposition 3 of Balbuzanov (2022).

*Proof of Theorem 3.* By the last result it suffices to show that the GCPS allocation $m = p(1)$ is $sd$-efficient. Aiming at a contradiction, suppose that $m' \in Q$, $m' \neq m$, and $m'_i \, sd(\succ_i) \, m_i$ for all $i$. Since $m' \, sd(\succ_i) \, m$ and $m' \neq m$ there is an agent $i$ and objects $o, o'$ such that $o' \succ_i o_0$ and $m'_{io'} > m_{io'}$, and there must have been a time at which $o'$ became unavailable to $i$.

Since the set of possible critical pairs is finite, we may assume that $t_0$ is the first time such that there is a $J \subset I$ and $P \subset O$ such that $p(t_0)$ satisfies (1) for $J$ and $P$ with equality and there is $i_0 \in J^c$ and $o_0 \in P$ such that $m'_{i_0 o_0} \neq m_{i_0 o_0}$. Since $m'$ satisfies (1) for $J$ and $P$ we have

$$\sum_{i \in J^c} \sum_{o \in P} m'_{io} \leq \sum_{i \in J^c} \sum_{o \in P} m_{io}.$$

Consequently there are $i_1 \in J^c$ and $o_1 \in P$ such that $m'_{i_1 o_1} < m_{i_1 o_1}$. Since $m' \, sd(\succ_i) \, m$ and $m' \neq m$ there is an object $o_2$ such that $o_2 \succ_{i_1} o_1$ and $m'_{i_1 o_2} > m_{i_1 o_2}$. Since $p_{i_1 o_1}(t_0) = m_{i_1 o_1} > m'_{i_1 o_1} \geq 0$, there must have been a time prior to $t_0$ at which $o_2$ became unavailable to $i_1$, but this contradicts the definition of $t_0$. □

*Proof of Proposition 3.* Suppose that $i, j \in I$, $\alpha_i \subset \alpha_j$, and $o_i \succ_i o_j$ for all $o_i \in \alpha_i$ and $o_j \in \alpha_j \setminus \alpha_i$. At each time during the allocation process, if there is a critical set of schools $P$ such that the only remaining schools that $i$ can consume are contained in $P$, but $j$ has some remaining school outside of $P$, then $j$ is consuming some element of $\alpha_j \setminus \alpha_i$. On the other hand, if at that time, for every critical $P$, either all the remaining schools for $j$ are contained in $P$ or there is some remaining school for $i$ that is outside $P$, then the set of remaining schools for $j$ that are contained in $\alpha_i$ is the set of remaining schools for $i$. In either case $i$ is consuming a school that she weakly prefers to the school that $j$ is consuming. Since this is true at all times during the allocation process, $GCPS_i(E, \succ) \, sd(\succ_i) \, GCPS_j(E, \succ)$. □

# D Eating Function Analysis

In this appendix we prove Theorems 4 and 5. Both proofs are based on a detailed analysis of the consequences of manipulation in terms of its effect on the continuous time eating process of BM, as generalized in KM and here. At this point we prepare both of the proofs by studying general aspects of this analysis.

We fix a school choice CEE $E = (I, O, 1, q, g)$ that satisfies the GMC and a profile $\succ = (\succ_j)_{j \in I}$ of strict preferences over $O$. Let $\Theta$ denote an artificial object (intuitively, eating nothing) that is not an element of $O$, and is available in infinite supply, and let $\tilde{O} = O \cup \{\Theta\}$. For each $j$ we extend $\succ_j$ to a strict preference over $\tilde{O}$ by specifying that $o \succ_j \Theta$ for all $o \in O$.

For $j \in I$ and $t \in [0, 1]$, an *eating schedule* on $[0, t)$ is a function $e_j \colon [0, t) \to \tilde{O}$ that is piecewise constant (i.e., changes objects finitely many times) and right continuous: for any $t' \in [0, t)$ there is an $\varepsilon > 0$ such that $e_j(t'') = e_j(t')$ for all $t'' \in [t', t' + \varepsilon)$. For such an $e_j$, $o \in \tilde{O}$, and $t' \in [0, t]$ let

$$p_{jo}(e_j, t') = \int_0^{t'} \mathbf{1}_{e_j(s)=o}\, ds.$$

An *eating function* on $[0, t)$ is a vector $e = (e_j)_{j \in I}$ of eating schedules on $[0, t)$. For $t' \in [0, t]$ let $p(e, t') \in \mathbb{R}_+^{I \times O}$ be the allocation with components $p_{jo}(e_j, t')$.

For $J \subset I$, $P \subset O$, and $t' \in [0, t]$ let

$$s_{(J,P)}(e, t') = \sum_{o \in P} q_o + \sum_{i \in J} \sum_{o \in P^c} g_{io} - \sum_{i \in J} r_i - \sum_{i \in J^c} \sum_{o \in P} p_{io}(e, t').$$

For $j \in I$, $o \in O$, $J \subset I$, and $P \subset O$ let

$$\tau_{jo}(e_j) = \sup\{\, t' : p_{jo}(e_j, t') < g_{jo} \,\} \quad \text{and} \quad \tau_{(J,P)}(e) = \sup\{\, t' : s_{(J,P)}(e, t') > 0 \,\}.$$

For $j \in I$ and $t' \in [0, t]$ let

$$\alpha_j(e, t') = \{\Theta\} \cup \alpha_j \setminus \Big( \{\, o : p_{jo}(e_j, t') \geq g_{io} \,\} \cup \bigcup_{J \subset I, P \subset O \,:\, s_{(J,P)}(e, t') \,\leq\, 0 \text{ and } j \,\in\, J^c} P \Big).$$

Note that $\alpha_j(e, \cdot)$ is right continuous. Let $e_j^{\succ}(e, t')$ be the $\succ_j$-best element of $\alpha_j(e, t)$. We say that $e_j$ is *myopic* for $e$ if, for all $t' \in [0, t)$, $e_j(t') = e_j^{\succ}(e, t')$.

We now fix a particular $i \in I$. We first show that any eating function for $i$ with $e_i(0) \in \alpha_i$ induces a well defined profile of myopic eating functions for the other agents, up to some time.

**Lemma 5.** For any $t \in [0, 1]$ and any eating schedule $e_i$ on $[0, t)$ for $i$, if $e_i(0) \in \alpha_i$, then there is a unique $\bar{t} \in (0, t]$ and unique eating schedules $e_j$ on $[0, \bar{t})$ for $j \neq i$ such that if $e_{-i} = (e_j)_{j \neq i}$ and $e = (e_i|_{[0,\bar{t})}, e_{-i})$, then:

(a) for all $t' \in [0, \bar{t})$, $e_i(t') \in \alpha_i(e, t')$;

(b) for each $j \neq i$, $e_j$ is myopic for $e$;

(c) either $\bar{t} = t$ or $e_i(\bar{t}) \notin \alpha_i(e, \bar{t})$.

*Proof.* For sufficiently small $\varepsilon > 0$, if, for each $j \neq i$, $e_j$ is the constant function on $[0, \varepsilon)$ with value $e_j^{\succ}(0)$, $e_{-i} = (e_j)_{j \neq i}$ and $e = (e|_{[0,\bar{t})}, e_{-i})$, then for all $t' \in [0, \varepsilon)$, $e_i(t') \in \alpha_i(e, t')$, and for each $j \neq i$, $e_j$ is myopic for $e$. Therefore there is a $\bar{t} \in (0, t]$ and a vector of eating schedules $e_{-i}$ on $[0, \bar{t})$ for $j \neq i$ such that if $e = (e_i|_{[0,\bar{t})}, e_{-i})$, then for each $j \neq i$, $e_j$ is myopic for $e$.

Suppose that $e'_{-i}$ is also a vector of eating schedules on $[0, \bar{t})$ for $j \neq i$ such that if $e' = (e_i|_{[0,\bar{t})}, e'_{-i})$, then for each $j \neq i$, $e'_j$ is myopic for $e'$. For each $j \neq i$ we have $e_j(0) = e_j^{\succ}(e, 0) = e_j^{\succ}(e', 0) = e'_j(0)$, so $e_{-i}$ and $e'_{-i}$ agree on the degenerate interval $[0, 0]$. If $\hat{t} \in [0, \bar{t})$ and $e_{-i}$ and $e'_{-i}$ agree on $[0, \hat{t}]$, then $\alpha(e, \hat{t}) = \alpha_j(e', \hat{t})$ for all $j \in I$, so for some $\varepsilon > 0$ they agree on $[0, \hat{t} + \varepsilon)$. Therefore, for the given $\bar{t}$, the vector $e_{-i}$ is unique.

If $\bar{t} < t$ and $e_i(\bar{t}) \in \alpha_i(e, \bar{t})$, then for some $\varepsilon > 0$ we can extend $e_{-i}$ to $[0, \bar{t} + \varepsilon)$ by setting $e_j(t') = e_j^{\succ_j}(e, \bar{t})$ for all $t' \in [\bar{t}, \bar{t} + \varepsilon)$, and each extended $e_j$ will be myopic for the extended $e$. It follows that there is a unique maximal $\bar{t}$, which satisfies (c). □

In the circumstance described in the last result we say that the profile $e_{-i}$ of eating schedules on $[0, \bar{t})$ is *induced* by $e_i$. An eating schedule $e_i \colon [0, t) \to \tilde{O}$ for $i$ is *feasible* if the time $\bar{t}$ of the last result is $t$.

Arguments similar to those used to prove Lemma 5 imply the next two results, so we do not provide proofs.

**Lemma 6.** There is a unique eating function $e$ on $[0, 1)$ such that each $e_j$ is myopic for $e$ on $[0, 1)$.

**Lemma 7.** Suppose that $e_i \colon [0, 1) \to \tilde{O}$ and $\bar{e}_i \colon [0, 1) \to \tilde{O}$ are eating schedules for $i$ such that $\{\, t : e_i(t) \neq \bar{e}_i(t) \,\} \subset \{\, t : \bar{e}_i(t) = \Theta \,\}$. If $e_i$ is feasible, then $\bar{e}_i$ is feasible.

We now fix particular objects $o^*, o^{**} \in \tilde{O}$. We begin with an eating schedule $e_i^{\rho_0}$ on $[0, 1)$ and an open interval $\mathcal{I}(\rho_0, \varepsilon) = (\rho_0 - \varepsilon, \rho_0 + \varepsilon) \subset (0, 1)$ such that $e_i^{\rho_0}(t) = o^*$ for $t \in (\rho_0 - \varepsilon, \rho_0)$ and $e_i^{\rho_0}(t) = o^{**}$ for $t \in [\rho_0, \rho_0 + \varepsilon)$. For each $\rho \in \mathcal{I}(\rho_0, \varepsilon)$ let $e_i^{\rho}$ be an eating schedule on $[0, \bar{t})$ such that $e_i^{\rho}$ agrees with $e_i^{\rho_0}$ on $[0, \rho_0 - \varepsilon]$, $e_i^{\rho}(t) = o^*$ for $t \in (\rho_0 - \varepsilon, \rho)$, and $e_i^{\rho}(t) = o^{**}$ for $t \in [\rho, \rho_0 + \varepsilon)$. We assume that $e_i^{\rho}$ is feasible for all $\rho$, we let $e_{-i}^{\rho}$ be the profile of eating schedules given by Lemma 5, and we set $e^{\rho} = (e_i^{\rho}, e_{-i}^{\rho})$.

We study two cases, which are relevant to Theorems 4 and 5 respectively. In the first $e^\rho$ is the eating function given by Lemma 6 for $E^\rho = (I, O, 1, q, g^\rho)$ and $\succ$ where $g^\rho$ differs from $g$ in the component for $i$ and $o^*$ in such a way that $p_{io^*}(e_i^\rho, \rho) = g^\rho_{io^*}$. We call this the *myopic case*.

In the second case $o^{**} = \Theta$ and each $e_i^\rho$ agrees with $e_i^{\rho_0}$ everywhere outside of $\mathcal{I}(\rho_0, \varepsilon)$. We call this the *constant case*. Note that in both cases, after consumption of a school $o$ by $j \neq i$ ceases, it does not resume later, and in the myopic case, after consumption of a school $o$ by $i$ ceases, it does not resume later.

If the times at which $j$ begins and finishes consuming $o$ are piecewise linear functions of $\rho$, then $\tau_{jo}(e^\rho)$ is a piecewise linear function of $\rho$. Similarly, if the times at which elements of $J^c$ start consuming elements of $P$, and the times prior to $\tau_{(J,P)}(e^\rho)$ at which they stop consuming them, are piecewise linear functions of $\rho$, then $\tau_{(J,P)}(e^\rho)$ is a piecewise linear function. By induction over increasing start and stop times, these are all piecewise linear functions.

We say that $\rho_0 \in (0,1)$ is *generic* if (possibly after replacing $\varepsilon$ with a smaller number) there are affine functions

$$t_0, t_1, \ldots, t_K \colon \mathcal{I}(\rho_0, \varepsilon) \to [0,1]$$

such that $0 \equiv t_0 < t_1 < \cdots < t_K \equiv 1$ and for each $j$ and $k = 1, \ldots, K$ there is an $o_{jk} \in \tilde{O}$ such that $e_j^\rho(t) = o_{jk}$ for all $\rho \in \mathcal{I}(\rho_0, \varepsilon)$ and $t \in [t_{k-1}(\rho), t_k(\rho))$. An interval of possible values of $\rho$ under consideration is partitioned into finitely many nongeneric values and finitely many open intervals whose elements are generic.

From now to the beginning of Subsection D.1 we assume that $\rho_0$ is generic. For each $k$ let $\sigma_k$ be the number such that

$$t_k(\rho) = t_k(\rho_0) + \sigma_k(\rho - \rho_0)$$

for all $\rho \in \mathcal{I}(\rho_0, \varepsilon)$. For each $j$ and $k = 1, \ldots, K$ there is a number $\kappa_{j,k-1}$ such that

$$p_{jo_{jk}}(e^\rho, t) = p_{jo_{jk}}(e^{\rho_0}, t) + \kappa_{j,k-1}(\rho - \rho_0)$$

for all $\rho \in \mathcal{I}(\rho_0, \varepsilon)$ and $t \in [t_{k-1}(\rho), t_k(\rho)) \cap [t_{k-1}(\rho_0), t_k(\rho_0))$.

Let $k_0$ be the integer such that $t_{k_0}$ is the identity function of $\mathcal{I}(\rho_0, \varepsilon)$. If $k < k_0$, then $\sigma_k = 0$ and $\kappa_{jk} = 0$ for all $j$. Clearly $\sigma_{k_0} = 1$, $\kappa_{ik_0} = -1$, and $\kappa_{jk_0} = 0$ for all $j \neq i$. If $o_{j,k+1} \neq o_{jk}$ and either $j \neq i$, or $j = i$ and we are in the myopic case, then (since consumption of an object does not resume after ceasing) we have $p_{jo_{j,k+1}}(e_j^\rho, t_k(\rho)) = 0$, so that $\kappa_{jk} = -\sigma_k$. In the constant case, if $o_{i,k+1} \neq o_{ik}$ and

$k \neq k_0$, then $\kappa_{jk} = 0 = -\sigma_k$.

The definition of genericity implies that:

(a) for each $j$ and $o$, if $\tau_{jo}(e_j^\rho) < 1$ for some $\rho \in \mathcal{I}(\rho_0, \varepsilon)$, then there is a $k$ such that $t_k(\rho) = \tau_{jo}(e_j^\rho)$ for all $\rho \in \mathcal{I}(\rho_0, \varepsilon)$;

(b) for each $J \subset I$ and $P \subset O$, if $\tau_{(J,P)}(e^\rho) < 1$ for some $\rho \in \mathcal{I}(\rho_0, \varepsilon)$, then there is a $k$ such that $t_k(\rho) = \tau_{(J,P)}(e^\rho)$ for all $\rho \in \mathcal{I}(\rho_0, \varepsilon)$.

If $t_k(\rho) = \tau_{jo}(e_j^\rho)$ for $\rho \in \mathcal{I}(\rho_0, \varepsilon)$, then $\sigma_k = -\kappa_{j,k-1}$, so that $\kappa_{jk} = \kappa_{j,k-1}$. Clearly $\kappa_{jk} = \kappa_{j,k-1}$ when $o_{j,k+1} = o_{jk}$.

For each $k = 1, \ldots, K$ let

$$\mathcal{P}_k = \{\, (J,P) : t_k(\rho) = \tau_{(J,P)}(e^\rho) \text{ for all } \rho \in \mathcal{I}(\rho_0, \varepsilon) \,\}.$$

For $(J,P) \in \mathcal{P}_k$ let $L_{k,(J,P)} = \{\, j \in J^c : o_{jk} \in P \,\}$. Let $k^*$ be the integer such that there is a $(J,P) \in \mathcal{P}_{k^*}$ such that $i \in J^c$ and $o^* \in P$, if such an integer exists. Let $k^{**}$ be the integer such that there is a $(J,P) \in \mathcal{P}_{k^{**}}$ such that $i \in J^c$ and $o^{**} \in P$, if such an integer exists. In what follows, the condition $k < k^*$ is to be understood as encompassing also the possibility that $k^*$ does not exist, and similarly for $k^{**}$.

**Lemma 8.** If $(J,P) \in \mathcal{P}_k$, then

$$\sigma_k = \frac{-1 - \sum_{j \in L_{k,(J,P)}} \kappa_{j,k-1}}{|L_{k,(J,P)}|}, \quad \sigma_k = \frac{1 - \sum_{j \in L_{k,(J,P)}} \kappa_{j,k-1}}{|L_{k,(J,P)}|}, \quad \text{or}$$

$$\sigma_k = \frac{- \sum_{j \in L_{k,(J,P)}} \kappa_{j,k-1}}{|L_{k,(J,P)}|},$$

according to whether $i \in J^c$ and $o^* \in P$, $i \in J^c$ and $o^{**} \in P$, or otherwise.

*Proof.* The proofs in the three cases are similar, and we give only the proof for the second case. The claim follows from the fact that the quantity

$$\sum_{j \in J^c} \sum_{o \in P} p_{jo}(e^\rho, t_k(\rho)) = \sum_{j \in J^c} \sum_{o \in P} p_{jo}(e^{\rho_0}, t_k(\rho_0)) + p_{io^*}(e^\rho, t_k(\rho)) - p_{io^*}(e^{\rho_0}, t_k(\rho_0))$$

$$+ \sum_{j \in L_{k,(J,P)} \setminus \{i\}} (\kappa_{j,k-1}(\rho - \rho_0) + t_k(\rho) - t_k(\rho_0))$$

$$= \sum_{j \in J^c} \sum_{o \in P} p_{jo}(e^{\rho_0}, t_k(\rho_0)) + (-1 + \sigma_k + \sum_{j \in L_{k,(J,P)}} \kappa_{j,k-1} + |L_{k,(J,P)} \setminus \{i\}|\sigma_k) \cdot (\rho - \rho_0)$$

does not depend on $\rho$. □

**Lemma 9.** $\sum_j \kappa_{jk^*} = \sum_j \kappa_{j,k^*-1} - 1$, $\sum_j \kappa_{jk^{**}} = \sum_j \kappa_{j,k^{**}-1} + 1$, and if $k \neq k^*, k^{**}$, then $\sum_j \kappa_{jk} = \sum_j \kappa_{j,k-1}$, .

*Proof.* We prove only the third assertion, since this clearly displays the ideas of the proofs of the other two assertions. Since the set of critical pairs is a lattice (Proposition 2) the set of $L_{k,(J,P)}$ such that $(J,P)$ is a minimal element of $\mathcal{P}_k$ is a partition of $\bigcup_{(J,P)\in\mathcal{P}_k} L_{k,(J,P)}$. For $j$ outside this union we have $\kappa_{jk} = \kappa_{j,k-1}$, either because $o_{jk} = o_{j,k-1}$ or because $t_k(\rho) = \tau_{jo}(e_j^\rho)$ for $\rho \in \mathcal{I}(\rho_0, \varepsilon)$. If $(J,P)$ is a minimal element of $\mathcal{P}_k$, then

$$\sum_{j\in L_{k,(J,P)}} \kappa_{jk} = -|L_{k,(J,P)}|\sigma_k = \sum_{j\in L_{k,(J,P)}} \kappa_{j,k-1}. \qquad \square$$

**Lemma 10.** If $k^* \leq k < k^{**}$, then $\kappa_{jk} \leq 0$ for all $j$, and $\sum_j \kappa_{jk} = -1$. If $k^{**} \leq k < k^*$, then $\kappa_{jk} \geq 0$ for all $j$, and $\sum_j \kappa_{jk} = 1$. If $k^* \leq k$ and $k^{**} \leq k$, then $\sum_j \kappa_{jk} = 0$, and if $I_k^- = \{\, j : \kappa_{jk} < 0\,\}$ and $I_k^+ = \{\, j : \kappa_{jk} > 0\,\}$, then $\sum_{j\in I_k^-} \kappa_{jk} \geq -1$ and $\sum_{j\in I_k^+} \kappa_{jk} \leq 1$.

*Proof.* The assertions concerning $\sum_j \kappa_{jk}$ follow from the last result by induction. The claims concerning the sign of $\kappa_{jk}$, and its bounds, also follow from induction on $k$, since if (for example) $k^* < k < k^{**}$, then for each $j$ we have either $\kappa_{jk} = \kappa_{j,k-1}$ (if $o_{jk} = o_{j,k-1}$ or $t_k(\rho) = \tau_{jo}(e_j^\rho)$ for $\rho \in \mathcal{I}(\rho_0, \varepsilon)$) or $\kappa_{jk} = -\sigma_k = \sum_{j'\in L_{k,(J,P)}} \kappa_{j',k-1}/|L_{k,(J,P)}|$ where $(J,P)$ is an element of $\mathcal{P}_k$ such that $j \in L_{k,(J,P)}$. When $k > k^*$ and $k > k^{**}$ this averaging cannot increase $-\sum_{j\in I_k^-} \kappa_{jk}$ or $\sum_{j\in I_k^+} \kappa_{jk}$, but it will decrease them if there is a minimal $(J,P) \in \mathcal{P}_k$ such that $\kappa_{j,k-1}$ is positive for some $j \in L_{k,(J,P)}$ and negative for others. $\square$

In preparation for the proof of Theorem 5 we mention that if $k < k^{**}$ and $(J,P) \in \mathcal{P}_k$, then

$$\tfrac{\partial}{\partial\rho}\tau_{(J,P)}(e^\rho) = \sigma_k \in [0,1] \quad \text{and} \quad \tfrac{\partial}{\partial\rho}s_{(J,P)}(e^\rho) = -\sum_{j\in L_{k,(J,P)}} \kappa_{j,k-1} \in [0,1]. \tag{2}$$

## D.1 The Proof of Theorem 4

For the given $i$, let $o^*$ be an element of $O \setminus \alpha_i$, let $\alpha_i' = \alpha_i \cup \{o^*\}$, and let $\succ_i'$ be a preference over $O$ that has $\alpha_i'$ as the set of schools weakly preferred to the safe school, and that agrees with $\succ_i$ on $\alpha_i$. We wish to show that the augmentation manipulation of reporting $\succ_i'$ rather than $\succ_i$ results in an allocation for $i$ that is weakly $sd(\succ_i)$ worse. Our method of analysis is to study how $i$'s allocation changes as the parameter $\rho = g_{io^*}$ varies continuously between 0 and 1. For $\rho \in [0,1]$ let $e^\rho$ be the eating function of the

GCPS mechanism when $g_{io^*} = \rho$. As we saw above, for each $o \in \alpha_i$, the probability that $i$ receives a seat in a school that is at least as good as $o$ is a piecewise linear function of $\rho$. Let $\mathcal{I}(\rho_0, \varepsilon) = (\rho_0 - \varepsilon, \rho_0 + \varepsilon)$ be an interval of generic values of $\rho$.

If $o \succ' o^*$, then $i$'s total consumption of schools preferred to $o$ is unaffected by $g_{io^*}$. For $\rho \in \mathcal{I}(\rho_0, \varepsilon)$ it may be the case that $i$ is excluded from consuming $o^*$ before $i$ has finished consuming $g_{io^*}$ units, and it is possible that $i$ does not finish consuming all $g_{io^*}$ units at time $1$. In both these cases small variations of $g_{io^*}$ do not change the allocation. Finally, if $o^* \succ_i' o$, then Lemma 10 implies that the total consumption of schools that are $\succ_i'$-weakly preferred to $o$ does not increase more rapidly than $g_{io^*}$ as this variable increases, so the total consumption of schools that are $\succ_i$-weakly preferred to $o$ does not increase as $g_{io^*}$ increases. Since this is the case for each of the finitely many generic intervals, it holds also for large increases, including the increase from $g_{io^*} = 0$ to $g_{io^*} = 1$.

## D.2 The Proof of Theorem 5

We now assume that $E$ is integral, so $g_{io} \in \{0, 1\}$ for all $i$ and $o$. Following KM, we will compare the eating schedule for $i$ resulting from truthful revelation with the eating schedule resulting from a manipulation by comparing both with the eating schedule that agrees with them at times when they agree with each other and consumes $\Theta$ during the times when they differ.

Let $\succ$ be a preference profile $\succ$ over $O$, let $\succ_i'$ be a deviant preference for $i$, and let $\succ' = (\succ_i', \succ_{-i})$. Lemma 6 gives unique eating functions $e^{\succ}$ and $e^{\succ'}$ that are generated by the GCPS procedure when agents report these preferences. Since $E$ satisfies the GMC, these eating functions are feasible. Let $\bar{e}_i$ be the eating schedule

$$\bar{e}_i(t) = \begin{cases} e_i^{\succ}(t), & \text{if } e_i^{\succ}(t) = e_i^{\succ'}(t), \\ \Theta, & \text{otherwise.} \end{cases}$$

Lemma 7 implies that $\bar{e}$ is feasible. Let $\bar{e}_{-i}$ be the profile of eating schedules induced by $\bar{e}_i$, and let $\bar{e} = (\bar{e}_i, \bar{e}_{-i})$. For $t \in [0, 1]$ let

$$\delta(t) = \int_0^t \mathbf{1}_{\bar{e}_i(s) \neq e_i(s)} ds.$$

**Lemma 11.** For all $(J, P)$ and $t \in [0, 1]$,

$$0 \le \tau_{(J,P)}(\bar{e}) - \tau_{(J,P)}(e) \le \delta(\tau_{(J,P)}(\bar{e})) \text{ and } 0 \le s_{(J,P)}(\bar{e}, t) - s_{(J,P)}(e, t) \le \delta(t).$$

*Proof.* In the obvious way we can create a path $\rho \mapsto e_i^\rho$ from $[0, \delta(1)]$ to the space of eating schedules, with $e_i^0 = e_i$ and $e_i^{\delta(1)} = \overline{e}_i$, that traverses each of the finitely many intervals in $[0,1]$ along which the value of $e_i$ is some $o^* \in O$ and the value of $\overline{e}_i$ is $\Theta$. Due to the piecewise linear nature of the problem, each of these intervals is a union of finitely many points and finitely many open "generic" intervals to which our earlier discussion in the case $k < k^{**}$ is applicable. The asserted inequalities are attained by integrating the inequalities (2) over these intervals. $\square$

Let
$$\beta(t) = \int_0^t \mathbf{1}_{e_i^{\succ'}(s) \succ_i e_i^{\succ}(s)} ds \quad \text{and} \quad \gamma(t) = \int_0^t \mathbf{1}_{e_i^{\succ}(s) \succ_i e_i^{\succ'}(s)} ds$$
be the sums of the lengths of time intervals, before time $t$, on which agent $i$'s consumption in the eating algorithm is $\succ_i$-better and $\succ_i$-worse, respectively, when the reported preferences change from $\succ$ to $\succ'$. Note that $\delta(t) = \beta(t) + \gamma(t)$.

Let $O' = \{o_1, o_2, \ldots, o_\ell\}$ be the set of objects $o$ such that for some time $t$, $o = e_i^{\succ'}(t)$ and $o \succ_i e_i^{\succ}(t)$, indexed so that $o_1 \succ_i' o_2 \succ_i' \cdots \succ_i' o_\ell$. For $l = 1, \ldots, \ell$ let
$$T_l = \inf\{\, t : e_i^{\succ'}(t) = o_l \text{ and } o_l \succ_i e_i^{\succ}(t) \,\}$$
be the first time $t$ when $o_l = e_i^{\succ'}(t)$ is $\succ_i$-preferred to $e_i^{\succ}(t)$. For each $l$ let
$$T_l' = \sup\{\, t : e_i^{\succ'}(t) = o_l \,\}$$
Clearly, $0 < T_1 < T_1' \leq T_2 < \cdots < T_{\ell-1}' \leq T_\ell < T_\ell' \leq 1$. Let $T_0 = 0$ and $T_{\ell+1} = 1$.

**Lemma 12.** For each $l = 1, \ldots, \ell$, $T_l' - T_l \leq \delta(T_l)/N_0$.

*Proof.* After time $T_l$ the object $o_l$ is not available to $i$ under the eating function $e^{\succ}$, so (because $E$ is integral) there is a pair $(J_l, P_l)$ such that $o_l \in P_l$, $i \in J_l^c$, and $\tau_{(J_l,P_l)}(e^{\succ}) = T_l$. Since $s_{(J_l,P_l)}(T_l, e^{\succ}) = 0$, Lemma 11 gives
$$s_{(J_l,P_l)}(T_l, e^{\succ'}) = s_{(J_l,P_l)}(T_l, e^{\succ'}) - s_{(J_l,P_l)}(T_l, e^{\succ})$$
$$= (s_{(J_l,P_l)}(T_l, e^{\succ'}) - s_{(J_l,P_l)}(T_l, \overline{e})) - (s_{(J_l,P_l)}(T_l, e^{\succ}) - s_{(J_l,P_l)}(T_l, \overline{e}))$$
$$\leq s_{(J_l,P_l)}(T_l, \overline{e}) - s_{(J_l,P_l)}(T_l, e^{\succ}) \leq \delta(T_l).$$
By assumption $i$ is one of at least $N_0$ students $j \in J_l^c$ such that $e_j^{\succ'}(t) = o_l$ for all $t \in [T_l, T_l')$, so $T_l' - T_l \leq s_{(J_l,P_l)}(T_l, e^{\succ'})/N_0 \leq \delta(T_l)/N_0$. $\square$

Let $\lambda = 1 + 1/N_0$.

**Lemma 13.** For all $l = 1, \ldots, \ell$, $T_l' - T_l \leq \gamma(1)(\lambda - 1)\lambda^{l-1}$.

*Proof.* We prove the lemma by induction on $l$. We have $\delta(T_1) = \gamma(T_1) \leq \gamma(1)$, so Lemma 12 implies that $T_1' - T_1 \leq \delta(T_1)/N_0 \leq \gamma(1)(\lambda - 1)$. Suppose that $l \geq 2$ and the induction hypothesis holds for $1, \ldots, l-1$. Then

$$\delta(T_l) = \gamma(T_l) + \beta(T_l) \leq \gamma(1) + \sum_{g=1}^{l-1} \beta(T_{g+1}) - \beta(T_g) = \gamma(1) + \sum_{g=1}^{l-1} T_g' - T_g$$

$$\leq \gamma(1)\Big(1 + (\lambda - 1)\sum_{g=0}^{l-2} \lambda^g\Big) = \gamma(1)\lambda^{l-1}.$$

Applying Lemma 12 again gives

$$T_l' - T_l \leq \delta(T_l)/N_0 \leq \gamma(1)(\lambda - 1)\lambda^{l-1}. \qquad \square$$

*Proof of Theorem 5.* We have

$$u_i(GCPS(\succ)) - u_i(GCPS(\succ')) = \int_0^1 u_i(e_i^{\succ}(s)) - u_i(e_i^{\succ'}(s))\, ds \geq d_i\gamma(1) - D_i\beta(1).$$

Since $\beta(T_1) = 0$, adding up the inequalities from Lemma 13 for gives

$$\beta(1) = \sum_{g=1}^{\ell} \beta(T_{g+1}) - \beta(T_g) = \sum_{g=1}^{\ell} T_g' - T_g \leq \gamma(1)(\lambda - 1)\sum_{g=1}^{\ell} \lambda^{g-1} = \gamma(1)(\lambda^\ell - 1).$$

Therefore

$$u_i(GCPS(\succ)) - u_i(GCPS(\succ')) \geq \gamma(1)\big(d_i - D_i(\lambda^\ell - 1)\big),$$

and since $\ell \leq |O|$, this is nonnegative if $d_i/D_i \geq \lambda^{|O|} - 1 = (1 + 1/N_0)^{|O|} - 1$. $\square$