DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

HOW TO CONTROL CONTROLLED SCHOOL CHOICE

Federico Echenique
California Institute of Technology

M. Bumin Yenmez
Carnegie Mellon University

SOCIAL SCIENCE WORKING PAPER 1366

December 2012

How to control controlled school choice

Federico Echenique M. Bumin Yenmez

Abstract

We characterize choice rules for schools that regard students as substitutes, while at the same time expressing preferences for the diversity composition of the student body. The stable (or fair) assignment of students to schools requires the latter to regard the former as substitutes. Such a requirement is in conflict with the reality of schools' preferences for a diverse student body. We show that the conflict can be useful, in the sense that certain unique rules emerge from imposing both considerations.

JEL classification numbers: D47,D71

Key words: Matching, Market Design, Gale-Shapley, School Choice, Diversity

How to control controlled school choice*

Federico Echenique M. Bumin Yenmez

> *. . . controlled choice provides local officials with a final student assignment policy that maximizes family choice and effective desegregation outcomes on a districtwide basis, provides stability of assignment . . . , and makes all schools and programs available to students of diverse racial, ethnic, and socioeconomic backgrounds (Alves and Willie, 1987, Page 75).*

1 Introduction

Recent reforms in school choice programs seek to install a stable (or fair) assignment of students to schools (Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu, Pathak, Roth, and Sönmez, 2005). This objective is severely compromised by school districts' concerns for diversity. Under diversity considerations, a stable assignment may not exist, and the mechanisms used in reformed school districts may not work. There is in fact a very basic tension between diversity considerations and the assumptions required in the theory of stable matching: diversity concerns will introduce complementarities in schools' preferences; the theory requires substitutability.

If a school is concerned with gender balance, for example, then it may admit a mediocre male applicant only because it allows the school to admit an excellent female applicant, while maintaining gender balance. The two students are thus complements, not substitutes, for the school. Complementarities in the school's choices of students are a problem because the theory, and the mechanism proposed in school choice programs, require that students are substitutes in schools' choices. We are far from the first to

*We thank Estelle Cantillon for her helpful comments, as well as seminar audiences at MIT/Harvard, ECARES, Sciences-Po, University of Arizona, University of California-Davis, University of California-Irvine, University of Montreal.

recognize this problem: Section 1.1 below discusses the relevant literature. The idea that diversity clashes with stability is very easy to recognize; in Section 1.2 we present a particularly simple example of the incompatibility between stability and diversity concerns.

Our paper seeks to reconcile diversity with the objective of seeking a stable matching of students to schools. We characterize the schools' choices that are compatible with *both* diversity considerations and the theory of stable matchings. There is so much tension between substitutability and diversity that one might think no choice rule can satisfy both. We prove that this does not need to be the case: We study the choices that satisfy certain normative axioms, one of them being substitutability, and show how combinations of axioms give rise to unique choice procedures, some of which are already implemented in practice. Our procedures allow schools to express concerns for diversity, while allowing the standard mechanism (the one used in the school choice programs guided by stable matching theory) to install a stable assignment of students to schools.[1]

We assume that students belong to one of multiple types. Types could be categories of gender, socioeconomic status, race or ethnicity. In all our results, there is an "ideal" or "target" distribution that plays a crucial role. For example, we axiomatize a rule that tries to minimize the (Euclidean) distance between the distribution over types in the student body, and some ideal distribution over types (see Section 4). A common consideration is that each school should have a share of White, Black, Hispanic, etc. children that matches, as close as possible, the distribution of races and ethnicities in the relevant population (Alves and Willie, 1987). Our rule operationalizes this consideration, where the population distribution is the ideal to be reached for. The axiomatization tells us what such a rule means, in terms of normative qualitative criteria.

In two other rules, the school reserves a number of seats for each type of students (hard and soft quotas, see Section 5). The number of seats reserved for each type is related to the target distribution over types. Yet another rule (Section 4) seeks to maximize a measure of diversity (for example the Theil measure of diversity, see Theil (1967); Foster and Sen (1997)); the target distribution enters as a parameter in the measure of diversity.

In Section 3 we give a brief overview of these rules and the corresponding axioms. The point of these results is that the rules result uniquely from the normative considerations

[1]We do not propose any new mechanisms: we want a theory that will work with the mechanisms that have already been accepted and adopted by multiple school districts. Indeed, these mechanisms have been accepted across many different market design problems (Roth, 2008), not only in the assignment of students to schools.

underlying our axioms.

We imagine that a school district can discuss a menu of axioms, and settle on the axioms that it deems most desirable. Basically, schools have given priorities, "preferences," over individual students. These priorities can result from test scores, or from the distance of the student's residence to the school and other objective criteria. The school also has preferences over the type composition of the student body: these preferences come from concerns over diversity. Now, the school or the district may combine these two preferences in different ways. Our results give recommendations on how the combination should be carried out so that the standard mechanism in matching theory will work. If a school, or a district, agrees on a set of axioms, then there will be a *unique* way of combining priorities and diversity preferences into a choice procedure for the school. In fact, as we explain in Sections 7 and 8, our rules are already similar to policies being implemented around the world.

In Section 3 we provide a discussion of the axioms that we use, and an overview of our results. In Sections 4-5 we present our characterization results. In Sections 7 and 8 we discuss actual implementations of affirmative action policies.

1.1 Related literature

Abdulkadiroğlu and Sönmez (2003) introduced matching theory as a tool in school choice and noted the problem with diversity concerns. Abdulkadiroğlu and Sönmez (2003) already raise the issue of diversity; they offer a solution based on hard quotas, one of the models we axiomatize below.

The last two years have seen multiple explorations into controlled school choice and diversity concerns. Kojima (2010) shows that affirmative action policies based on majority quotas may hurt minority students. To overcome this difficulty, Hafalir, Yenmez, and Yildirim (2011) propose affirmative action based on minority reserves. They show that the outcome of the deferred acceptance algorithm (DA) with minority reserves Pareto dominates DA with majority quotas. More generally, Ehlers, Hafalir, Yenmez, and Yildirim (2011) study affirmative action policies when there are both type-specific upper and lower bounds. They propose solutions based on whether the bounds are hard or soft. In contrast, our paper seeks to endogenize the rules and consider (possibly) all of them. Part of our research deals with the results uncovered by Hafalir, Yenmez, and

Yildirim (2011). There are other papers that consider specific choice rules (Westkamp, 2010; Kominers and Sönmez, 2012; Erdil and Kumano, 2012).

In contrast with the other papers in the literature, our focus is not on the market as a whole, but rather on the preferences or choices of individual schools. We imagine the mechanism is fixed (the deferred acceptance mechanism), and that we can design schools' choices to satisfy certain normative axioms.

We focus on school preferences, but student preferences may also induce problems: for example students may care about their colleagues. These problems are treated in Echenique and Yenmez (2007) and Pycia (2012); they are outside the scope of the present analysis. We focus here on diversity, and the effects of diversity on standard stable matching theory. Our exercise pins down reasonable circumstances in which schools may be concerned about diversity, and where the theory still remains useful because schools satisfy gross substitutes.

1.2 Motivating example

In this example, we demonstrate the basic conflict between diversity concerns and the gross substitutes. Suppose that there are two schools, c_1 and c_2, and two students, s_1 and s_2. Students are of different "type." For example, s_1 and s_2 could be of different gender, race or ethnicity.

School c_1 can admit two students, but it is constrained to mimic the population representation of each type. So it must admit either both students or none. School c_2 has a single empty seat. It prefers to admit student s_1 over student s_2.

The students have preferences over schools as well: s_1 prefers c_1 over c_2, while s_2's favorite school is c_2. The table below summarizes the agents' preferences.

c_1	c_2	s_1	s_2
$\{s_1, s_2\}$	s_1	c_1	c_2
	s_2	c_2	c_1

Now it is easy to see that no assignment of students to schools is fair, or stable. For example, if both students are assigned to school c_1 then s_2 might ask for the empty slot in school c_2. School c_2 finds s_2 acceptable, so the pair (c_2, s_2) can "block" this assignment (equivalently, s_2's claim to the empty seat is "justified"). Similarly, if s_2 is assigned to c_2 then s_1 would have no place, as school c_1 cannot admit an unbalanced student body. Then s_1 would claim s_2's spot in school c_2. Since s_1 has a higher priority than s_2 at that school, (c_2, s_1) can "block" this assignment (or s_1's claim is "justified") and thus the assignment of s_2 to c_2 is unfair. Finally, if s_1 is assigned to school c_2 and s_2 is unassigned, then both students would prefer school c_1, and school c_1 prefers to get both of them. Therefore, $(c_1, \{s_1, s_2\})$ can block the assignment.

Thus, there exists no stable or fair assignment of students to schools in this example. The intuition is that c_1's preferences for diversity cause complementarities between students that make it impossible to have a stable assignment.

2 Model

2.1 Notational conventions

For any vector $x \in \mathbf{Z}^d_+$, let $||x|| \equiv \sum_{i=1}^d x_i$ be the sum of its coordinates. For any $x, y \in \mathbf{Z}^d_+$, let $x \wedge y \equiv (\min\{x_1, y_1\}, \ldots, \min\{x_d, y_d\})$ and $x \vee y \equiv (\max\{x_1, y_1\}, \ldots, \max\{x_d, y_d\})$ be the infimum and supremum of x and y, respectively. For a finite set A, $|A|$ denotes the cardinality of A.

2.2 Schools' admissions choices

We consider the admissions choices of an individual school or college. A school's admissions policy is described by a choice rule that determines which students to admit from a pool of applicants. Our model is therefore one of the most basic models in microeconomics: see for example Moulin (1991), or Chapter 2 in Mas-Colell, Whinston, and Green (1995). Later in the paper we study the market-wide implications of our results.

Let $\mathcal{S}$ be a nonempty finite set: the set of all ***students***. A ***choice rule*** is a function $C : 2^{\mathcal{S}} \setminus \{\emptyset\} \to 2^{\mathcal{S}}$ such that $C(S)$ is a subset of S, for all $S \subseteq \mathcal{S}$. The interpretation of C is that, if a school had the ability to admit its students out of the set S of students, then it would choose $C(S)$ to be its student body.

We shall assume that there is a positive number q such that $|C(S)| \leq q$ for all $S \subseteq \mathcal{S}$. The number q is the ***capacity*** of the school: the number of available seats that it has.

The set of students $\mathcal{S}$ is partitioned into students of different "types," these can be based on gender, socioeconomic factors, or race and ethnicity. Formally, there exists a set $T \equiv \{t_1, \ldots, t_d\}$ of ***types***, and a type function $\tau : S \to T$; $\tau(s)$ is the type of student s. Let $\mathcal{S}^t$ be the set of type-t students, i.e., $\mathcal{S}^t \equiv \{s \in \mathcal{S} : \tau(s) = t\}$. Similarly, for any set of students $S \subseteq \mathcal{S}$, let $S^t \equiv S \cap \mathcal{S}^t$.

We use a function $\xi : \mathcal{P}(\mathcal{S}) \to \mathbf{Z}_+^d$ to describe how many students of each type a given set of students has. More formally, let

$$\xi(S) \equiv (|S^{t_1}|, \ldots, |S^{t_d}|) \in \mathbf{Z}_+^d,$$

which consists of the number of students of each type in S. We term $\xi(S)$ the ***distribution*** of students in S.

We assume that the school is not large enough that it could admit all students of a given type: $q \leq |\mathcal{S}^t|$ for all $t \in T$.[2]

3 Characterizations of choice rules: Overview

We shall define the main substantive axioms in our study, and give an overview of our results.

The main axiom that we are interested in is the following:

Axiom 1. *Choice rule C satisfies* ***gross substitutes (GS)*** *if $s \in S \subseteq S'$ and $s \in C(S')$ imply that $s \in C(S)$.*

Gross substitutes requires that a student chosen from a set S' must also be chosen from a subset of $S \subseteq S'$ if the student is a member of the subset. Gross substitutes was first studied by Kelso and Crawford (1982) and Roth (1984): It is sufficient for the existence of stable matchings and for the Gale-Shapley algorithm to find a stable matching. It is also in some sense necessary for these properties to hold (Hatfield and Milgrom, 2005; Hatfield and Kojima, 2008).[3]

We impose gross substitutes as a basic constraint on choice rules because, as we have emphasized above, we want to ultimately achieve stable (or fair) assignments of students

[2] This assumption is reasonable, but not important for our results. We only use it because it makes it easier to write some of our proofs. As far as we know, none of our results depend on it.

[3] Note that gross substitutes is formally identical to Sen's α. The interpretation is different, though, because here $C(S)$ is the chosen subset of $C(S)$, not a set of alternatives that are "equally good."

to schools. By looking at other axioms we obtain four different basic models of controlled school choice.

Aside from gross substitutes, we can classify the axioms in two categories, diversity and rationality axioms. The following table summarizes our main results:

Model		Diversity				Rationality		
	GS	Mon	DD	Eff	RM	t-WARP	SARP1	SARP 2
Ideal point	✓	✓				✓		
Schur	✓		✓	✓		✓		
Soft quotas	✓			✓		✓	✓	
Hard quotas	✓				✓	✓		✓

The first category of axioms constrain distributions over types, we think of them as diversity axioms. We have looked at four such axioms: The monotonicity axiom (Mon in the table) says that an increase in the distribution over the set of applicants should result in an increase in the distribution over the admitted students. The distribution dependence axiom (DD) is a weaker form of monotonicity, and requires that if two sets of applicants have the same distribution, then the sets of admitted students should also have the same distribution. The efficiency axiom (Eff) states that a student should never be rejected if there is an empty seat. Rejection maximality (RM) requires that if a student is rejected from a school that has space for him, then a maximal number of students of his type must have been achieved.

The second category of axioms impose basic rationality criteria: they are versions of standard axioms from revealed preference theory. We use them here to elicit a priority order over individual students that is compatible with the choice rule. The type-weak axiom of revealed preference (tWARP) deals with comparisons between students of the same type. We look at two versions of the strong axiom of revealed preference (SARP1 and SARP2), which become relevant once we compare students of different types.[4]

Our results deal with four basic models of choice. As we discuss in Sections 7 and 8, these are close to systems that are already used in many places to achieve diversity.

The first two models are somewhat rigid: the "ideal point" model tries first to achieve a distribution over types that is as close as possible (in Euclidean distance) to some ideal distribution over types. Given such a choice, it selects the best available, or highest

[4]One final axiom, irrelevance of rejected students, was omitted from the table. It is a basic rationality criterion, and it is satisfied by all our models. Indeed, it is hard to imagine a normative model of choice that would violate irrelevance of rejected students.

priority, students of each type. In the Schur model the distribution does not try to approximate some ideal point but instead it seeks to maximize some measure of the degree of diversity of the school.

The first two models are rigid because, given a set of applicants, they first formulate the distribution of the choice set, and then admit the highest priority students of every type without exceeding the distribution. In the next two models, hard and soft quotas, there is some tradeoff between the priority of students of different types and the distribution over types. A student of one type may be admitted over a student of a different type because he has a higher priority. Of course, the degree to which this can happen is limited by diversity considerations.

In the model of soft quotas, a school reserves a number of seats for each type. The school then tries to fill these reserved seats; some of them may be unfilled if there are not enough applicants of a given type. For the remaining seats, which are not taken yet, students compete "openly". This model is flexible because the distribution over types is partially determined by the students' priorities.

In the model of hard quotas, instead of seats being reserved for a type, there is an upper bound, or quota, on how many students of a given type may be accepted. Student compete "openly" for seats until they hit the quota on their types. After that happens, a student may be turned down in favor of a lower priority candidate whose quota has not been obtained yet.

4 Ideal points and Schur concavity

We analyze the first two models discussed in Section 3. Both of these models involve two basic rationality axioms.

We want a priority order over individual students to play a role in the school's choices, and for that we need to make sure that the choice satisfies basic revealed preference axioms. The two axioms we use are versions of standard properties in the decision theoretic literature, used in the study of abstract choice rules and social choice (see, e.g., Moulin (1991)).[5]

[5]Rationality could instead require a preference relation over sets of students. We are focusing on priorities over individual students because it seems to be what most schools use in forming their preferences.

Axiom 2. *Choice rule C satisfies the* ***type-weak axiom of revealed preference (t-WARP)*** *if, for any s, s', S and S' such that $\tau(s) = \tau(s')$ and $s, s' \in S \cap S'$,*

$$s \in C(S) \text{ and } s' \in C(S') \setminus C(S) \text{ imply } s \in C(S).$$

The type-weak axiom of revealed preference is necessary for the existence of some underlying priority ordering over students. We need it to ensure that a school admits the best students of each type, given the underlying priority order.

Our second rationality axiom simply says that a rejected student may be made unavailable without affecting the set of chosen students. It has been used before in the matching context by Alkan (2002), and Alkan and Gale (2003); and by Fleiner (2003) and Aygün and Sönmez (2012) for markets with contracts.

Axiom 3. *Choice rule C satisfies* ***irrelevance of rejected students (IRS)*** *if $C(S') \subseteq S \subseteq S'$ implies that $C(S) = C(S')$.*

To be clear, gross substitutes and the two rationality axioms capture standard properties of choice. Gross substitutes is important given that we want to apply our results to matching markets.

Next, we impose axioms capturing a concern for diversity. Different axioms deliver different choice rules: we focus on those that are generated by ideal points and Schur concave monotone functions.

4.1 Ideal points

A school may have an ideal distribution that it tries to achieve. For example, it may strive for perfect gender balance, or for a distribution over races and ethnicities that match those in the population. Here we characterize those rules that try to minimize the Euclidean distance from the distribution of admitted students to the ideal distribution. In Sections 7 and 8 we give examples of how actual school districts' policies reflect this concern.

Choice rule C is ***generated by an ideal point*** if there is a vector $z^* \in \mathbf{Z}^d_+$ with $||z^*|| \leq q$ and a strict priority $\succeq$ over $\mathcal{S}$ such that, for any $S \subseteq \mathcal{S}$, (1) $\xi(C(S))$ is the closest vector to z^*, among those in $B(\xi(S))$ where

$$B(x) \equiv \{z \in \mathbf{Z}_+^d : z \leq x \text{ and } ||z|| \leq q\};$$

and (2) the students of type t in $C(S)$ have higher priority than any student of type t in $S \setminus C(S)$, for any t.

Our next axiom states that an increase in the number of applicants of every type should give rise to an increase in the admissions of every type.

Axiom 4. *Choice rule C satisfies **distribution-monotonicity (Mon)** if $\xi(S) \leq \xi(S')$ implies that $\xi(C(S)) \leq \xi(C(S'))$.*

One consequence of distribution-monotonicity is to say that a school prioritizes its distribution. It "first" fixes a target distribution, and then chooses the best student body to fit that distribution. It is, therefore, a strong assumption, and possibly questionable (the models of soft and hard quotas in Section 5 relax this assumption). We believe that it is a good approximation to how many schools operate in actuality because many schools have diversity targets independently of the quality-composition of the body of applicants.

Importantly, distribution-monotonicity is compatible with most forms of pure diversity concerns. Including diversity policies that imply a failure of gross substitutes, and non-existence of stable matchings. Note that the example in Section 1.2 satisfies the distribution-monotonicity axiom.

Theorem 1. *Choice rule C satisfies GS, t-WARP, IRS, and Mon, if and only if it is generated by an ideal point.*

The result in Theorem 1 is surprising because the requirement of gross substitutes has nothing to do with diversity. Rather, Mon says that the school has diversity as a primary objective, but there are many ways in which diversity can be implemented. The tension between gross substitutes and diversity is important enough, however, that when we put the four axioms together, only ideal point rules survive.

Remark 1. The type-weak axiom of revealed preference alone does not suffice to give a rationalizing priority relation because it only rules out revealed preference cycles of length two. Normally, in a revealed-preference exercise, one needs to rule out cycles of any length. As a result one needs the strong axiom of revealed preference (or Richter's notion of consistency, see Richter (1966)).[6] It is interesting that the axiom of gross substitutes "aids" the weak axiom, and allows us to rule out cycles of any length.[7]

4.2 Schur concave

The distance to an ideal distribution is a reasonable criterion, but it may lead to inefficiencies. A school who is deeply committed to diversity may leave some seats empty (or under-report its capacity) when the newcomers would upset the distribution over gender, or race/ethnicity, of the student body. That said, it may be reasonable to require schools to be efficient in the sense of never leaving a seat empty if they can fill it. To this end, we now substitute the monotonicity axiom that we used above for an efficiency axiom: the school is required to fill all seats that it can fill.

Axiom 5. *Choice rule C satisfies **efficiency** if $C(S) = S$ when $|S| \leq q$, and $|C(S)| = q$ when $|S| > q$.*

We still need to ensure that the school cares primarily about diversity: that it sets a diversity objective independently of its body of applicants. Thus, we use the following weakening of the distribution-monotonicity axiom.

Axiom 6. *Choice rule C satisfies **distribution-dependence** if $\xi(S) = \xi(S')$ implies that $\xi(C(S)) = \xi(C(S'))$.*

When a choice rule is distribution-dependent, then for any two sets with the same distribution the set of admitted students also have the same distribution. However, in contrast with distribution-monotonicity, it does not say anything about two sets which have different distributions.

As a result of these axioms, choice is driven by a measure of diversity. Researchers studying diversity often consider numerical measures of diversity (such as entropy, or Theil's index). For example, the ecological diversity studied by Weitzman (1992) is a special case of entropy. These numerical measures are often Schur concave: a property that we shall not define here. Instead, we shall use a canonical construction of a Schur concave function (see Marshall, Olkin, and Arnold (2010)). A school satisfying our axioms will seek to maximize the sum of values of a monotone increasing and concave function.

[6] Alternatively, one may have to observe choice from all sets with two or three elements, but such a condition is not useful in the present model. Ehlers and Sprumont (2008) is one study of behavior described by WARP, allowing for cyclic choices. In our case it turns out that cycles are ruled out by the interaction of t-WARP with GS and IRS.

[7] The reason is partly that GS coincides with Sen's α.

More formally, we say that C is ***Schur-generated*** if there is a point $z^* \in \mathbf{Z}^d_+$ with $||z^*|| \leq q$, an increasing and concave function $g : \mathbf{R} \to \mathbf{R}$, and a strict priority $\succeq$ over $\mathcal{S}$ such that

1. $\sum_{t=1}^d g(x_t - z^*_t)$ achieves a maximum in $B(\xi(S))$ at $\xi(C(S))$;
2. $\xi(C(S)) = \xi(C(S'))$ for any S and S' with $\xi(S) = \xi(S')$; and
3. students of type t in $C(S)$ have higher priority than any student of type t in $S \setminus C(S)$.

Theorem 2. *Choice rule C satisfies GS, t-WARP, IRS, efficiency, and distribution-dependence, if and only if it is Schur-generated.*

The interpretation of this model is as follows. Suppose that $z^* = 0$. Then the maximization of $\sum_{t=1}^d g(x_t)$ involves values of x_t which are as close to each other as possible. That is, it seeks to obtain equal representation of all types in the school. Otherwise, when $z^* > 0$, then the maximum is going to be achieved at a point when $x \geq z^*$ and $x_t - z^*_t$ are as close to each other as possible. Equivalently, the school tries to achieve a distribution of students z^* and tries to get the same number of students in excess.

5 Hard and soft quotas

The model in Section 4 is restrictive in the sense that diversity considerations pin down a unique distribution, and then a priority relation is used to choose students within each type. This is restrictive because, given the distribution of the applicant body, only one distribution is allowed for the admitted students. In this section we consider schools that may have a more flexible commitment to diversity.

In particular, we may want to allow a school to trade off students of different types. Such a trade off may directly conflict with diversity concerns: it may involve rejecting an undesirable type 1 student in favor of a good type 2 student, thereby affecting the distribution over types.

5.1 Hard quotas

A school may want to limit the number of admitted students who has the same type, so it may have a type-specific limit or quota for each type. However, as long as these quotas are not exceeded, the school does not differentiate between students with different types. This is the model studied in Abdulkadiroğlu and Sönmez (2003).

Choice rule $\mathcal{C}$ is ***generated by hard quotas*** if there exist a strict priority $\succeq$ over $\mathcal{S}$, and a vector $(r_t)_{t\in T} \in \mathbf{Z}^d_+$, such that for any $S \subseteq \mathcal{S}$,

1. $|C(S)^t| \leq r_t$;
2. if $s \in C(S)$, $s' \in S \setminus C(S)$ and $s' \succ s$, then it must be the case that $\tau(s) \neq \tau(s')$ and $\left|C(S)^{\tau(s')}\right| = r_{\tau(s')}$; and
3. if $s \in S \setminus C(S)$, then either $|C(S)| = q$ or $\left|C(S)^{\tau(s)}\right| = r_{\tau(s)}$.

In this case, r_t is an upper bound on the number of students of type t that the school can accept. The school considers all students and chooses the highest ranked ones conditional on not exceeding any of the quotas. In particular, if $||r|| \leq q$ then this model is equivalent to the ideal point model.

Say that S is ***ineffective*** for t if there is S' such that $|S^t| = |S'^t|$ with $|C(S)^t| < |C(S')^t|$. In words, a set is ineffective for type t when the school does not accept the maximum number of type t students among the sets with the same number of type t applicants. This notion of ineffectiveness is crucial in our axiom below.

Axiom 7. *Choice rule C satisfies the* ***effective strong axiom of revealed preference (E-SARP)*** *if there are no sequences $\{s_k\}_{k=1}^K$ and $\{S_k\}_{k=1}^K$, of students and sets of students, respectively, such that, for all k*

1. *$s_{k+1} \in C(S_{k+1})$ and $s_k \in S_{k+1} \setminus C(S_{k+1})$;*
2. *$\tau(s_{k+1}) = \tau(s_k)$ or S_{k+1} is ineffective for $\tau(s_k)$.*

(using addition mod K).

E-SARP rules out certain cycles in the revealed preference of the choice rule, but it is careful as to where it infers a revealed preference from choice. The subtlety in the definition is the second part that requires either $\tau(s_{k+1}) = \tau(s_k)$ or that S_{k+1} is ineffective for $\tau(s_k)$. In the first case, when s_{k+1} and s_k have the same type, it is revealed that s_{k+1} has a higher priority than s_k. However, when they have different types, s_{k+1} is revealed preferred to s_k only when S_{k+1} is ineffective for $\tau(s_k)$ implying that the school could admit more students of type $\tau(s_k)$. It is easy to see that E-SARP implies t-WARP.

Our next axiom is a diversity axiom. It states that whenever a student of type t is rejected from a set of students S when there is an empty seat in the school, then it must be the case that the school has admitted the most number of type t students for any possible set of applicants.

Axiom 8. *Choice rule C satisfies* ***rejection maximality (RM)*** *if $s \in S \setminus C(S)$ and $|C(S)| < q$ imply that $\left|C(S)^{\tau(s)}\right| \geq \left|C(S')^{\tau(s)}\right|$ for every S'.*

In words, if a student of type t is rejected when there is an empty seat, then it must be that the school has admitted the most number of type t students. Rejection maximality is the main axiom we use to construct the hard quotas.

Theorem 3. *Choice rule C satisfies GS, E-SARP, and RM if and only if C is generated by hard quotas.*

Remark 2. We characterize choice rules that are generated by hard quotas in terms of GS, E-SARP and RM. However, it is easy to see that if a choice rule C is generated by hard quotas then it also satisfies IRS. We reconcile this observation with the result above by showing that if a choice rule satisfies GS and RM then it also satisfies IRS.

Suppose that choice rule C satisfies GS and RM. Let S' and S be such that $C(S') \subseteq S \subseteq S'$. By GS, $C(S') \subseteq C(S)$. Suppose for contradiction that there exists $s \in C(S) \setminus C(S')$. This implies that $s \in S' \setminus C(S')$ and $|C(S')| < q$. By RM, $\left|C(S')^{\tau(s)}\right| \geq \left|C(\hat{S})^{\tau(s)}\right|$ for any $\hat{S}$. Letting $\hat{S} = S$ yields, $\left|C(S')^{\tau(s)}\right| \geq \left|C(S)^{\tau(s)}\right|$, which is a contradiction since $C(S') \subseteq C(S)$ and $s \in C(S) \setminus C(S')$. Therefore, there does not exist $s \in C(S) \setminus C(S')$, so $C(S) = C(S')$ and IRS is satisfied.

Even when $||r|| > q$ a choice rule that is generated by hard quotas r can be inefficient. For example, suppose that all applicants have the same type and their quota is less than the capacity of the school. In this case, the school is not going to fill its capacity. Next, we impose efficiency and IRS instead of rejection maximality and get a different model that can be characterized by *soft quotas*.

5.2 Soft Quotas

Suppose that there are type-specific bounds that the school would like to implement. In other words, the school limits the number of students of each type that it admits. However, if the school cannot fill its capacity by imposing these bounds, then it can still admit more students until the capacity is filled or the set of applicants is exhausted.

Therefore, these bounds are not hard quotas as we have defined above but soft quotas. We capture this notion below.

More formally, choice rule C is ***generated by soft quotas*** if there exist a strict priority $\succeq$ over $\mathcal{S}$, and a vector $(r_t)_{t\in T} \in \mathbf{Z}_+^d$ with $||r|| \leq q$, such that for any $S \subseteq \mathcal{S}$,

1. $|C(S)^t| \geq r_t \wedge |S^t|$;
2. if $s \in C(S)$, $s' \in S \setminus C(S)$ and $s' \succ s$, then it must be the case that $\tau(s) \neq \tau(s')$ and $\left|C(S)^{\tau(s)}\right| \leq r_{\tau(s)}$; and
3. if $s \in S \setminus C(S)$, then $|C(S)| = q$.

In words, the school reserves r_t seats for students of type t. Given a pool of students S, it admits the best $r_t \wedge |S^t|$ students according to priority $\succeq$. In a second stage, it admits the best (according to priority $\succeq$) students regardless of type, among the remaining students.

Say that $t \in T$ is ***saturated*** at S if there is S' such that $|S^t| = |S'^t|$ with $S'^t \setminus \mathcal{C}(S')^t \neq \emptyset$. The interpretation is that when there are $|S^t|$ students the school is not obliged to accept them all out of diversity considerations. Our first axiom builds on this notion.

Axiom 9. *Choice rule C satisfies the **adapted strong axiom of revealed preference (A-SARP)** if there are no sequences $\{s_k\}_{k=1}^K$ and $\{S_k\}_{k=1}^K$, of students and sets of students, respectively, such that, for all k*

1. *$s_{k+1} \in C(S_{k+1})$ and $s_k \in S_{k+1} \setminus C(S_{k+1})$;*
2. *$\tau(s_{k+1}) = \tau(s_k)$ or $\tau(s_{k+1})$ is saturated at S_{k+1}*

(using addition mod K).

The adapted SARP rules out the existence of certain cycles in revealed preference, where again we are careful as to when we infer the existence of a revealed preference. It is stronger than t-WARP. The difference between E-SARP and A-SARP is the second component of the definition, when we require that $\tau(s_{k+1})$ is saturated at S_{k+1}. When this happens, even though the school could admit fewer type $\tau(s_{k+1})$ students, it accepts more. Thus in the revealed preference, s_{k+1} is preferred to s_k even if they have different types. This axiom allows us to construct a priority order over students.

Theorem 4. *Choice rule C satisfies GS, A-SARP, Eff and IRS if and only if C is generated by soft quotas.*

6 Implications for school choice and matching markets

A ***matching market*** is a tuple $\langle \mathcal{C}, \mathcal{S}, (\succ_s)_{s\in\mathcal{S}}, (C_c)_{c\in\mathcal{C}}\rangle$, in which $\mathcal{C}$ is a finite set of ***schools***, $\mathcal{S}$ is a finite set of ***students***, for each $s \in \mathcal{S}$; $\succeq_s$ is a strict preference order over $\mathcal{C} \cup \{s\}$ where $\{s\}$ is the outside option for student s,[8] and for each $c \in \mathcal{C}$; C_c is a choice rule over $\mathcal{S}$.

A ***matching*** (or an ***assignment***) μ is a function on the set of agents such that

1. $\mu(c) \subseteq \mathcal{S}$ for all $c \in \mathcal{C}$ and $\mu(s) \in \mathcal{C} \cup \{s\}$ for all $s \in \mathcal{S}$;

2. $s \in \mu(c)$ if and only if $\mu(s) = c$ for all $c \in \mathcal{C}$ and $s \in \mathcal{S}$.

In a matching market, we would like to find *stable* matchings that satisfy indiviual rationality and fairness properties that we formalize below.

Definition 1. *A matching μ is **stable** if*

1. *(**individual rationality**) $C_c(\mu(c)) = \mu(c)$ for all $c \in \mathcal{C}$, $\mu(s) \succeq_s \{s\}$ for all $s \in \mathcal{S}$; and*

2. *(**no blocking**) there exists no (c, S') such that $S' \not\subseteq \mu(c)$ such that $S' \subseteq C_c(\mu(c) \cup S')$ and $c \succeq_s \mu(s)$ for all $s \in S'$.*

Stability requires both individual rationality and no blocking. First, individual rationality for schools requires that no school can be better off by rejecting some of the admitted students; whereas for students it only requires that each student prefers their assigned schools to their outside options. Second, no blocking requires that there exists no coalition of agents who can beneficially rematch among themselves. This is the standard definition of stability used in many-to-one matching problems (Roth and Sotomayor, 1990).

For matching markets, stability has proved to be a useful solution concept because mechanisms that find stable matchings are successful in practice (Roth, 2008). Moreover, finding stable matchings is relatively easy. In particular, the deferred acceptance algorithm of Gale and Shapley (1962) finds a stable matching and the algorithm has other attractable properties. Therefore, it also serves as a recipe for market design. For example, it has been adapted in New York and Boston school districts (see Abdulkadiroğlu,

[8]The outside option for student s can be going to a private school or being homeschooled.

Pathak, Roth, and Sönmez (2005) and Abdulkadiroğlu, Pathak, and Roth (2005)). For completeness, we provide a description of the student-proposing deferred acceptance algorithm.

Deferred Acceptance Algorithm (DA)

Step 1 Each student applies to her most preferred school. Suppose that S_c^1 is the set of students who applied to school c. School c tentatively admits students in $C_c(S_c^1)$ and permanently rejects the rest. If there are no rejections, stop.

Step k Each student who was rejected at Step $k-1$ applies to their next preferred school. Suppose that S_c^k is the set of new applicants and students tentatively admitted at the end of Step $k-1$ for school c. School c tentatively admits students in $C_c(S_c^k)$ and permanently rejects the rest. If there are no rejections, stop.

The algorithm ends in finite time since at least one student is rejected at each step.[9]

Usually the only strategic component of a matching market is the student preference profile $(\succ_s)_{s\in\mathcal{S}}$; schools' choice rules are fixed by laws and regulations. Therefore, we only worry that students and their families state their preferences truthfully. To this end, we consider a group strategyproofness concept.

Let $P_{\mathcal{S}}$ be the space of student preference profiles $(\succ_s)_{s\in\mathcal{S}}$ and $\mathcal{M}$ be space of matchings between $\mathcal{C}$ and $\mathcal{S}$. Therefore, a ***mechanism*** is a function $\Phi : P_{\mathcal{S}} \to \mathcal{M}$.

Definition 2. *Mechanism Φ is **group incentive compatible for students** if there exists no group of students S', preference profiles $(\succ_s)_{s\in\mathcal{S}}$, and $(\succ'_s)_{s\in S'}$ such that*

$$\Phi((\succ'_s)_{s\in S'}, (\succ_s)_{s\in\mathcal{S}\setminus S'}) \succ_s \Phi((\succ_s)_{s\in\mathcal{S}})$$

for all $s \in S'$.

In words, a mechanism is group incentive compatible for students if there exists no group of students who can jointly manipulate their preferences to be matched with better schools.

The following axiom for choice rules plays a critical role in establishing the desirable properties of the deferred acceptance algorithm.

[9] For a history of the deferred acceptance algorithm, see Roth (2008).

Axiom 10. *Choice rule C satisfies* ***the law of aggregate demand (LAD)*** *if $S \subseteq S'$ implies $|C(S)| \leq |C(S')|$.*

The law of aggregate demand requires that the number of students chosen from a subset of a set of students S' should not be bigger than the number of students chosen from S'. This property was first introduced in Alkan (2002) and Alkan and Gale (2003) for matching markets without transfers (or contracts), and by Fleiner (2003) and Hatfield and Milgrom (2005) for markets with contracts. All of the choice rules that we have studied in Sections 4 and 5 satisfy LAD. In particular, it is easy to see that distribution monotonicity or efficiency implies LAD.

The following result is well-known (see Roth and Sotomayor (1990), Hatfield and Milgrom (2005), Hatfield and Kojima (2009), and Aygün and Sönmez (2012)). We state it here to highlight the role of the properties that we have studied in Sections 4 and 5.

Theorem 5. *Suppose that schools' choice rules satisfy IRS and GS, then DA produces the stable matching that is simultaneously the best stable matching for all students. Suppose, furthermore, that choice rules satisfy LAD, then DA is group incentive compatible for students and each school is matched with the same number of students in any stable matching.*

We refer the outcome of DA, as the student optimal stable matching. Below we compare the outcomes of DA when schools have different choice rules.

Theorem 6. *Consider two substitutable choice rule profiles $(C_c)_{c\in\mathcal{C}}$ and $(C'_c)_{c\in\mathcal{C}}$. Suppose that $C_c(S) \subseteq C'_c(S)$ for every $S \subseteq \mathcal{S}$ and $c \in C$. Let μ and μ' be DA outcomes with $(C_c)_{c\in\mathcal{C}}$ and $(C'_c)_{c\in\mathcal{C}}$, respectively. Then $\mu'(s) \succeq_s \mu(s)$ for all s.*

Therefore, if choice rule C' selects more students than choice rule C, then all students weakly prefer DA with C' to DA with C. In fact, a slightly more general statement is true: Suppose that $(C')_{c\in\mathcal{C}}$ is a substitutable choice rule profile and μ' is the outcome of DA with this choice rule profile. Then if $(C)_{c\in\mathcal{C}}$ is a choice rule profile such that $C_c(S) \subseteq C'_c(S)$ for every $S \subseteq \mathcal{S}$, $c \in C$ and μ is a stable matching with $(C)_{c\in\mathcal{C}}$, then $\mu'(s) \succeq_s \mu(s)$ for all s. In other words, we do not need that $(C)_{c\in\mathcal{C}}$ is a substitutable choice rule profile. We make this assumption to make the statement easier. Otherwise, we need to assume that there exists a stable matching with respect to $(C)_{c\in\mathcal{C}}$.

In particular, we can compare the outcome of DA with choice rules that are generated by soft quotas and hard quotas:

Corollary 1. *Suppose that C_c is generated by hard quotas $(r_t^c)_{t\in T}$ and C'_c is generated by soft quotas $(r_t^c)_{t\in T}$ where $||r^c|| \leq q_c$ for every school c. Let μ and μ' be the DA outcomes with $(C_c)_{c\in\mathcal{C}}$ and $(C'_c)_{c\in\mathcal{C}}$, respectively. Then $\mu'(s) \succeq_s \mu(s)$ for all s.*

This follows directly from Theorem 6 above. Indeed, if student s is rejected by C'_c from S, i.e., if $s \in S \setminus C'_c(S)$, then $\xi(C'_c(S))_{\tau(s)} \geq r_{\tau(s)}$ and for any $s' \in C'_c(S)$ with $\tau(s') = \tau(s)$ we have $s' \succ_c s$. Therefore, s cannot be chosen by choice rule C from S since there are at least $r_{\tau(s)}$ students of type $\tau(s)$ who have higher ranking than s in S.

Corollary also implies that all students weakly prefer DA using Schur-generated choice rules to DA using ideal point generated choice rules with the same reference points. In fact, we can say more than that for particular reference points:

Theorem 7. *Suppose that C_c satisfies GS and distribution-dependence and that $d = 2$. Let $z_c^* = \xi(C_c(\mathcal{S}))$ for each c. Let μ be the DA outcome using choice profile $(C_c)_{c\in C}$, μ^i the DA outcome using choice profile with rules that are generated by ideal points z_c^* for each c, and μ^s be the DA outcome using choice rules that are Schur-generated z_c^* for each c. Then*

$$\mu^s(s) \succeq_s \mu(s) \succeq_s \mu^i(s),$$

for all s.

Therefore, with this particular choice of z^*, we get that students prefer the outcome of DA using the Schur-generated choice rules, to DA using the original choice profile to DA using the ideal point generated choice rules.

Finally, in the next result, we consider matching markets in which schools' choice rules are generated by hard quotas. For this market, we establish a type specific "rural hospital theorem":

Theorem 8. *Suppose that for each school c, choice rule C_c is generated by hard quotas r_c. If there exists a stable matching μ such that $|\mu^t(c)| < r_c^t$ and $|\mu(c)| < q_c$ then for any stable matching μ', $\mu'^t(c) = \mu^t(c)$.*

This result follows immediately from Kojima (2012). He shows that the same conclusion holds when schools preferences satisfy a more general condition called separable with affirmative action constraints.

In particular, if schools' choice rules are generated by ideal points then each school's student distribution is the same in all stable matchings, i.e., the diversity of schools are the same in all stable matchings. Therefore, as long as the the school district implements an algorithm that finds a stable matching, each school incoming student composition in terms of types are exactly the same.

7 Controlled school choice in the US

The legal background on diversity in school admissions is complicated. Since the landmark 1954 *Brown vs. Board of Education* supreme court ruling, which ended school segregation, many school districts have attempted to achieve more integrated schools. The current legal environment is summarized in the 2011 guidelines issued by the US departments of justice and education: "Guidance on the voluntary use of race to achieve diversity and avoid racial isolation in elementary and secondary schools." (There is a separate set of guidelines for college admissions.) We shall not summarize these guidelines here, but suffice it say here that they are perfectly compatible with the theory developed in this paper.

In particular, the "race neutral" approaches described in the guidelines can be carried out through our methods (race neutrality goes into the definition of types). We proceed to briefly describe some of the best know programs in the US.

7.1 Jefferson County

The Jefferson County (KY) School District is prominent in promoting diversity among its schools, and the litigation surrounding its admissions policies serve partly as basis for the guidelines mentioned above. Starting from the early 1970s, the student assignment plan used in the Jefferson County went through major changes. First, in order to avoid segregation, a racial assignment plan was used and students were bused to their schools. In early 1990s a school choice system was implemented, allowing parents to state their preferences. In 1996, schools were required to have between 15 and 50 percent of African-American students. In 2002 a lawsuit was filed against the Jefferson County School District because it had a racial admissions policy. After a litigation process, the case came before the US Supreme Court. The Supreme Court in 2007 ruled in favor of the plaintiffs, and decided that race cannot be the only factor to use for admissions.

Following this ruling, the Jefferson County switched to an assignment plan that considered the socioeconomic status of parents: Using census data, the school district divided the county into two regions and required all schools to have 15 and 50 percent of their students to be from the first region.

The Jefferson County is undergoing yet another change at the moment. The new assignment plan, which was accepted by the school district to be implemented in 2013/14 admissions cycle, divides students into three types: Type 1, Type 2 and Type 3. These types are determined by educational attainment, household income, and percentage of white residents in the census block group that the student lives in. Then each school is assigned a diversity index, defined as the average of student types. The new admissions policy requires each school to have a diversity index between 1.4 and 2.5.

These two assignment policies are in conflict with the gross substitutes (GS) axiom, so it would be incompatible with a school choice plan that would seek to install a stable matching. It should be clear, however, that the rules proposed in our paper can achieve similar objectives to the ones in the current policies, while satisfying gross substitutes.

7.2 Chicago

Chicago Public Schools also strive for diversity (Pathak and Sönmez, 2012; Kominers and Sönmez, 2012). To this end, they have an affirmative action policy that uses socioeconomic status to divide students into four types: Tier 1, Tier 2, Tier 3 and Tier 4. Students who would like to attend selective high schools take a centralized exam that is used to determine a score of each student. Each school allocates 40 percent of their seats to the students with the highest rank and then 15 percent of the seats are allocated to each tier separately. In particular, if a school is divided into two fictitious schools one representing open seats and the other representing the rest, this affirmative action policy can be viewed as the hard quotas / ideal point model as follows: For the first fictitious school, the quota for each type can be the capacity of the school whereas for the second fictitious school the quota for each type can be 15 percent of the total school capacity. In particular, the actual implementation of the affirmative action policy is very similar to this. By Corollary 6, if the Chicago school district switched to a soft quota policy that uses of choice rules generated by soft quotas for the second fictitious school with the same quotas, all students would weakly benefit.[10]

Alternatively, the Chicago system could be modified to satisfy the model of soft

[10] Kominers and Sönmez (2012) has a different counterfactual in which students either rank the fictitious school representing the open slots first or last.

quotas. This would involve first filling the spaces that are reserved to each type, and then having types compete openly for the remaining spaces. Depending on the distribution of scores for each type, the quotas could be calibrated to achieve the effect desired by the school district. The result has the advantage of fitting directly into the existing stable matching school choice mechanism.

8 Controlled School Choice in Other Countries

Affirmative action can be found in many countries around the world (Sowell, 2004). We focus here on affirmative action in school admissions. Some of the affirmative action policies implement preferential policies whereas some of them implement policies based on quotas. The preferential policies resemble the soft quota model that we study above, while policies based on quotas are similar to hard quotas model with regional variations in actual implementation. There are many countries that have similar policies including but not limited to Brazil, China, Germany, Finland, Macedonia, Malaysia, Norway, Romania, Sri Lanka, and the United States. Below we discuss two particular examples: college admissions in India and high school admissions in French-speaking Belgium.

8.1 Indian College Admissions

In India the caste system divides society into hereditary groups, castes ("types" in our model). Historically, it enforces a particular division of labor and power in society, and places severe limits on socioeconomic mobility. To overcome this, the Indian constitution has since 1950 implemented affirmative action. It enforces that the "scheduled castes" (SCs) and "scheduled tribes" (STs) are represented in government jobs and public universities proportional to their population percentage in the state that they belong to. These percentages change from state to state. For example, in Andhra Pradesh, each college reserves 15 percent of its seats for SCs, 6 percent for STs, 35 percent for other 'backward classes' and the remaining 44 percent is left open for all students.

The college admissions to these public schools is administered by the state, and it works as follows. Students take a centralized exam that determines their ranking. Then students are called one by one to make their choices from the available colleges. In each college, first the open seats are filled. Afterwards the reserved seats are filled only by students for whom the seats were reserved. This model corresponds to the situation described above for Chicago. Therefore, this affirmative action policy fits into our hard quotas / ideal point model in which we replace each school with two copies, the first representing the open seats and the second representing the rest. For the first copy of the college, each student is treated the same and the choice rule picks the best available students regardless of their caste. Similar to the Chicago school district, if a soft quota policy was used, all students would be weakly better off.

The choice rule can also be generated by soft quotas in which each quota is greater than the school's capacity. But the second copy of the college implements a choice rule that is generated by the hard quota / ideal point model described in the previous paragraph.[11]

8.2 High Schools in French-Speaking Belgium

In French-speaking Belgium, high school admissions is done to promote diversity. However, in contrast with many examples we have seen thus far, the target of affirmative action policy is the set of students who have attended "disadvantaged primary schools." The administration announces these primary schools, which may change each year depending on supply and demand. Each school is required to reserve at least 15 percent of their seats to students from disadvantaged primary schools, and also some seats for students living in the neighborhood of the school. If a reserved seat for either group cannot be filled then it can also be allocated to other students as long as there is no student from the privileged group willing to take that seat. This choice corresponds to the soft quotas model described above.

9 Auxiliary lemmas

The general approach we follow here is the following: We translate considerations related to distributions into results on functions on $\mathbf{Z}_+^d$. Therefore, most of our results follow from mapping C into a function or a correspondence defined on $\mathbf{Z}_+^d$. In particular, it turns out that some of our axioms have interesting counterparts as properties of such functions and correspondences.

[11]For an empirical study of affirmative action policies in Andhra Pradesh see Bagde, Epple, and Taylor (2011).

9.1 Notational Convention

We first lay out some notational conventions. For $A \subseteq \mathbf{Z}_+^d$, $\partial A \equiv \{z \in A : z' \gg z \Rightarrow z' \notin A\}$, where $z' \gg z$ if and only if $z'_t > z_t$ for all t and, similarly, let $\partial^M A \equiv \{z \in A : z' > z \Rightarrow z' \notin A\}$.

Let e_t denote the unit vector in $\mathbf{Z}_+^d$ with 0 in all its entries except that corresponding to t, in which it has 1.

9.2 Ideal point model

We shall first introduce some simple lemmas related to functions on $\mathbf{Z}_+^d$. Hopefully the discussion is suggestive of how we use these lemmas in proving our main results.

Let $f : \mathbf{Z}_+^d \to \mathbf{Z}_+^d$. We say that f is ***monotone increasing*** if $y \leq x$ implies that $f(y) \leq f(x)$; f is ***within budget*** if $f(x) \in B(x) = \{y : 0 \leq y \leq x, ||y|| \leq q\}$; that f satisfies ***gross substitutes*** if

$$y \leq x \Rightarrow f(x) \wedge y \leq f(y).$$

A function f is ***generated by an ideal point*** if there is $z^* \in \mathbf{Z}_+^d$ such that $||z^*|| \leq q$, and $f(x)$ minimizes the Euclidean distance to z^* among the vectors in $B(x)$.

Lemma 1. *Function f is monotone increasing, within budget and satisfies gross substitutes if and only if it is generated by an ideal point.*

We need the following lemma:

Lemma 2. *Let $z^* \in \mathbf{Z}_+^d$ satisfy $||z^*|| \leq q$. Then $x \wedge z^*$ is the unique minimizer of the Euclidean distance to z^* among the vectors in $B(x)$.*

Proof of Lemma 2. First note that $x \wedge z^* \in B(x)$. The distance from z to z^* is minimized if $\sum_t (z_t - z_t^*)^2$ is minimized. The lemma follows from the observation that one can minimize, for each t, $(z_t - z_t^*)^2$ by setting $z_t = \min\{x_t, z_t^*\}$: when $\min\{x_t, z_t^*\} = z_t^*$ this is trivial, and when $\min\{x_t, z_t^*\} = x_t$ then there are no $z \in B(x)$ with $z_t > x_t$. Since $z_t = \min\{x_t, z_t^*\}$ for every t, we get $z = x \wedge z^*$. $\square$

Proof of Lemma 1. We first show that if f is generated by an ideal point z^* with $||z^*|| \leq q$, then it is monotone increasing, within budget and it satisfies gross substitutes. Suppose that the ideal point is z^*. By Lemma 2, $f(x) = x \wedge z^*$. Then $f(x) \leq x$ and $||f(x)|| \leq ||z^*|| \leq q$, so f is within budget. Next we show monotonicity:

$$y \leq x \Rightarrow y \wedge z^* \leq x \wedge z^* \Rightarrow f(y) \leq f(x).$$

Last we show gross substitutes. Let $y \leq x$. Then,

$$f(x) \wedge y = (x \wedge z^*) \wedge y = (x \wedge y) \wedge z^* = y \wedge z^* = f(y).$$

It will be useful to consider an additional property. We say that f satisfies the ***boundary condition*** if $f(x) \in \partial B(x)$.

We now turn to proving that the axioms are sufficient for generation by an ideal point. We suppose that $f : \mathbf{Z}_+^d \to \mathbf{Z}_+^d$ is a function satisfying monotonicity, gross substitutes, and it is within budget. We show that it must be generated by some ideal point. We consider two different cases, the case when f satisfies the boundary condition and when it does not.

First, suppose that f satisfies the boundary condition. Let $\hat{x}$ such that $\hat{x}_t \geq q$ for all t and $z^* \equiv f(\hat{x})$. Note that $\sum_t z_t^* = q$ because

$$f(\hat{x}) \in \partial B(\hat{x}) = \{z \in \mathbf{Z}_+^d : \sum_t z_t = q\},$$

by the choice of $\hat{x}$. We show that, for all y, $f(y)$ minimizes the distance to z^* in $B(y)$.

Note if $y \leq x$ then the monotonicity of f, and that $f(y) \leq y$ implies that $f(y) \leq y \wedge f(x)$. Thus the gross substitute axiom becomes:

$$y \leq x \Rightarrow f(x) \wedge y = f(y). \tag{1}$$

Now, $\hat{x} \leq \hat{x} \vee y$, so $z^* = f(\hat{x}) \leq f(\hat{x} \vee y)$, as f is monotone increasing. Then $\sum_t z_t^* = q$ and $f(\hat{x} \vee y) \in \partial B(\hat{x} \vee y)$ implies that $z^* = f(\hat{x} \vee y)$. Now, $y \leq \hat{x} \vee y$ and the substitutes condition gives us that

$$f(y) = y \wedge f(\hat{x} \vee y) = y \wedge z^*.$$

By Lemma 2, $f(y)$ minimizes the distance to z^* in $B(y)$.

We finish the proof by considering the case when f does not satisfy the boundary condition. In this case there is z^* such that $f(z^*) \notin \partial B(z^*)$. We shall prove that f is generated by ideal point $f(z^*)$.

Let $x \in \mathbf{Z}^d_+$. Note that $z^* \wedge x \leq z^*$, monotonicity, and gross substitutes, imply (using equation (1)) that

$$f(z^* \wedge x) = (z^* \wedge x) \wedge f(z^*) = (z^* \wedge f(z^*)) \wedge x = f(z^*) \wedge x.$$

Similarly, $z^* \wedge x \leq x$ gives us that

$$f(z^* \wedge x) = (z^* \wedge x) \wedge f(x) = (x \wedge f(x)) \wedge z^* = f(x) \wedge z^*.$$

Thus, $f(z^*) \wedge x = f(x) \wedge z^*$. Now observe that $f(z^*) \notin \partial B(z^*)$ means that $f(z^*) \ll z^*$; so $f(z^*) \wedge x \ll z^*$. Then $f(z^*) \wedge x = f(x) \wedge z^*$ is only possible if $f(z^*) \wedge x = f(x)$. By Lemma 2 $f(x)$ minimizes the distance to $f(z^*)$ in $B(x)$. $\square$

Remark 3. When the ideal point z^* is such that $\sum_t z^*_t = q$, then the ideal point rule also satisfies the boundary condition. To see this, suppose first that $z^* \leq x$. Then $\sum_t z^*_t = q$ implies that $f(x) = x \wedge z^* = z^* \in \partial B(x)$. If, on the other hand, $z^* \nleq x$ then there is t such that $x_t = (x \wedge z^*)_t = f(x)_t$; so if $z \gg f(x)$ then $z \notin B(x)$, as $z_t > f(x)_t = x_t$. Therefore, $f(x) \in \partial B(x)$.

9.3 Schur concavity

We say that $f : \mathbf{Z}^d_+ \to \mathbf{Z}^d_+$ is ***Schur-generated*** if there is $z^* \in \mathbf{Z}^d_+$ such that $||z^*|| \leq q$ and a monotone increasing and concave function $g : \mathbf{R} \to \mathbf{R}$, such that $f(x)$ is a maximizer of $\sum_{t=1}^d g(x_t - z^*_t)$ in the set $B(x)$ for all x. Similarly, f is ***efficient*** if $f(x) \in \partial^M(B(x))$ for all x.

Lemma 3. *A function f is efficient and satisfies gross substitutes if and only if it is Schur-generated.*

Proof. Let f satisfy the two axioms. Let $\hat{x}$ be such that $\hat{x}_t > q$ for all t, and let $z^* \equiv f(\hat{x})$. For any $\alpha \in \mathbf{R}$, let

$$g(\alpha) \equiv \alpha \wedge q + (\alpha - \alpha \wedge q)/2.$$

Notice that g is strictly monotone increasing and concave. Let $\nu(x) \equiv \sum_t g(x_t + q - z_t^*)$. Since g is monotone increasing and concave, so is ν.

We shall prove that $f(y)$ maximizes ν in $B(y)$. To prove this, we show that $f(y) \geq z^* \wedge y$. Note that this suffices because it says that $f(y)_t \geq z_t^*$ when $y_t > z_t^*$, and $f(y)_t = y_t$ (as $f(y) \leq y$) when $y_t \leq z_t^*$; by definition of ν and the axiom of efficiency, ν is maximized by such an f.

Now, that $f(y) \geq z^* \wedge y$ follows from gross substitutes in the case that $y \leq \hat{x}$. Suppose then that $y \nleq \hat{x}$.

First, $y \wedge \hat{x} \leq \hat{x}$ and gross substitutes imply that $f(y \wedge \hat{x}) \geq z^* \wedge (y \wedge \hat{x}) = z^* \wedge y$, as $z^* = z^* \wedge \hat{x}$.

Second, $y \wedge \hat{x} \leq y$ and gross substitutes imply that

$$f(y \wedge \hat{x}) \geq f(y) \wedge (y \wedge \hat{x}) = (f(y) \wedge y) \wedge \hat{x} = f(y) \wedge \hat{x} = f(y).$$

Then $y \nleq \hat{x}$ implies that $||y|| > q$, so efficiency of f implies that $||f(y)|| = q$. Then $f(y \wedge \hat{x}) \geq f(y)$ and $||f(y \wedge \hat{x})|| \leq q$ give us that $f(y \wedge \hat{x}) = f(y)$. We showed above that $f(y \wedge \hat{x}) \geq z^* \wedge y$, so we obtain that $f(y) \geq z^* \wedge y$ as desired. □

10 Proofs from Section 4

10.1 Proof of Theorem 1

Suppose that C satisfies the axioms. We shall prove that it is generated by an ideal point. To this end, we show that there exist an ideal point z^* and a strict priority $\succ$ such that the choice function created by these coincides with C. The result follows essentially from Lemma 1 above.

Define f as follows. For any x, let S be such that $x = \xi(S)$ and let $f(x) = \xi(C(S))$. By distribution-monotonicity we know that $\xi(S) = \xi(S') \Rightarrow \xi(C(S)) = \xi(C(S'))$, so the particular choice of S does not matter; f is well defined. Moreover, when $y \leq x$ we have $f(y) \leq f(x)$, again by distribution-monotonicity. So f is a monotone increasing function. In addition, $f(x) \leq x$, so f is within budget. Let z^* be as defined in the proof of Lemma 1. Since $f(z^*) = z^*$, we have that $|z^*| \leq q$.

Define a binary relation R by saying that $s\,R\,s'$ if $\tau(s) = \tau(s')$ and there is some $S \ni s, s'$ such that $s \in C(S)$ and $s' \notin C(S)$. We shall prove that R is transitive.

Lemma 4. *If C satisfies GS, t-WARP and IRS, then R is transitive.*

Proof. Let $s\,R\,s'$ and $s'\,R\,s''$; we shall prove that $s\,R\,s''$. Let S' be such that $s', s'' \in S'$, $s' \in C(S')$, and $s'' \notin C(S')$. Consider the set $S' \cup \{s\}$. First, note that $s \in C(S' \cup \{s\})$. The reason is that if $s \notin C(S' \cup \{s\})$ then $C(S' \cup \{s\}) = C(S') \ni s'$, by irrelevance of rejected student. Thus $s'\,R\,s$, in violation of the type weak axiom. Second, note that $s'' \notin C(S' \cup \{s\})$, as $s'' \notin C(S')$ and C satisfies gross substitutes. □

The relation R is transitive. Thus it has an extension to a linear order $\succ$ over S. For any S, and any $s, s' \in S$ with $\tau(s) = \tau(s')$ we have that $s \succ s'$ when $s \in C(S)$ while $s' \notin C(S)$.

Lemma 5. *If C satisfies GS then $\xi(\mathcal{S}) \geq x \geq y$ implies $f(x) \wedge y \leq f(y)$.*

Proof. Suppose that C satisfies GS. Let $\xi(\mathcal{S}) \geq x \geq y$ and $S' \subseteq \mathcal{S}$ be such that $\xi(S') = x$. Construct S with $\xi(S) = y$ as follows. If $y_t \geq \xi(C(S'))_t$, then $S^t \supseteq C(S')^t$. However, if $y_t < \xi(C(S'))_t$, then $S^t \subseteq C(S')^t$. In the former case, $C(S)^t \supseteq C(S')^t$ by gross substitutes. In the later case, $C(S)^t = S^t$ by gross substitutes. In both cases, $\xi(C(S))_t \geq \min\{\xi(S)_t, \xi(C(S'))_t\}$, which implies $f(y) \geq f(x) \wedge y$. □

By Lemma 5, C satisfies gross substitutes implies that f satisfies gross substitutes. In addition, f is also monotone increasing and within budget, therefore, f is generated by an ideal point rule with z^* by Lemma 1. Then C is generated by the ideal point z^* and priority order $\succ$.

Conversely, let C be generated by an ideal point z^* and $\succeq$. It is immediate that C satisfies t-WARP. Define f as above. Here, f is well defined because for any S and S' such that $\xi(S) = \xi(S') = x$, $\xi(C(S))$ is the closest vector to z^* among those in $B(x)$ and $\xi(C(S'))$ is the closest vector to z^* among those in $B(x)$. Therefore, $\xi(C(S)) = \xi(C(S'))$ and so f is well defined.

To show that C satisfies distribution-monotonicity, let $y = \xi(S)$ and $x = \xi(S')$ such that $y \leq x$. By Lemma 2, $f(x) = x \wedge z^*$ and $f(y) = y \wedge z^*$. Then, $f(x) = x \wedge z^* \leq y \wedge z^* = f(y)$, and, therefore, $\xi(C(S)) \leq \xi(C(S'))$. Hence, C satisfies distribution-monotonicity.

To show that C satisfies irrelevance of rejected students, let $C(S') \subseteq S \subseteq S'$, $\xi(S') = x$ and $\xi(S) = y$. By construction, $\xi(C(S')) = x \wedge z^*$ and $\xi(C(S)) = y \wedge z^*$. Therefore,

$$\xi(C(S')) \geq \xi(C(S)).$$

On the other hand, by assumption $\xi(C(S')) \leq \xi(S) = y$ and also $\xi(C(S')) = x \wedge z^* \leq z^*$. The last two inequalities imply that

$$\xi(C(S')) \leq y \wedge z^* = \xi(C(S)).$$

Hence, $\xi(C(S')) = \xi(C(S))$, that is the same number of type t students are chosen from S and S' for each t. Since students are chosen according to $\succeq$ and $C(S') \subseteq S \subseteq S'$ we conclude that $C(S') = C(S)$.

To see that C satisfies gross substitutes, let $s \in S \subseteq S'$, $\tau(s) = t$, $\xi(S) = y$ and $\xi(S') = x$. As we have shown above, $f(x) = x \wedge z^*$ and $f(y) = y \wedge z^*$. If $f(y)_t \geq f(x)_t$, then more type t students are chosen in S compared to S'. Since $s \in C(S')$ and by construction of the choice rule we derive that $s \in C(S)$. On the other hand, if $f(y)_t < f(x)_t$, then $f(y)_t < z^*_t$ since $f(x)_t = (x \wedge z^*)_t \leq z^*_t$. Since $f(y)_t = (y \wedge z^*)_t$, we derive that $f(y)_t = y_t$. That means all type t students are chosen from S, so $s \in C(S)$. Hence, C satisfies gross substitutes.

10.2 Proof of Theorem 2

Let C satisfy the axioms. By the distribution-dependence axiom we can define f as in the proof of Theorem 1. We can also define a strict preference $\succeq$ to act as priority order, by the same argument as in the proof of Theorem 1.

Now, by Lemma 5, f satisfies the assumption of gross substitutes for functions. The axiom of efficiency for C implies that f satisfies efficiency for functions. By Lemma 3, f is Schur-generated. This implies that C is Schur-generated.

Conversely, suppose that C is Schur-generated. It is easy to see that C satisfies t-WARP and distribution dependence. We show that C also satisfies GS, IRS and efficiency. For any $x \leq \xi(\mathcal{S})$, let S be such that $\xi(S) = x$ and define $f(x) \equiv \xi(C(S))$. Since C is Schur-generated, f is well defined and Schur-generated. By Lemma 3, f is efficient and satisfies gross substitutes. That f is efficient implies C is efficient.

To see that C satisfies gross substitutes, let $s \in S \subseteq S'$, $\tau(s) = t$, $\xi(S) = y$ and $\xi(S') = x$. Since f satisfies GS, we have

$$\min\{f(x)_t, y_t\} \leq f(y)_t.$$

If $y_t \leq f(x)_t$, then GS implies $y_t \leq f(y)_t$, which is equivalent to $y_t = f(y)_t$. Hence, $s \in C(S)$. On the other hand, if $y_t \geq f(x)_t$, then $f(x)_t \leq f(y)_t$, so more type t students are chosen from S compared to S'. Since $s \in C(S')$ and C satisfies t-WARP, this implies $s \in C(S)$.

To show that C satisfies irrelevance of rejected students, let $C(S') \subseteq S \subseteq S'$, $\xi(S') = x$ and $\xi(S) = y$. Since C satisfies GS, $C(S') \subseteq C(S)$. Suppose that there exists $s \in C(S) \setminus C(S')$. Then $\xi(C(S)) > \xi(C(S'))$ but this is a contradiction since $\xi(C(S))$ maximizes $\sum_{t=1}^{d} g(x_t - z_t^*)$ in $B(\xi(S))$ and $\xi(C(S'))$ maximizes $\sum_{t=1}^{d} g(x_t - z_t^*)$ in $B(\xi(S'))$ with $B(\xi(S')) \supseteq B(\xi(S))$.

11 Proofs from Section 5

11.1 Proof of Theorem 3

Suppose that C satisfies the axioms. We start by showing that C is generated by hard quotas.

Let $r_t \equiv |C(\mathcal{S}^t)|$ and $S \subseteq \mathcal{S}$. First we prove that $|C(S)^t| \leq r_t$. If $r_t = q$, then $|C(S)^t| \leq r_t$ holds trivially. Suppose that $r_t < q$. Since $|\mathcal{S}^t| > q$, there exists $s \in \mathcal{S}^t \setminus C(\mathcal{S}^t)$. By rejection maximality, for every S, $r_t = |C(\mathcal{S}^t)| \geq |C(S)^t|$. Therefore, $|C(S)^t| \leq r_t$.

Let $\succ^*$ be defined as follows: $s \succ^* s'$ if there exists $S \supseteq \{s, s'\}$ such that $s \in C(S)$, $s' \notin C(S)$ and either $\tau(s) = \tau(s')$ or S is ineffective at $\tau(s')$. By the strong axiom of revealed preference, $\succ^*$ has a linear extension $\succ$ to $\mathcal{S}$.

We now show that if $s \in C(S)$, $s' \in S \setminus C(S)$ and $s' \succ s$, then it must be the case that $\tau(s) \neq \tau(s')$ and $\left|C(S)^{\tau(s')}\right| = r_{\tau(s')}$. If $\tau(s) = \tau(s')$, then $s \succ^* s'$ and $s \succ s'$, which is a contradiction with the fact that $\succ$ is an extension of $\succ^*$. So $\tau(s) \neq \tau(s')$. To prove that $\left|C(S)^{\tau(s')}\right| = r_{\tau(s')}$ suppose, towards a contradiction, that $\left|C(S)^{\tau(s')}\right| \neq r_{\tau(s')}$. We shall prove that S is ineffective for $\tau(s')$, which will yield the desired contradiction, as $\succ$ is an extension of $\succ^*$. Note that $\left|C(S)^{\tau(s')}\right| < r_{\tau(s')}$ by definition of r. Let $S' \equiv S^{\tau(s')}$. We consider three cases. First, $|C(S')| = q$ then $|C(S)| < |C(S')|$ (as there is $s \in C(S)$ with $\tau(s) \neq \tau(s')$), so S is ineffective for $\tau(s')$. Second, consider the case when $|C(S')| < q$ and $|C(S')| < |S'|$. Then, by rejection maximality,

$$|C(S')| = r_{\tau(s')} > \left|C(S)^{\tau(s')}\right|.$$

Hence S is ineffective for $\tau(s')$. Third, consider the case when $|C(S')| < q$, and $|C(S')| = |S'|$. Then $|C(S')| > \left|C(S)^{\tau(s')}\right|$, as $s' \in S^{\tau(s')} \setminus C(S)^{\tau(s')}$. Thus S is ineffective for $\tau(s')$, which implies $s \succ^* s'$. Since $\succ$ is a linear extension of $\succ^*$, we get $s \succ s'$, a contradiction.

Finally, we need to show that if $s \in S \setminus C(S)$, then either $|C(S)| = q$ or $\left|C(S)^{\tau(s)}\right| = r_{\tau(s)}$. By rejection maximality, $s \in S \setminus C(S)$ and $|C(S)| < q$ implies that $\left|C(S)^{\tau(s')}\right| \geq \left|C(S')^{\tau(s')}\right|$ for any S'. We already proved that r_t is the supremum of $\{\left|C(\hat{S})^t\right| : \hat{S} \subseteq \mathcal{S}\}$. So we get that $\left|C(S)^{\tau(s')}\right| = r_{\tau(s')}$. Hence, $s \in S \setminus C(S)$ implies either $|C(S)| = q$ or $\left|C(S)^{\tau(s')}\right| = r_{\tau(s')}$.

To finish the proof, suppose that C is generated by hard quotas. Then it is easy to see that C satisfies the strong axiom of revealed preference and rejection maximality. We show that it also satisfies gross substitutes. Suppose that $s \in S \subseteq S'$ and $s \in C(S')$. For each type t, let $S(t; r_t) \subseteq S^t$ be the r_t highest ranked type t students in S (if $|S^t| \leq r_t$ then $S(t; r_t) = S^t$). Define $S'(t; r_t)$ analogously. Since $s \in C(S')$, we have $s \in S'(\tau(s), r_{\tau(s)})$ and the ranking of s in $\cup_t S'(t; r_t)$ is no more than q. Since $S \subseteq S'$, the preceding statements also hold for S instead of S', which implies that $s \in C(S)$.

11.2 Proof of Theorem 4

The proof requires the following lemma.

Lemma 6. *Let C satisfy GS. If $y \in \mathbf{Z}^d_+$ is such that $\hat{f}(y)_t < y_t$ then $\hat{f}(y+e_{t'})_t < y_t + 1_{t=t'}$*

Proof. Let y and t be as in the statement of the lemma. Let S be such that $\xi(S) = y$ and $\xi(C(S))_t < \xi(S)_t = y_t$. Such a set S exists because $\hat{f}(y)_t < y_t$. Let $s' \notin S$ be an arbitrary student with $\tau(s') = t'$. Note that

$$\emptyset \neq S^t \setminus C(S)^t \subseteq (S \cup \{s'\})^t \setminus C(S \cup \{s'\})^t,$$

as $\mathcal{C}$ satisfies GS. Then we cannot have $\xi(C(S\cup\{s'\}))_t = y_t + 1_{t=t'}$ because that would imply $(S\cup\{s'\})^t \setminus C(S\cup\{s'\})^t = \emptyset$. Then

$$y_t + 1_{t=t'} > \xi(C(S\cup\{s'\}))_t \geq \hat{f}(y+e_{t'})_t.$$

$\square$

Using Lemma 6, we can construct the vector r of minimum quotas as follows. Let $\bar{x} = \xi(\mathcal{S})$. The lemma implies that if $\hat{f}(y_t, \bar{x}_{-t})_t < y_t$ then $\hat{f}(y'_t, \bar{x}_{-t})_t < y'_t$ for all $y'_t > y_t$. Then there is $r_t \in \mathbf{N}$ such that $y_t > r_t$ if and only if $\hat{f}(y_t, \bar{x}_{-t}) < y_t$. This uses the assumption we made on the cardinality of $\mathcal{S}^t$, which ensures that $\hat{f}(y)_t < y_t$ if y_t is large enough. Note that we may have $r_t = 0$.

First we prove that $S \subseteq \mathcal{S}$ with $|S^t| \leq r_t$ then $S^t = C(S)^t$. Observe that, for any x and t, $\hat{f}(r_t, x_{-t}) = r_t$. To see this note that if there is x and t such that $\hat{f}(r_t, x_{-t}) < r_t$ then Lemma 6 would imply that $\hat{f}(r_t, \bar{x}_{-t}) < r_t$, in contradiction with the definition of r. In fact, we can say more: For any x, t, and y_t, if $y_t \leq r_t$ then $\hat{f}(r_t, x_{-t}) = r_t$ and Lemma 6 imply that $\hat{f}(y_t, x_{-t}) = y_t$. Therefore, letting $S \subseteq \mathcal{S}$ with $|S^t| \leq r_t$ we have that

$$\left|C(S)^t\right| \geq \hat{f}(y)_t = y_t,$$

where $y = \xi(S)$. Since $y_t = |S^t| \geq |C(S)^t|$ we have that $S^t = C(S)^t$.

Second we prove that, if $|S^t| > r_t$, then $|C(S)^t| \geq r_t$. Let $\tilde{S} = C(S)$. Assume, towards a contradiction, that $\left|\tilde{S}^t\right| < r_t$. Let $S' = \tilde{S} \cup S''$, where $S'' \subseteq S^t \setminus \tilde{S}^t$ is such that $|S''^t| = r_t$. By irrelevance of rejected students, $C(S') = C(S)$. Thus,

$$\hat{f}(\xi(S'))_t \leq \left|C(S')^t\right| = \left|C(S)^t\right| < r_t.$$

Since $\xi(S')_t = |S''^t| = r_t$, we obtain a contradiction with the definition of r_t above.

Consider the following binary relation. Let $s \succ^* s'$ if there is S, at which $\{s\} = \{s, s'\} \cap C(S)$ and $\{s, s'\} \subseteq S$, and either $\tau(s) = \tau(s')$ or $\tau(s)$ is saturated at S. By the adapted strong axiom, $\succ^*$ has a linear extension $\succeq$ to $\mathcal{S} \cup \{\emptyset\}$.

Third we prove that C is consistent with $\succeq$, as stated in the definition. Let $s \in C(S)$ and $s' \in S \setminus C(S)$. If $\tau(s) = \tau(s')$ then $s \succ^* s'$ by definition of $\succ^*$; hence $s \succeq s'$. If $\tau(s) \neq \tau(s')$ then we need to consider the case when $|S^t| > r_t$ where $t = \tau(s)$. The construction of r_t implies that $r_t = \hat{f}(|S^t|, \bar{x}_{-t}) < |S^t|$. Therefore, there exists $S' \subseteq \mathcal{S}$

such that if

$$S' = S^t \cup \left(\cup_{\tilde{t} \neq t} \mathcal{S}^{\tilde{t}} \right)$$

then $S'^t \setminus C(S')^t \neq \emptyset$. Thus t is saturated at S. Since $s \in C(S)$ and $s' \in S \setminus C(S)$, we get $s \succeq s'$, as $\succeq$ extends $\succ^*$.

It remains to show that if C is generated by soft quotas, then it satisfies the axioms. It is immediate that it satisfies efficiency, IRS and adapted-SARP.

To see that it satisfies gross substitutes, let $S \subseteq S'$ and $s \in S \setminus C(S)$. Then $\left|S^{\tau(s)}\right| > r_{\tau(s)}$, so $\left|S'^{\tau(s)}\right| > r_{\tau(s)}$. Moreover, $s \in S \setminus C(S)$ implies that there are $r_{\tau(s)}$ students in $S^{\tau(s)}$ ranked above s. So s could only be admitted at the second step in the construction of C. Let $C^{(1)}(S)$ be the set of students that are accepted in the first step, S^* be the set of students that are considered in the second step and q^* be the number of remaining seats to be allocated in the second step. Again, $s \in S \setminus C(S)$ implies that there are q^* students ranked above s in S^*. Consider the following procedure for S'. In the first step for each t we accept $\xi(C^{(1)}(S))_t$ highest ranked students of type t. And in the second step we consider all remaining students. It is clear that s cannot be admitted in the first step since $S'^{\tau(s)} \supseteq S^{\tau(s)}$ and that there are at least $r_{\tau(s)}$ students ranked above $s \in S^{\tau(s)}$. Moreover, in the second step there are more higher ranked students of each type compared to S^*, so s can also not be admitted in the second step since there are only q^* seats left. If s cannot be admitted with this procedure, then it cannot be in $C(S')$ because for each $t \neq \tau(s)$, $\xi(C^{(1)}(S))_t \leq r_t$. Therefore, $s \in S' \setminus C(S')$.

12 Proofs from Section 6

12.1 Proof of Theorem 6

We start with the following lemma.

Lemma 7. *If C satisfies GS and (c, S) blocks a matching μ, then for every $s \in S \setminus \mu(c)$ $(c, \{s\})$ ν-blocks μ.*

Proof. Since (c, S) blocks μ, we have $S \subseteq C_c(\mu(c) \cup S)$. Let $s \in S \setminus \mu(c)$, by substitutability $s \in C(\mu(c) \cup S)$ implies $s \in C(\mu(c) \cup \{s\})$. Therefore, $(c, \{s\})$ blocks μ. $\square$

We proceed with the proof of Theorem 6.

Since we use two different choice rule profiles and stability depends on the choice rules, we prefix the choice rule profile to stability, individual rationality and no blocking to avoid confusion. For example, we use C-stability, C-individual rationality and C-no blocking.

By Theorem 5, DA produces the stable matching that is the best stable matching for all students simultaneously. Denote the outcome of DA with C and C' by μ and μ', respectively. Since $C_c(\mu(c)) = \mu(c)$ by C-individual rationality of μ by every school c, $C'_c(\mu(c)) \supseteq C_c(\mu(c))$ by the assumption, and $C'_c(\mu(c)) \subseteq \mu(c)$ by definition of the choice rule we get $C'(\mu(c)) = \mu(c)$. Therefore, μ is also C'-individually rational for schools. Since student preference profile is fixed, μ is also C'-individually rational for students. If μ is a C'-stable matching, then μ' Pareto dominates μ since μ' is the student-optimal C'-stable matching. Otherwise, if μ is not a C'-stable matching, then there exists a C'-blocking pair. Whenever there exists such a blocking pair, there also exists a blocking pair consisting a school and a student by Lemma 7. In such a situation, we apply the following improvement algorithm. Let $\mu^0 \equiv \mu$.

Step k Consider blocking pairs involving school c_k and students who would like to switch to c_k, say $S^k_{c_k} \equiv \{s : c_k \succ_s \mu^{k-1}(s)\}$. School c_k accepts $C'(\mu^{k-1}(c_k) \cup S^k_{c_k})$ and rejects the rest of the students. Let $\mu^k(c_k) \equiv C'(\mu^{k-1}(c_k) \cup S^k_{c_k})$ and $\mu^k(c) \equiv \mu^{k-1}(c)$ for $c \neq c_k$. If there are no more blocking pairs, then stop and return μ^k, otherwise go to Step $k+1$.

We first prove by induction that no previously admitted student is ever rejected in the improvement algorithm. For the base case when $k = 1$ note that $C'(\mu(c_1) \cup S^1_{c_1}) \supseteq C(\mu(c_1) \cup S^1_{c_1})$ by assumption and $C(\mu(c_1) \cup S^1_{c_1}) = \mu(c_1)$ since μ is C-stable. Therefore, $C'(\mu(c_1) \cup S^1_{c_1}) \supseteq \mu(c_1)$, which implies that no students are rejected at the first stage of the algorithm. Assume, by mathematical induction hypothesis, that no students are rejected during Steps 1 through $k-1$ of the improvement algorithm. We prove that no student is rejected at Step k of the algorithm. There are two cases to consider.

First, consider the case when $c_n \neq c_k$ for all $n \leq k-1$. Since μ is C-stable, we have $C(\mu(c_k) \cup S^k_{c_k}) = \mu(c_k)$ (as students in $S^k_{c_k}$ prefers c_k to their schools in μ). By assumption, $C'(\mu(c_k) \cup S^k_{c_k}) \supseteq C(\mu(c_k) \cup S^k_{c_k})$ which implies $C'(\mu(c_k) \cup S^k_{c_k}) \supseteq \mu(c_k)$. Since $\mu(c_k) \supseteq \mu^{k-1}(c_k)$ we have $C'(\mu^{k-1}(c_k) \cup S^k_{c_k}) \supseteq \mu^{k-1}(c_k)$ by substitutability. In this case no student is rejected at Step k.

Second, consider the case when $c_k = c_n$ for some $n \leq k-1$. Let n^* be the last step smaller than k in which school c_k was considered. Since each student's match is either the same or improved at Steps 1 through $k-1$, we have $\mu^{n^*-1}(c_k) \cup S_{c_k}^{n^*} \supseteq \mu^{k-1}(c_k) \cup S_{c_k}^{k}$. By construction $\mu^{n^*}(c_k) = C'(\mu^{n^*-1}(c_k) \cup S_{c_k}^{n^*})$ which implies $\mu^{k-1}(c_k) \subseteq C'(\mu^{k-1}(c_k) \cup S_{c_k}^{k})$ by substitutability and the fact that $\mu^{n^*}(c_k) \supseteq \mu^{k-1}(c_k)$ (since n^* is the last step before k in which school c_k is considered). Therefore, no student is rejected at Step k.

Since no student is ever rejected by the improvement algorithm, it ends in a finite number of steps. Moreover, the resulting matching does not have any C'-blocking pair. By construction, it is also C'-individually rational. This shows that there exists a C'-stable matching that Pareto dominates μ. Since μ' is the student-optimal C'-stable matching, we have that μ' Pareto dominates μ for students.

12.2 Proof of Theorem 7

For each school c, let f be defined as in the proof of Theorem 1 for choice rule profile C_c: such f is well defined because C_c satisfies distribution-dependence. Similarly, let f^i be the corresponding function in the ideal point model, given ideal point z_c^*; and let f^s be the f corresponding function in the Schur model, given parameter z_c^*.

Let S be a set of students and $y \equiv \xi(S) \leq \xi(\mathcal{S})$ be the type distribution of S. Then, by gross substitutes of C_c and Lemma 2 we have that

$$f^i(y) = y \wedge z^* \leq f(y).$$

Moreover,

$$f(y) \leq f^s(y)$$

is implied by the fact that $f^s(y) \in \partial^M B(y)$ because (a) if $y_t \leq z_t^*$ then $f(y)_t = f^s(y)_t$, and (b) if $y_t > z_t^*$ then can choose $f^s(y) \in \partial^M B(y)$ that max. ν and satisfies $f(y)_t = f^s(y)_t$.

Since $f^i(y) \leq f(y) \leq f^s(y)$, the conclusion follows from Theorem 6.

References

Abdulkadiroğlu, A., P. Pathak, A. Roth, and T. Sönmez (2005): "The Boston Public School Match," *American Economic Review*, 95(2), 368–371.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review*, 95(2), pp. 364–367.

Abdulkadiroğlu, A., and T. Sönmez (2003): "School choice: A mechanism design approach," *American Economic Review*, 93(3), 729–747.

Alkan, A. (2002): "A class of multipartner matching models with a strong lattice structure," *Economic Theory*, 19(4), 737–746.

Alkan, A., and D. Gale (2003): "Stable schedule matching under revealed preference," *Journal of Economic Theory*, 112(2), 289 – 306.

Alves, M. J., and C. V. Willie (1987): "Controlled choice assignments: A new and more effective approach to school desegregation," *The Urban Review*, 19, 67–88.

Aygün, O., and T. Sönmez (2012): "Matching with Contracts: The Critical Role of Irrelevance of Rejected Contracts," Working Paper.

Bagde, S., D. Epple, and L. J. Taylor (2011): "Dismantling the Legacy of Caste: Affirmative Action in Indian Higher Education," Working Paper.

Echenique, F., and M. B. Yenmez (2007): "A Solution to Matching with Preferences over Colleagues," *Games and Economic Behavior*, 59(1), 46–71.

Ehlers, L., I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim (2011): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," mimeo, Carnegie Mellon University.

Ehlers, L., and Y. Sprumont (2008): "Weakened WARP and top-cycle choice rules," *Journal of Mathematical Economics*, 44(1), 87–94.

Erdil, A., and T. Kumano (2012): "Prioritizing Diversity in School Choice," Working paper.

Fleiner, T. (2003): "A Fixed-Point Approach to Stable Matchings and Some Applications," *Mathematics of Operations Research*, 28(1), 103–126.

Foster, J., and A. Sen (1997): *On economic inequality*. Oxford University Press.

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

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2011): "Effective affirmative action in school choice," Forthcoming, Theoretical Economics.

Hatfield, J., and F. Kojima (2008): "Matching with contracts: Comment," *The American Economic Review*, 98(3), 1189–1194.

Hatfield, J., and P. Milgrom (2005): "Matching with Contracts," *American Economic Review*, 95(4), 913–935.

Hatfield, J. W., and F. Kojima (2009): "Group incentive compatibility for matching with contracts," *Games and Economic Behavior*, 67(2), 745–749.

Kelso, A. S., and V. P. Crawford (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 50, 1483–1504.

Kojima, F. (2010): "School choice: Impossibilities for affirmative action," Forthcoming, Games and Economic Behavior.

Kojima, F. (2012): "The "rural hospital theorem" revisited," *International Journal of Economic Theory*, 8(1), 67–76.

Kominers, S. D., and T. Sönmez (2012): "Designing for Diversity: Matching with Slot-Specific Priorities," Working Paper.

Marshall, A., I. Olkin, and B. Arnold (2010): *Inequalities: theory of majorization and its applications*. Springer.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory*. Oxford University Press, New York.

Moulin, H. (1991): *Axioms of cooperative decision making*, vol. 15. Cambridge University Press.

Pathak, P., and T. Sönmez (2012): "School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation," forthcoming, American Economic Review.

Pycia, M. (2012): "Stability and Preference Alignment in Matching and Coalition Formation," *Econometrica*, 80(1), 323–362.

Richter, M. K. (1966): "Revealed Preference Theory," *Econometrica*, 34(3), 635–645.

Roth, A., and M. Sotomayor (1990): *Two-sided Matching: A Study in Game-Theoretic Modelling and Analysis*, vol. 18 of *Econometric Society Monographs*. Cambridge University Press, Cambridge England.

Roth, A. E. (1984): “Stability and Polarization of Interests in Job Matching,” *Econometrica*, 52(1), 47–57.

——— (2008): “Deferred Acceptance Algorithms: History, Theory, Practice and Open Questions,” *International Journal of Game Theory*, 36(3), 537–569.

Sowell, T. (2004): *Affirmative Action Around the World: An Empirical Study*. New Haven: Yale University Press.

Theil, H. (1967): *Economics and information theory*. North-Holland Amsterdam.

Weitzman, M. L. (1992): “On diversity,” *The Quarterly Journal of Economics*, 107(2), 363–405.

Westkamp, A. (2010): “An analysis of the German university admissions system,” Forthcoming, Economic Theory.