# Essays on Matching Markets

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Alexander Westkamp
aus Bonn

Bonn 2009

Dekan: Prof. Dr. Christian Hilgruber

Erstreferent: Prof. Dr. Benny Moldovanu

Zweitreferent: Prof. Paul Heidhues, Ph. D.

Tag der mündlichen Prüfung: 21.07.2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
(http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

# Acknowledgments

In writing this thesis I have greatly benefited from the help and encouragement from a number of people.

First and foremost, I appreciate the guidance and support of Benny Moldovanu who has been a great supervisor. Without him this thesis would not have come into being as it is. I want to thank him for valuable perspectives on my work and economic theory in general. Paul Heidhues agreed to be the second supervisor of this thesis (Thank You!) and I hope it is sufficiently well written to make it an enjoyable read for him. Konrad Mierendorff has been my office mate for several years and I owe special thanks to him for enduring earlier drafts of parts of this thesis, many insightful discussions, and companionship during my graduate student years. Furthermore, at Bonn I want to thank Urs Schweizer, for his enduring support and his continuous efforts to manage and run the graduate school, and Thomas Gall, for a helpful discussion about the second chapter of this thesis.

Outside of Bonn, I am very grateful to Al Roth for many discussions and his support during my academic year at Harvard and beyond. During my stay abroad I also met Lars Ehlers, who has become not only a co-author (the second chapter of this thesis is based on a joint project), but also a friend. It was a great pleasure to work with him on our joint projects and I benefitted a great deal from his comments on my other research projects. Other economists that I had the pleasure to talk to and discuss my as well as their work with are Peter Coles, Itay Fainmesser, Bettina Klaus, Fuhito Kojima, Markus Mobius, and Utku Unver.

Outside of the academic world I am most grateful to my parents, my wonderful wife, and my grandparents who have always supported me in all I do. Whenever I had doubts about my work, they encouraged me and gave me the firm belief that somehow “everything would come together”. I guess the completion of this thesis proves them right.

Finally, I want to thank my fellow students in the BGSE and my friends for making life outside academics worthwile.

# Contents

# Introduction

Matching is a central problem to economics: Workers need to be matched to firms, objects need to be allocated to bidders in multi-object auctions, students need to be matched to public schools or universities, and so on. Following the seminal paper of Gale and Shapley (1962), theoretical models have provided important, practically relevant, insights into the strategic structure of matching markets and have uncovered important links between the above problems. While the literature has long moved beyond Gale and Shapley's famous model of a monogamous marriage market, their ideas have remained central to the literature. The central element of theoretical matching models is the cooperative solution concept of *pairwise stability*. Roughly speaking, this postulates that a matching of agents can persist in a market of self-interested agents if and only if there is no pair of agents who are not matched to each other but who would mutually prefer to form a partnership. Empirical and experimental evidence, Roth and Sotomayor (1991) is an excellent survey, suggests that stability may be a key determinant for the success and longevity of market mechanisms. Thus, an important practical question is how market rules have to be designed in order to achieve stable outcomes. For an important class of matching models, a simple and intuitive class of procedures, the *deferred acceptance algorithms*,[1] always find stable matchings and can thus be seen as a blueprint for stable market rules.

In recent years, this theory has been successfully applied to the design of centralized matching institutions.[2] Following the advice of matching theorists some of these markets have replaced malfunctioning (centralized) assignment procedures with variants of the deferred acceptance algorithms. Within the realm of the theoretical models that had been studied, these algorithms ensured not only that the outcome was stable with respect to the reported preferences, but also that revealing preferences truthfully was a dominant strategy for (some of the) agents. However, although the markets to which the theory was supposed to be applied seemed reasonably close

[1]For a survey of the theoretical and applied history of these algorithms see Roth (2008).

[2]For a survey of earlier design efforts see Roth (2002). A more recent survey, which includes the design of school choice systems and centralized exchanges for live donor organ transplants, is Sonmez and Unver (2008).

to existing theoretical models, the applied literature has often encountered complex constraints and problems that had largely been ignored by the theoretical literature. For example, one of the major challenges in Roth and Peranson (1999)'s effort to redesign the matching market for medical students in the United States was to adopt the deferred acceptance algorithm to take into account that some students are in a relationship with each other and desire a *pair* of positions not too far apart.[3] This is one of many examples - we will encounter more below - showing that applied matching-market design often requires tailoring the simple and intuitive concepts of theoretical models to the complex realities at hand. As Roth (2002) [p.1342] argues, *[...] we need to foster a still unfamiliar kind of design literature in economics, whose focus will be different than traditional game theory and theoretical mechanism design* In particular the first two chapters of this thesis contribute to this research agenda.

The first chapter of this thesis analyzes the German university admissions system, where places for medicine and related subjects are assigned via a centralized clearinghouse. The system has to deal with three conflicting goals that are dictated by policymakers and legal constraints: First, applicants who did exceptionally well in high-school should be given relative freedom in choosing a university. Secondly, applicants who have unsuccessfully participated in the procedure several times must be given a chance of admission. Finally, universities should be able to evaluate applicants according to their own criteria. The current assignment procedure adapts to these goals by dividing the capacity of each university into three parts and assigning places sequentially starting with those places reserved for excellent high-school graduates. Interestingly, the procedure is based on the well known Boston and College Optimal Stable Matching Mechanisms. We argue that this system induces a very complicated revelation game for applicants. In particular, some applicants face a difficult trade-off between securing a match early in the procedure and taking the risk of participating in later parts of the procedure to obtain a more preferred university. Assuming that universities do not act strategically we derive a characterization of complete information Nash equilibrium outcomes. Our main result is that the set of equilibrium outcomes coincides with the set of matchings that satisfy a notion of stability adapted to the constraints of the German admissions system. A major problem of the current procedure is that it supports outcomes which are Pareto dominated with respect to applicants' preferences. As a response we develop a variant of the student proposing deferred acceptance algorithm that allocates all places simultaneously while maintaining the

[3]This *couples problem*, for which the set of stable matchings can be empty, has spawned a literature of its own. For example, Klaus and Klijn (2005) derive conditions under which existence of a stable matching is guaranteed in couples problems.

three quotas of the current system. This algorithm relies on a transformation of the German admissions system to a related college admissions problem. We show that despite the complex institutional constraints this related problem is “well behaved” so that known results from the theory of two-sided matching apply. In particular, our version of the student proposing deferred acceptance algorithm is *strategy-proof for applicants*, i.e. makes it a dominant strategy for them to report their preferences truthfully, and Pareto dominates any equilibrium outcome of the current procedure with respect to the true preferences of applicants. Outside the specific context of the German system, we discuss how our approach can be used to implement affirmative action constraints in school choice problems.

In the second chapter, which is based on joint work with Lars Ehlers, we study the school choice problem with indifferences in priority orders. In this problem a set of students has to be assigned among a set of public schools. Each school has an exogenously given priority ordering of students that e.g. represents political or social preferences about who should be given prioritized access to the school. In this context, stability (with respect to student preferences and school priorities) can be understood as a fairness criterion which ensures that no student ever envies another student for a school at which she has higher priority. If students can never have identical priority for a given school, as is likely if priorities are based on e.g. scores from centralized exams, the student proposing deferred acceptance algorithm produces a student optimal stable matching and is strategy-proof (Gale and Shapley (1962),Dubins and Freedman (1981),Roth (1982)). However, large indifference classes in priority structures are the rule rather than an exception in real-life school choice problems. For these problems, a matching mechanism has to specify how *ties* between equal priority students are broken. The main problem here is that tie-breaking introduces additional stability constraints to the problem which can lead to decreased student welfare. A counterexample in Erdil and Ergin (2008) shows that, in some instances, any strategy-proof and stable mechanism incurs additional welfare loss due to tie-breaking. An important question is whether this negative result is an exception or the rule in school choice problems with indifferences in priority orders. We call a priority structure *solvable*, if there is a strategy-proof (for students) matching mechanism that never incurs any welfare loss due to tie-breaking. In the second chapter we introduce a model in which schools are either completely indifferent between all students or have a strict ordering of students. One interpretation of this model is that strict orderings arise from subject tests at *specialized schools*, while *non-specialized schools* offer general educational training and therefore do not discriminate between students. The model is a natural starting point for analyzing matching with indifferences and provides important insights into problems with large indifference classes. Furthermore, this model presents a unified perspective on two problems that have been studied extensively in the literature. Within the *(non-)specialized schools* model, we analyze when a strategy-proof and stable matching mechanism exists that never incurs any welfare loss due to tie-breaking. Our main results relate the existence of such mechanisms to the priority structure of specialized schools. For the case where no school can admit more than one student, we provide a full characterization of solvable priority structures. Of course, schools can typically admit more than one student and we derive weaker sufficient conditions for solvability in case of general capacity vectors. The conditions are easy to test and connect the capacity vector with the *amount of allowable variability* in the priority structure of specialized schools. Our existence proofs are constructive and are based on a new version of the student proposing deferred acceptance algorithm with preference-based tie-breaking. In particular, our results show that there is often scope for preference based tie-breaking and it is not sufficient to restrict attention to exogenous tie-breaking rules.

The first two chapters of this thesis are concerned with a class of two-sided matching markets in which only one side of the market (universities/schools) can be matched to more than one partner. Abstracting from the above applications, the theory of these *many-to-one two-sided* matching markets excludes a wealth of interesting applications that one might want to study using the tools of matching theory. For example, in labor markets it is not uncommon that workers are looking for several jobs, i.e. not all two-sided matching markets of interest need to be many-to-one. Furthermore, in many matching markets intermediaries facilitate exchange or partnership formation between agents, i.e. markets may fail to be two-sided. Recently, Ostrovsky (2008) introduced the *supply chain model*, which allows for these features. In this model agents are located in an exogenously given vertically ordered network[4] and have preferences over sets of trading relationships, or *contracts*, with their neighbors. Ostrovsky showed that a generalized notion of pairwise stability, called *chain stability*, can be satisfied for a natural domain of preferences. This existence result suggests that this new stability concept could play an important role in extending the theory of two-sided matching markets. However, unlike pairwise stable matchings in the two-sided models considered above, chain stable allocations may not be immune to all coordinated deviations and can even fail to be efficient. This is a major obstacle for extending the theory of two-sided matching since it questions the *cooperative foundation* of Ostrovsky's stability concept. In the third chapter, we

[4]Here, a vertically ordered network means a directed graph of connections between the agents that has no *directed cycles*.

take his basic model as given and analyze the relationship between chain stability, efficiency, and some important competing concepts of stability. In a first step, we characterize the largest class of supply chain models for which chain stable allocations are efficient and immune to all coalitional deviations. The characterization is based on properties of the exogenously given network structure that agents interact in and our main condition rules out certain kinds of trading cycles. A major difference to most other papers in the literature is that we do not impose additional restrictions on preferences but work with the most general domain of preferences for which the existence of a chain stable allocation is known. A major benefit of our approach is that we are able to derive two justifications for the use of chain stability in the unrestricted model: First, whenever a minimal stability requirement can (always) be reconciled with efficiency, chain stable outcomes are also guaranteed to be efficient. Second, if chain stable outcomes fail to be immune to some coalitional deviations, there does not (in general) exist any outcome that is immune to all coalitional deviations. The relationship between chain stability and the classical (cooperative) solution concept of the core is also studied. We characterize the largest class of supply chain models for which these two concepts yield identical predictions. Examples show that this class is strictly smaller than the class for which chain stable outcomes are efficient and immune to any coalitional deviation. Before proceeding, it is important to stress that in contrast to the first two chapters, the third chapter takes an entirely cooperative game theory view of the economy. However, we believe that our results lay the foundation for future studies that focus on the non-cooperative implementation of chain stable allocations as they provide a cooperative rationale for using this stability concept.

Before proceeding to our contributions to the literature, the next two sections formally introduce the basic language and terminology of (two-sided) matching theory, and summarize most of the classical results. Readers proficient in this theory may want to skip these sections but we hope that they provide useful compendium to the three main chapters of this thesis.

## I.1 The College Admissions Problem with Responsive Preferences

In a *college admissions problem* (Gale and Shapley (1962)) two finite sets of students and colleges have to be matched to each other. Each student is interested in receiving a place at one of the colleges. Each college has a fixed upper bound on the maximal number of students it can admit. Students have preferences over available colleges and the option of not attending a college. Colleges have preferences over entering classes of students. In this section we assume that college preferences over groups of students are *responsive* (Roth (1985)) to a ranking of individual students. More formally, a college admissions problem with responsive preferences consists of

- a finite set of students $I$
- a finite set of colleges $C$,
- a capacity vector $(q_c)_{c \in C}$
- a profile of strict student preferences $R_I = (R_i)_{i \in I}$[5] and
- a profile of strict college preferences $R_C = (R_c)_{c \in C}$.

We write $cR_ic'$ if $i$ *weakly prefers* college $c$ over college $c'$ and $cP_ic'$ denotes that $i$ *strictly prefers* $c$ over $c'$ (i.e. $cR_ic'$ and $c \neq c'$). We denote by $cP_ii$ that $i$ strictly prefers being assigned to $c$ over not receiving a place at any university. In this case we say that college $c$ is *acceptable* to student $i$. Sometimes we write preferences in the form $P_i : c_1, \ldots, c_k$, which means that $c_lP_ic_{l'}$ for all $l < l' \leq k$ and that $i$ finds only colleges $c_1, \ldots, c_k$ acceptable.

The notation for preferences of colleges is exactly the same as for the students. Apart from having a strict preference relation $R_c$ over individual students (and the option of leaving a place unfilled), college $c$ has a strict ranking $R_c^\#$ over subsets of $I$. For this section we assume that $R_c^\#$ is *responsive* (Roth (1985)) to its ranking of individual students $R_c$: If $J \subset I$ and $i, j \in I \setminus J$ then

(i) $J \cup \{i\} P_c^\# J \cup \{j\}$ if and only if $iP_cj$, and
(ii) $J \cup \{i\} P_c^\# J$ if and only if $iP_cc$.

Note that there may be several preferences over groups of students that are responsive to the same ranking of individual students. However, for our purpose it does not matter which *responsive extension* of the ranking of individual students is used. This is the reason for including the ranking of individual students and not the preferences over groups in the formulation of the problem.

A *matching* is an assignment of students to colleges that respects capacity constraints of colleges. More formally, a matching is a mapping $\mu$ from $I \cup C$ into itself such that (i) $\mu(i) \in C \cup \{i\}$ for all $i \in I$, (ii) $\mu(c) \subseteq I$ and $|\mu(c)| \leq q_c$ for all $c \in C$, and (iii) $i \in \mu(c)$ if and only if $\mu(i) = c$. Student $i$ is *unassigned under matching* $\mu$ if $\mu(i) = i$. We assume throughout that agents only care about their own partner(s) in a matching so that their preferences over

[5] More formally, for all students $i \in I$, $R_i$ is a complete, reflexive, transitive, and antisymmetric binary relation on $C \cup \{i\}$. The same remark applies to college preferences.

matchings are congruent with their preferences over potential partners.[6] The sets of students and colleges as well as the capacity vector are assumed to be fixed so that we can think of a college admissions problem with responsive preferences as being given by a profile of student and college preferences $R = (R_I, R_C)$.

A main interest of matching theory is to predict which matchings will occur when self-interested agents form partnerships. The key concept in the literature in this respect is *(pairwise) stability* as introduced by Gale and Shapley (1962). Given a college admissions problem $R$, a matching $\mu$ is pairwise stable if

(i) no student is matched to an unacceptable college, that is, $\mu(i) R_i i$ for all $i \in I$,

(ii) no college prefers to reject some of its assigned students, that is, for all $c \in C$, $i P_c c$ for all $i \in \mu(c)$, and

(iii) there is no student-college pair that blocks $\mu$, that is, there is no pair $(i, c)$ such that $c P_i \mu(i)$ and either $i P_c j$ for some $j \in \mu(c)$ or $i P_c c$ and $|\mu(c)| < q_c$.

If a matching was not stable, we would expect agents to act upon their incentives to form new partnerships and *block* the matching. Note that stability is a cooperative solution concept which remains agnostic as to how the market is supposed to reach such an equilibrium. For the domain of responsive preferences a stable matching always exists. Furthermore, pairwise stability is equivalent to *core stability* (Roth and Sotomayor (1991)) so that there is no group of agents who can block a pairwise stable matching.[7] In particular, a stable matching is efficient with respect to the preferences of students *and* colleges. Two stable matchings are of central interest to matching theory which can be found by applying the two variants of the *deferred acceptance algorithm* introduced by Gale and Shapley (1962). This class of algorithms is central to the theory of two-sided matching and also provides an important point of departure for the first two chapters of this thesis.[8]

### The Student Proposing Deferred Acceptance Algorithm

Given a profile of student and college preferences the *student proposing deferred acceptance algorithm* (SDA) proceeds as follows.

In the first round, every student applies to her favorite acceptable college. For each

[6]For matching models with externalities see Dutta and Masso (1997) and Echenique and Yenmez (2007).

[7]More precisely, the set of pairwise stable matchings coincides with the *core defined by weak domination.* Here, a matching is in the core, if there is no group of agents who can obtain a matching that all agents involved weakly (and at least one strictly) prefer(s) by forming partnerships only among themselves.

[8]An excellent survey of the theoretical and applied history of the deferred acceptance algorithms is Roth (2008).

college $c$, the $q_c$ most preferred acceptable students (or all acceptable students if there are fewer than $q_a$) are placed on the waiting list of $c$ and the rest are rejected.

In the $t$th round, those applicants who were rejected in round $t-1$ apply to their next best acceptable college. For each college $c$, the $q_c$ most preferred acceptable students among the new applicants and those in the waiting list are placed on the new waiting list of $c$ and the rest are rejected.

The algorithm ends when all unmatched students have proposed to all acceptable colleges. Only at this point are assignments finalized (hence the term deferred acceptance). Given a college admissions problem $R$, let $f^I(R)$ denote the matching chosen by the SDA. As shown by Gale and Shapley (1962) the matching $f^I(R)$ is the unanimously most preferred stable matching for students and the unanimously least preferred stable matching for colleges: If $\tilde{\mu}$ is any other stable matching for the college admissions problem $R$, then $f_i^I(R)R_i\tilde{\mu}(i)$ for all students $i \in I$, and $\tilde{\mu}(c)R_c^{\#}f_c^I(R)$ for all colleges $c \in C$.[9] The next algorithm reverses the roles of students and colleges in the deferred acceptance procedure.

**The College Proposing Deferred Acceptance Algorithm**

Given a profile of student and college preferences the *college proposing deferred acceptance algorithm* (CDA) proceeds as follows

In the first round, every college offers admission to its $q_c$ most preferred acceptable students. Each student $i$ temporarily holds on to her most preferred offer and rejects all other offers.

In the $t$th round, every college that had $k$ of its offers rejected in round $t-1$ offers admission to the $q_c - k$ most preferred acceptable students that have not rejected one of its offers in earlier rounds. Each student $i$ temporarily holds on to her most preferred offer among the one she was holding at the end of round $t-1$ and the ones she receives in round $t$.

The algorithm ends when all colleges with unfilled capacity have offered admittance to all acceptable students. Only at this point are assignments finalized. Let $f^C(R)$ denote the matching chosen by the CDA for the college admissions problem $R$. This matching has diametrically

[9]It does not matter which responsive extension of $R_c$ is used in this comparison since they all yield the same ranking of stable matchings (Roth and Sotomayor (1989)).

opposed properties to the SDA in the sense that $f^C(R)$ is the college optimal and student pessimal stable matching given $R$.

An important strand of the matching literature is concerned with the design of centralized clearinghouses for matching markets. A centralized matching institution can be thought of as a *(deterministic) matching mechanism* that collects preferences from the agents to determine a matching. More formally, a matching mechanism is a mapping $f$ that associates a matching to each college admissions problem $R$.[10] Given a college admissions problem $R$, $f_i(R)$ denotes the college assigned to student $i \in I$ by $f$ (if any). Similarly, $f_c(R)$ denotes the set of students assigned to college $c \in C$. A matching mechanism is *stable*, if it selects a stable matching for each college admissions problem. We have already encountered two stable matching mechanisms above: $f^I$ is the student optimal stable matching mechanism (SOSM) and $f^C$ is the college optimal stable matching mechanism (COSM).

Given that preferences about potential partners are typically private information, a matching mechanism has to provide participants with the right incentives to reveal their private information. Ideally, it should be in a participant's best interest to submit her true preferences irrespective of her expectations about the behavior of others. A mechanism $f$ is *strategy-proof* if there is no college admissions problem for which some student or college can benefit from misrepresenting preferences. More formally, this requires that for all college admissions problems $R$, $f_i(R) R_i f_i(R'_i, R_{-i})$ for all $i \in I$ and all $R'_i$, and $f_c(R) R^{\#}_c f_c(R'_c, R_{-c}))$ for all $c \in C$ and all $R'_c$. Unfortunately, a result by Roth (1982) shows that a stable mechanism cannot always provide all participants with dominant strategy incentives to reveal their true preferences.[11][12] However, in some applications the abilities to misrepresent preferences are not symmetric between the two market sides. For example, colleges often base their admission decisions on verifiable student characteristics, e.g. performance in standardized tests, and thus have little scope for strategic manipulation once their admission criteria have been announced. For such applications, the following result by Dubins and Freedman (1981) and Roth (1982) is particularly useful: The student optimal stable mechanism $f^I$ is *strategy-proof for students*, that is, for all college admis-

[10] Note that we restrict attention to mechanisms that elicit only a ranking of individual students from colleges. Given that we concentrate on stable mechanisms and the case of responsive preferences, this restriction is innocuous since the set of stable matchings only depends on the ranking of individual students.

[11] For a general class of matching problems that includes the marriage model as a special case, Sonmez (1999) shows that a strategy-proof and stable matching mechanism exists if and only if the set of stable matchings is a singleton.

[12] Several authors have studied weaker incentive compatibility concepts: See Kara and Sonmez (1996), Kara and Sonmez (1997), and Sonmez (1997) as well as the references therein for results on the Nash implementability of (subsolutions of) the stable rule. See Alcalde and Romero-Medina (2000) for two simple sequential mechanism that implement the set of stable matchings in subgame perfect equilibrium.

sions problems $R$, $f^I_i(R) R_i f^I(R'_i, R_{-i})$ for all $i \in I$ and all $R'_i$.[13] Roth (1985) shows that there is no analogous result for colleges. In particular, it may sometimes be beneficial for colleges to submit a false ranking of individual students to the COSM.[14]

## I.2 The School Choice Problem with Strict Priorities

A *school choice problem* (Abdulkadiroglu and Sönmez (2003)) is conceptually almost identical to the college admissions problem. The main and only difference is that instead of having preferences over entering classes of students, colleges, or schools as they will be called from now on, are exogenously assigned a *priority ordering* of students. This priority ordering may result for example from test scores or social criteria such as distance from a school. More formally, a school choice problem consists of

- a finite set of students $I$,
- a finite set of schools $S$,
- a vector of capacities $q = (q_c)_{c \in C}$,[15]
- a profile of strict priority orders of schools $\succ = (\succ_s)_{s \in S}$, and
- a profile of strict student preferences $R = (R_i)_{i \in I}$.[16]

Everything but the priority orderings has exactly the same interpretation as in the college admissions problem. The priority ordering of school $s$, $\succ_s$, is a strict ordering of $I$. For two students $i, i' \in I$, $i \succ_s i'$ denotes that $i$ has *strictly higher priority* for $s$ than $i'$. For example, if schools assign priorities according to distance then $i \succ_s i'$ means that $i$ lives closer to $s$ than $i'$. The sets of students and schools, the capacity vector, and the priority structure of schools are assumed to be fixed so that we can think of a school choice problem as being given by a profile of student preferences $R$. A matching is defined precisely as in the college admissions problem[17] and a matching mechanism is a mapping that assigns a matching to each school

[13] Alcalde and Barbera (1994) show that the SOSM is the only stable mechanism that is strategy-proof for students.

[14] For these results it does not matter whether colleges are allowed to state their full preferences over subsets of students or only their ranking of individual students. The counterexample in Roth (1985) shows that colleges can manipulate even when a mechanism elicits only rankings of individual students.

[15] Many authors require that the total available capacity is greater than the number of students (cf Abdulkadiroglu and Sönmez (2003)). It is inconsequential for the results below whether this assumption is satisfied or not.

[16] It will be clear from context whether we are dealing with a college admissions or a school choice problem. For economy of notation we will thus denote a preference profile of students in the school choice problem by $R$.

[17] Most authors prefer to define a matching as a mapping from $I$ to $S$ in the school choice problems to emphasize that schools are objects here. However, we find that using the same formulation as for the college admissions problem leads to a more compact notation.

choice problem/profile of student preferences. Since only students possess private information, a mechanism is *strategy-proof* if it is strategy-proof for students.

As in the college admissions problem, a major goal of the literature on school choice problems is to design matching mechanisms that satisfy certain desirable properties. There is an important difference to the college admissions problem, where stability was identified as a constraint that a matching mechanism *has* to satisfy in order to ensure orderly participation. In the school choice problem, there are several competing desirable properties that have been proposed in the literature.

First of all, given a school choice problem $R$ a matching $\mu$ is *efficient*, if there is no other matching $\tilde{\mu}$ such that $\tilde{\mu}(i) R_i \mu(i)$ for all students $i \in I$ and $\tilde{\mu}(i) P_i \mu(i)$ for at least one $i \in I$. Note that since schools are objects, efficiency only conditions on students' preferences. A matching mechanism is efficient, if it selects an efficient matching for each school choice problem. One of the most commonly used efficient mechanisms in school choice problems is the so called *Boston mechanism* that determines a matching using the following algorithm.[18]

**The Boston mechanism**

Givn a profile of student preferences (and the fixed priority structure), the Boston mechanism proceeds as follows

> In the first round, only students' top choices are considered. Each school $s$ admits the $q_s$ highest priority students who have it as their top choice (or all students if there are fewer than $q_s$). All other students are rejected.
>
> In the $t$th round, only students' $t$th choice schools are considered. Each school $s$ admits the $q_s^t$ highest priority students who have it as their $t$th choice (or all students if there are fewer than $q_s^t$), where $q_s^t$ is the number of empty places at $s$ after round $t-1$. All other students are rejected.

Let $f^{BOS}(R)$ denote the outcome this algorithm chooses for the school choice problem $R$. This mechanism was used to assign students to public schools in Boston until 2005 (Abdulkadiroglu, Pathak, Roth, and Sönmez (2006)) and also has an important role in the German university admissions system that will be analyzed in Chapter 1. A major problem of the

[18]The term Boston mechanism usually refers to the below algorithm with a particular priority structure used for Boston public schools (Abdulkadiroglu and Sönmez (2003)). However, for the properties of this mechanism it is inconsequential which priority structure is used and we use the same term in our description.

Boston mechanism is that students lose their priority for a school unless they rank it sufficiently high (Abdulkadiroglu and Sönmez (2003)) so that students sometimes have an incentive to submit a false preference relation to the mechanism.

There are other efficient mechanisms which provide straightforward incentives to students, the most prominent being the *top trading cycles* mechanism originally developed by Shapley and Scarf (1974).[19] However, depending on the priority structure, there can exist school choice problems $R$ such that for any efficient matching $\mu$ there is a student $i$ and a school $s$ such that $sP_i\mu(i)$ and $i \succ_s i'$ for some $i' \in \mu(s)$. If a student's priority for a school is an absolute right to be admitted to the school before any student with lower priority can be admitted, student $i$ (or her parents) could take legal actions to enforce her priority for school $s$. Furthermore, while priorities cannot be interpreted as measuring the welfare of schools, they often formally represent social or political preferences about the admission process. For these reasons, honoring the stability constraints imposed by the priority structure is an important goal in school choice problems. A matching $\mu$ is called *stable* for the school choice problem $R$ (with the priority structure $\succ$), if

(i) is *individually rational*, if $\mu(i)R_ii$ for all students $i \in I$,

(ii) *eliminates justified envy*, if there is no student school pair $(i, s)$ such that $sP_i\mu(i)$ and $i \succ_s i'$ for some $i' \in \mu(s)$, and

(iii) is *non-wasteful*, if there is no student school pair $(i, s)$ such that $sP_i\mu(i)$ and $|\mu(s)| < q_s$.

Clearly, this notion of stability is equivalent to stability in a college admissions problem (with responsive preferences) where $\succ_s$ is taken to be school $s$' ranking of individual students. For the case of strict priorities, this implies in particular that the SOSM (of the associated college admissions problem) is strategy-proof and *constrained efficient* in the sense that for all school choice problems all students weakly prefer its outcome to any other stable matching.[20]

A problem that we will return to in Chapter 2 is that equal priorities at schools are not necessarily a knife-edge case in the school choice problem. For example, cities are sometimes partitioned into walking zones and students have higher priority for any school in their walking zone than any student living in another part of the city. Here, it is not the political will to discriminate between students within the same walking zone and we have to be careful in breaking ties between students in order to prevent additional welfare loss due to unnecessary stability constraints. In other cases, for example if priorities are at least partly determined by

[19] A comprehensive recent survey that includes these mechanisms and also discusses potential applications to markets for organ transplants is Sonmez and Unver (2008).

[20] See Ergin (2002) for an elegant characterization of priority structures for which the SOSM achieves full efficiency.

test scores as in the first chapter of this thesis, ties in priority orders are much less likely.

# Chapter 1

# An Analysis of the German University Admissions System

## 1.1 Introduction

According to German legislation, every student who obtains the *Abitur* (i.e., successfully finishes secondary school) or some equivalent qualification is entitled to study any subject at any public university. In accordance with this principle of free choice university admission was a completely decentralized process prior to the 1960s: a student with the appropriate qualification could just enroll at the university of her choice. Problems emerged in the early 1960s when some universities had to reject a substantial number of applicants for medicine and dentistry. Rejections were usually based on some measure of the quality of the *Abitur*, mostly the average degree. This often led to a threshold for average grades, called *Numerus Clausus*, such that applicants with higher averages were not admitted.[1] The problem quickly spread to other disciplines and many universities had to establish local admission criteria. This resulted in a very complicated decentralized admission procedure that forced students to spend more time on maximizing their chances of admission than to figure out which university fitted their needs.[2] To solve these problems a centralized clearinghouse, the *Zentralstelle für die Vergabe von Studienplätzen* (ZVS), was established in 1973. Ever since its introduction the ZVS has been subject to immense public scrutiny and political debates. These debates led to gradual changes in the assignment procedure, with the last major revision in 2005.

In this chapter we analyze the most recent version of the ZVS procedure that is used to

[1] In Germany average grades range from 1.0 to 6.0, with 1.0 representing the best possible average grade. Hence, *high* average grades indicate a *bad* performance in secondary school.

[2] In a study of the admission procedures at 40 universities, Scheer (1999) finds that 70 different admission criteria were used.

allocate places for medicine and related subjects. The procedure consists of three steps that sequentially allocate parts of total capacity: In step one twenty percent of available places at each university can be allocated among applicants with exceptional average grades. This is implemented by first using average grades to *select* as many applicants as places can be allocated in step one and then assigning selected applicants (henceforth *top-grade applicants*) to universities using the *Boston mechanism*. In this mechanism the priority of a top-grade applicant for a university is determined by average grade and subordinated social criteria such as distance between hometown and the university. In step two a completely analogous procedure is used to allocate up to twenty percent of available places at each university among applicants who have unsuccessfully participated in previous assignment procedures (henceforth *wait-time applicants*) on basis of average grades and social criteria. In the third step all remaining places - this includes in particular all places that could have been but were not allocated in the first two steps - are assigned among remaining applicants according to criteria chosen by the universities using the *college (university) proposing deferred acceptance algorithm* (CDA). In a sense, the ZVS procedure tries to have the *best of both worlds* by using the applicant proposing Boston mechanism for the priority based steps of the procedure (steps 1 and 2) and letting universities take an active role in the last step of the procedure, where they are able to evaluate applicants themselves.

With respect to reported preferences, the ZVS procedure can lead to very undesirable allocations. For example, an applicant assigned in the first step may prefer a university at which she could have been admitted in the third step. However, the procedure is highly manipulable so that reported need not correspond to true preferences. In particular, prospective students can submit one ranking for each step of the procedure which allows them to condition their reports on the different admission criteria and assignment mechanisms used in the three steps of the procedure. In general, applicants have to make a difficult trade-off between securing a match in an early step and staying in the procedure in hope of obtaining a better assignment in a later step. We argue that given the structure of the German university admissions system it is reasonable to assume that universities do not act strategically. Under this assumption we show that the set of (complete information) equilibrium outcomes coincides with the set of matchings that are *stable* with respect to the true preferences of applicants and the admissions environment. Here, stability roughly means that if an applicant prefers some university to the assignment received, she could not have been admitted at that university no matter which of the three different admission criteria are considered. Using well known results from the theory of two-sided matching, we show that the ZVS procedure supports equilibrium outcomes that are Pareto dominated with respect to applicants' preferences. We also briefly consider the case of incomplete information. Two simple examples point out the problems associated with (i) the sequential allocation of places, and (ii) allowing universities to use their position in applicants' preference rankings as an admission criterion for the last step of the procedure.

Given the deficiencies of the existing procedure, we develop a proposal for a redesign of the current system in the second part of the paper. The approach is to take the basic university admissions environment, in particular the legal constraints, as given and look for better alternatives *within* this environment. The main idea is to assign all places simultaneously while keeping the three quota system of the current ZVS procedure. We introduce a version of Gale and Shapley (1962)'s student proposing deferred acceptance algorithm (SDA) in which places initially reserved for top-grade and wait-time applicants stay open for qualifying applicants throughout the procedure. As in the current ZVS procedure, if in some round of the algorithm, the supply of places for top-grade applicants exceeds demand at a particular university, the excess capacity can be allocated on basis of the admission criteria chosen by this university among all interested applicants. In contrast to the current procedure top-grade applicants may, however, reclaim any place that was initially reserved for them in later rounds of the modified SDA. Thus, quotas are *floating* in the sense that the number of places allocated according to fixed priorities and universities' own admission criteria, respectively, may vary across different rounds of the procedure. We show that the SDA with floating quotas produces a matching that is as favorable as possible to applicants subject to the stability constraints of the German university admissions environment. A major benefit of the new procedure is that it provides applicants with dominant strategy incentives to submit their true preferences if universities are not allowed to use their position in applicants' rankings. If universities are not strategic, the outcome chosen by the modified SDA thus (weakly) Pareto dominates any equilibrium outcome of the current ZVS procedure with respect to the true preferences of applicants. Outside the context of the German system, we discuss how the proposed procedure can be used to implement affirmative action plans in school choice problems while ensuring a non-wasteful allocation of school places.

This chapter is structured as follows: After discussing the related literature, we describe the current ZVS procedure and illustrate it by means of a simple example in section 1.2. In section 1.3 we analyze the revelation game induced by the ZVS procedure under the assumption that universities do not act strategically. In section 1.4 we develop advice for a potential redesign of the current system. In section 1.5 we conclude and discuss our results. Some proofs, further details about the current procedure, data on the evaluation process, and a short history of the

ZVS procedure are relegated to Appendix A.1.

## Related Literature

Since the assignment procedure analyzed in this paper combines the Boston mechanism with the college optimal stable matching mechanism, the theoretical and applied literature concerned with these two algorithms is closely related.

There are three incidents of real-life matching procedures that were found to be equivalent to (one of the versions of) a deferred acceptance algorithm: Roth (1984a) showed that the matching algorithm used to match graduating medical students to their first professional position in the US from 1951 until the late 1990s was equivalent to the CDA. In a similar vein, Balinski and Sönmez (1999) showed that the mechanism used to assign Turkish high school graduates to public universities, the *multi-category serial dictatorship*, was also equivalent to this mechanism. More recently, Guillen and Kesten (2008) have shown that the mechanism used to allocate on campus housing among students of the Massachusetts Institute of Technology is equivalent to the SDA. Our study contributes to this literature by reporting another case of a real-life assignment procedure that uses the CDA. However, in the German system this mechanism is combined with the well known Boston mechanism which, to the best of our knowledge, is the first time a combined use of these two popular mechanisms has been observed and analyzed.

The Boston mechanism has been extensively studied in the matching literature since Abdulkadiroglu and Sönmez (2003)'s influential study of school choice systems. Ergin and Sönmez (2006) analyze the preference revelation game between students induced by the Boston mechanism. They show that the set of pure strategy equilibrium outcomes coincides with the set of stable matchings for the school choice problem. Hence, strategic incentives of students "correct" the instabilities of the Boston mechanism.[3] Well known results from the theory of two-sided matching markets then imply that the SDA outcome weakly dominates any equilibrium outcome of the Boston mechanism with respect to the true preferences of students. In an empirical investigation of the Boston mechanism, Abdulkadiroglu, Pathak, Roth, and Sönmez (2006) find strong evidence that many students try to manipulate the mechanism. They argue that the strategic choices of some families hurt other families who strategize suboptimally.[4] In an ex-

[3]Kojima (2008) shows that this result also holds for *generalized priority structures*, which are formally equivalent to substitutable preferences over subsets of students. These generalized priority structures can accommodate e.g. affirmative action constraints, which are often present in real-life applications of the school choice problem.

[4]A theoretical argument in this vein is provided by Pathak and Sönmez (2008). They consider a model in which students are either fully strategic or naive in the sense that they always submit their true ranking of schools. The main result is that the equilibrium outcomes of the Boston mechanism correspond to the set of

perimental comparison of school choice mechanism, Chen and Sönmez (2006) found that the student optimal stable matching mechanism outperformed the Boston mechanism in terms of efficiency. According to Abdulkadiroglu, Pathak, Roth, and Sönmez (2006) these theoretical, empirical, and experimental results were instrumental in convincing school choice authorities in Boston to replace their assignment procedure with the student optimal stable matching mechanism.[5] In our study we show that Ergin and Sönmez (2006)'s equilibrium characterization has a natural extension to the more complicated German admission system. Furthermore, it is shown that the SDA can be accommodated to the specific constraints of the German market. The associated matching mechanism always selects an applicant optimal stable (for the German market) matching and is strategy-proof for applicants. This shows that at least the theoretical arguments in favor of deferred acceptance algorithms remain valid despite the complex constraints in Germany. We view this as an important initial step in convincing German authorities to change their assignment procedure.

Another study of the German university admissions system is Braun, Dwenger, and Kübler (2008). Using data for the winter term 2006/2007 they find considerable support for the hypothesis that applicants try to manipulate the ZVS procedure.[6] Our paper, which was drafted independently of this empirical study, complements this research since it shows precisely how these findings can be explained by applicants' strategic incentives. A major benefit of the more theoretical approach is that we are not only able to design a promising alternative but can also compare it directly to the equilibrium outcomes of the current procedure.

## 1.2 The German University Admissions System

The ZVS assigns places for medicine and related subjects.[7] There is a separate assignment procedure for each course of study and applicants have to decide in which one of these procedures to participate prior to their application. The assignment of places in all courses of studies proceeds in three sequential steps.

1. In the first step, *Step E*, (up to) one fifth of total places at each university are allocated among applicants with an *exceptional* qualification, that is, an excellent, or very low,

stable matchings for a modified school choice problem in which strategic students have higher priority than naive students.

[5]For a more positive perspective on the Boston mechanism in some special symmetric environments see Miralles (2008) and Featherstone and Niederle (2008).

[6]There have been some minor changes in the procedure since then, which we detail in Appendix A.1.

[7]For biology and psychology, only those universities that still offer a diploma certificate allocate places via the centralized ZVS procedure.

average grade in school leaving examinations.

2. In the second step, *Step W*, (up to) one fifth of total places at each university are allocated among applicants with an exceptionally long waiting time, that is, a long time since obtaining their high-school degree.
3. In the third step, *Step U*, all remaining places are allocated among applicants not assigned in steps E and W on basis of universities' preferences.

In the following, let $\mathcal{A}$ be the set of applicants interested in a particular course of studies and let $\mathcal{U}$ denote the set of universities offering this course. In order to participate in the centralized assignment procedure applicants have to submit an ordered (preference) list of universities for each step of the procedure. There is no consistency requirement on the three lists and the list submitted for step $i \in \{E, W, U\}$ is used only to determine assignments in step $i$. All three preference lists are submitted simultaneously. For step W, applicants can rank as many universities as they want. For steps E and U at most six universities can be ranked. Let $Q_a = (Q_a^E, Q_a^W, Q_a^U)$ denote the preference lists that $a \in \mathcal{A}$ submitted to the ZVS. An applicant *applies for a place in step $i$* if she ranks at least one university for step $i$ of the procedure. Let $q_u$ denote the total number of places that university $u$ has to offer. Let $q_u^E = q_u^W = \frac{1}{5}q_u$ and $q^E = q^W = \frac{1}{5}\sum_{u\in\mathcal{U}} q_u$ denote the number of places at university $u$ and the total number of places available in steps E and W, respectively. To avoid integer problems we assume that, for all $u \in \mathcal{U}$, $q_u$ is a multiple of five. With these preparations, the ZVS procedure can be described as follows.[8]

**Step E: Assignment for excellent applicants**

**(Selection)** Select $q^E$ applicants from those that applied for a place in step E. If there are more than $q^E$ such applicants, order applicants lexicographically according to (i) average grade, (ii) time since obtaining qualification, (iii) completion of military or civil service, (iv) lottery. Select the $q^E$ highest ranked applicants in this ordering.

**(Assignment)** Apply the Boston mechanism to determine assignments of selected applicants. University $u$ can admit at most $q_u^E$ applicants, the preference relation of a selected applicant $a$ is $Q_a^E$, and an applicant's priority for a university is determined lexicographically by (i) average grade, (ii) social criteria,[9] (iii) lottery. Denote the matching produced in step E by $f^{ZVSE}(Q^E)$.

[8] The main reference for this description are the *Vergabeverordnung ZVS, Stand: WS 2008/2009* and *Merkblatt M1 - M10* , which can be found at www.zvs.de. The following is a simplified version of the actual assignment procedure and some omitted details can be found in Appendix A.1.

[9] In this category, applicants are ordered lexicographically according to the following criteria: 1. Being severely disabled. 2. Main residence with spouse or child in the district or a district-free city associated to this

**Step W: Assignment for wait-time applicants**

**(Selection)** Select $q^W$ applicants from those that applied for a place in this step and have not been assigned in step E. If there are more than $q^W$ such applicants, order applicants lexicographically according to (i) time since obtaining qualification, (ii) average grade, (iii) completion of military or civil service, (iv) lottery. Select the $q^W$ highest ranked applicants in this ordering.

**(Assignment)** Apply the Boston mechanism to determine assignments of selected applicants. University $u$ can admit at most $q_u^W$ applicants, the preference relation of a selected applicant $a$ is $Q_a^W$, and an applicant's priority for a university is determined lexicographically by (i) social criteria, (ii) average grade, (iii) lottery. Denote the matching produced in step W by $f^{ZVSW}(Q^W)$.

**Step U: Assignment according to universities' preferences**

For each university $u \in \mathcal{U}$, all remaining places are allocated in this step. Let $q_u^U = q_u - \left|f_u^{ZVSE}(Q^E)\right| - \left|f_u^{ZVSW}(Q^W)\right|$, that is, the total capacity of this university minus places assigned in steps E and W.

(**Preference Formation**) Each university $u$ evaluates applicants who have not been assigned in the two previous steps and listed $u$ in their preference list for step U. Each university $u$ submits the results of this evaluation in form of a strict ranking $R_u$ of individual applicants and the option of leaving a place unfilled.

(**Assignment**) Apply the college proposing deferred acceptance algorithm to determine an assignment of applicants to universities. University $u$ can admit at most $q_u^U$ applicants, the preference relation of an applicant $a$ is given by $Q_a^U$, and the preference relation of university $u$ over individual applicants is given by $R_u$. Denote the matching produced in this step of the procedure by $f^{ZVSU}(Q^U, R_{\mathcal{U}})$.

### 1.2.1 An example

In the following, we illustrate the ZVS procedure by calculating the chosen assignment in a simple example. This will also serve as a first step in the analysis of the procedure since the example already highlights some of the problems.

place of study. 3. Granted request for preferred consideration of top choice. 4. Main residence with parents in the area associated with this place of study. Note that, in contrast to the selection stage, an applicant's priority may thus differ across universities.

Suppose that $\mathcal{A} = \{a_1, \ldots, a_9\}$ and $\mathcal{U} = \{u_1, u_2, u_3\}$. For simplicity, assume that each university has three places to allocate among students and that one place at each university is available in all three steps of the ZVS procedure.[10] Applicants are indexed in increasing order of their average grades, so that $a_i$ has the $i$th best average grade among $a_1, \ldots, a_9$. The applicants with the longest waiting time are $a_7, a_8, a_9$. Assume that applicants rank available universities as follows:[11]

| $R_{\mathcal{A}}$ | $R_{a_1}$ | $R_{a_2}$ | $R_{a_3}$ | $R_{a_4}$ | $R_{a_5}$ | $R_{a_6}$ | $R_{a_7}$ | $R_{a_8}$ | $R_{a_9}$ |
|---|---|---|---|---|---|---|---|---|---|
| | $u_1$ | $u_1$ | $u_3$ | $u_2$ | $u_2$ | $u_3$ | $u_2$ | $u_2$ | $u_1$ |
| | $u_2$ | $u_3$ | $u_2$ | $u_1$ | $u_3$ | $u_2$ | $u_1$ | $u_1$ | $u_2$ |
| | $u_3$ | $u_2$ | $u_1$ | $u_3$ | $u_1$ | $u_1$ | $u_3$ | $u_3$ | $u_3$ |

We now calculate the assignment chosen by the ZVS procedure under the assumption that all applicants submit their preferences truthfully for each step of the procedure.

In Step E, applicants $a_1, a_2, a_3$ are selected since they have the best average grades. The Boston mechanism produces the following assignment

$$f^{ZVSE}(R_{\mathcal{A}}) = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_1 & a_2 & a_3 \end{array}.$$

Since the Boston mechanism is used to determine assignments this matching is efficient with respect to the preferences of $a_1, a_2, a_3$. Note that this assignment is not stable: $a_2$ would strictly prefer a place at $u_3$, has a better average grade than $a_3$ and yet $a_3$ was assigned a place at $u_3$. Applicant $a_2$ would have been assigned a place at $u_3$ if she had ranked this university first (since she has a better average grade), but she loses her priority over student $a_3$ by ranking $u_3$ second. Thus, $a_2$ has an incentive to overreport her preference for $u_3$ in this example.

Next, we calculate the assignment in step W. Given the above description, $a_7, a_8$, and $a_9$ are eligible for a place in this step. To pin down assignments, assume that the priority ordering in the assignment stage of step W is $a_8, a_7, a_9$ at university $u_1$ (applicants are listed in decreasing order of priority), $a_9, a_7, a_8$ at $u_2$, and $a_7, a_8, a_9$ at $u_3$.[12] In this case, the Boston mechanism

[10] It is unproblematic to let each universities' capacity be some multiple of five if one includes more applicants to take the additional places. Larger examples do not facilitate the understanding of the mechanism and all the points made below apply equally well to larger, more realistic settings. This point applies to all examples considered in this chapter.

[11] This notation means that e.g. applicant $a_1$ strictly prefers $u_1$ over $u_2$ over $u_3$ and that only these universities are acceptable to her.

[12] This ranking could result e.g. if $a_8$ lives in the vicinity of $u_1$ and $a_9$ lives in the vicinity of $u_2$.

produces the following assignment

$$f^{ZVSW}(R_{\mathcal{A}}) = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_9 & a_7 & a_8 \end{array}.$$

As in step E the matching is efficient with respect to the preferences of $a_7, a_8, a_9$. Similar to above, $a_8$ would have been better off claiming that her most preferred university is $u_1$. In addition, $a_8$ and $a_2$ would both benefit if they were allowed to trade their places.[13] Hence, the matchings chosen in the first two steps of the ZVS procedure are not necessarily efficient with respect to the preferences of all applicants selected in steps E <u>and</u> W.

Finally, we calculate the assignment in step U. To pin down assignments, assume that $R_{u_1} : a_4, a_5, a_6$, $R_{u_2} : a_6, a_5, a_4$, and $R_{u_3} : a_4, a_5, a_6$. The college proposing deferred acceptance algorithm in step U produces the following assignment

$$f^{ZVSU}(R_{\mathcal{A}}, R_{\mathcal{U}}) = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_4 & a_6 & a_5 \end{array}.$$

Note that this is the college/university optimal stable matching in a college admissions problem with participants $\{a_4, a_5, a_6, u_1, u_2, u_3\}$ and preferences as given above if universities' preferences are responsive to $R_{\mathcal{U}}$.

## 1.3 Analysis of the assignment procedure: Strategic Incentives

The example in the last section showed that applicants sometimes have an incentive to manipulate the ZVS procedure by submitting a ranking of universities that does not correspond to their true preferences. Strategic behavior is encouraged by the ability to submit three different preference lists. In its official information brochures the ZVS makes it very clear to applicants that they should choose submitted preference lists carefully in order to maximize their chances of admission. These materials, available at www.zvs.de, even contain examples where profitable manipulations are explicitly calculated. Braun, Dwenger, and Kübler (2008) provide empirical evidence showing that applicants do act upon the incentives to manipulate the assignment procedure. In order to evaluate the performance of the university admissions system it is thus

[13] One argument against this trade is that it would mean that university $u_2$ is stuck with two of the three applicants with the worst average grades.

important to analyze the strategic incentives induced by the ZVS mechanism.

An important question is whether universities are strategic players or not. Given that universities take a passive role in steps E and W, the only possibility for strategizing is the evaluation process in step U that we now describe in some detail. Prior to the application deadline, each university has to decide on the criteria that it will use to rank applicants in step U. A university can use detailed information about applicants that is provided by the ZVS. This information includes its position in the preference rankings submitted by applicants for step U, average grades, waiting time, and so on. A university may also gather additional information about applicants for example by conducting interviews or evaluating letters of motivation. If a university has announced that it will use only "objective" criteria such as its rank in applications or average grades, there is no scope for manipulation since unrightfully (according to the criteria set by the university) rejected applicants could sue the university.[14] If a university uses "subjective" criteria such as performance in interviews, there might be some scope for strategic manipulation. However, we will assume that universities do not act strategically and always submit their ranking of applicants truthfully. This is not without loss of generality, but (i) only a limited number of universities use subjective criteria,[15] and (ii) there have not been reports about universities strategically manipulating their lists of applicants in step U. For these reasons the assumption of non-strategic universities is a useful approximation and we will concentrate on the strategic incentives of applicants in the following. This is not to say that universities do not act strategically at all. Rather, the game induced by the ZVS procedure can be viewed as a two-stage game where in the first stage universities (strategically) announce their evaluation criteria and in the second stage applicants submit their rankings of universities. In this paper we focus for the most part on the game between applicants.

Before proceeding to the analysis, it is useful to formally summarize those factors that will be taken as fixed and to define a few terms that will be used throughout the whole remainder of this chapter. The *university admission environment* consists of the sets of applicants $\mathcal{A}$, the set of universities $\mathcal{U}$, the vector of universities' capacities $q$, and the criteria which determine applicants' priorities in steps E and W. For our purpose these criteria can be summarized as follows: Let $\tilde{\mathcal{A}} \subseteq \mathcal{A}$ be the set of applicants who would be selected in step E if all applicants applied for a place in step E. Remember that selection is related to submitted preference lists only in so far that no applicant is considered who did not rank any university for step E. For

[14]This is quite likely given that the assignment procedure is subject to immense public scrutiny. The information brochure of the ZVS actually includes many advertisements for lawyers specialized in suing universities over their admission decisions.

[15]In Pharmacy for example, only 2 out of 22 universities employ subjective criteria. See Appendix A.1.

$u \in \mathcal{U}$, let $\succ_u^E$ denote the ordering of $\tilde{\mathcal{A}}$ that results by applying the criteria of the assignment stage in step E. In this chapter we assume that this ordering is *strict* so that no two top-grade applicants have equal (top-grade) priority for some university. This can be extended to an ordering of all applicants by setting $u \succ_u^E a$ for all $a \notin \tilde{\mathcal{A}}$. Let $\succ^E$ be the resulting profile of orderings. The profile of orderings for step $W$ is defined analogously and is denoted by $\succ^W$. As above, we assume that no two wait-time applicants have equal (wait-time) priority for any university. The admissions environment can thus be summarized by $(\mathcal{A}, \mathcal{U}, q, \succ^E, \succ^W)$ and it is taken to be fixed throughout the analysis. Applicant $a \in \mathcal{A}$ is a *top-grade* (*wait-time*) applicant if $a \succ_u^E u$ ($a \succ_u^W u$) for all $u \in \mathcal{U}$. Next, we define feasible assignments for the university admissions environment. Since the total capacity is divided into three parts, a matching has to specify not only to which university an applicant is matched, but also which of the three types of places she receives. More formally, we have the following.

**Definition 1.** *A* ***matching for the university admission environment*** *is a three-tuple of matchings $\mu = (\mu^E, \mu^W, \mu^U)$ that respects capacity constraints for all three steps and assigns at most one place to each applicant, that is,*

*(i) for all $i \in \{E, W, U\}$ and all $a \in \mathcal{A}$, $\mu^i(a) \in \mathcal{U} \cup \{a\}$*

*(ii) $|(\mu^E(a) \cup \mu^W(a) \cup \mu^U(a)) \cap \mathcal{U}| \leq 1$ for all $a \in \mathcal{A}$,*

*(iii) for $i \in \{E, W\}$ and all $u \in \mathcal{U}$, $|\mu^i(u)| \leq q_u^i$,*

*(iv) for all $u \in \mathcal{U}$, $|\mu^U(u)| \leq q_u^U(\mu^E, \mu^W) = q_u - \left|\mu^E(u)\right| - \left|\mu^W(u)\right|$, and*

*(v) for all $i \in \{E, W, U\}$, all $a \in \mathcal{A}$, and all $u \in \mathcal{U}$, $a \in \mu^i(u)$ if and only if $\mu^i(a) = u$.*

We usually suppress the dependency of $q_u^U$ on $\mu^E$ and $\mu^W$. Applicant $a$'s *assignment under matching* $\mu$ is $u$ if for some $i \in \{E, W, U\}$, $\mu^i(a) = u$, and it is $a$ if for all $i \in \{E, W, U\}$, $\mu^i(a) = a$. We denote $a$'s assignment under $\mu$ by $\mu(a)$ and $a$ is *unassigned* if $\mu(a) = a$.

### 1.3.1 Complete Information

In this section we assume that the admissions environment, applicants' true preferences, and the rankings of universities to be used for step U are all common knowledge among applicants. While the assumption of complete information might seem strong, applicants can often rely on the outcomes of past assignment procedures to estimate their chances of admission. If universities rely solely on objective criteria such as average grade, these estimates are usually quite reliable so that complete information is a reasonable approximation to the actual information structure in this case. However, in case a university uses subjective criteria this means that an applicant knows exactly how she would perform e.g. in an interview at some university, irrespective of whether the interview took place or not. We will return to this point in the next section. In order to simplify the analysis and to be able to characterize complete information equilibrium outcomes we make four additional assumptions.

**Strict University Rankings:** The ZVS procedure allows universities to restrict attention to applicants who have ranked them sufficiently high. We assume that the criteria a university uses apart from such *ranking constraints* induce a strict ordering $\succ_u^U$ of $\mathcal{A} \cup \{u\}$.[16] For example, a university $u$ might consider only applicants who ranked it first and order these applicants according to their average grades. In this case, $\succ_u^U$ would simply list applicants in increasing order of their average grades.

**No Empty Lists:** All applicants always rank at least one university for each part of the procedure. Without this assumption the set of applicants selected in steps E and W would depend on the profile of submitted rankings. The empirical evidence in Braun, Dwenger, and Kübler (2008) offers strong support in favor of the assumption. The assumption is restrictive, since it can be shown that it may - in very rare cases - be in an applicant's best interest to manipulate the set of students eligible for a place in steps E or W. An example demonstrating this point can be found in Appendix A.1.

**Disjoint Sets:** No applicant who is selected in step E can also be selected in step W. Without this assumption the set of applicants selected in step W would depend on the assignment in step E. The assumption is reasonable since applicants with very good average grades are typically assigned a place of study before they could become eligible for step W.[17]

**Strict Preferences and Early Assignment:** If an applicant could be matched to the same university in steps E/W and U, she prefers to be assigned in E/W. This assumption makes sense since assignments in steps E/W are determined more than one month before step U is conducted (to give universities enough time to evaluate applicants). The additional time an applicant gains by receiving a place in steps E/W is valuable extra time to search

[16]We use the same notation as for priority orderings here to point out that this ordering is taken to be fixed throughout.

[17]For example, in medicine an applicant needed an average grade of at least 1.3 to be selected in step E, while an applicant must have waited at least 5 full years in order to be considered in step W. At 18 out of the 34 universities offering medicine, a student with an average grade of 1.3 would have been guaranteed to receive a place in step U for the winter term 2008/2009. This data is publicly available at www.zvd.de.

for an apartment, move, and so on. Formally, we assume that each applicant $a$ has a (true) strict ranking $R_a$ of $\mathcal{U} \cup \{a\}$ and a preference ranking $\tilde{R}_a$ over (three tuples of) matchings such that she strictly prefers $\mu$ over $\tilde{\mu}$, denoted by $\mu \tilde{P}_a \tilde{\mu}$, if and only if either $\mu(a) P_a \tilde{\mu}(a)$, or for some $i \in \{E, W\}$ and $u \in \mathcal{U}$, $\mu^i(a) = \tilde{\mu}^U(a) = u$ .

We now describe the revelation game between applicants induced by the ZVS mechanism. For each applicant $a$, the set of admissible *strategies* consists of three-tuples of rankings $Q_a = (Q_a^E, Q_a^W, Q_a^U)$, such that, for $i \in \{E, U\}$, $Q_a^i$ is an ordered list containing at least one and at most six universities in decreasing preference, and $Q_a^W$ is an ordered list containing at least one university. Let $\mathcal{Q}$ denote the set of admissible strategies and let $\mathcal{Q}^{|\mathcal{A}|}$ be the set of all possible strategy profiles. Given a profile of applicant reports $Q \in \mathcal{Q}^{|\mathcal{A}|}$, the admissions environment $(\mathcal{A}, \mathcal{U}, q, \succ^E, \succ^W)$, and the preferences of universities $\succ^U$, let $f^{ZVS}(Q)$ be the outcome of the ZVS procedure. Denote by $f_a^{ZVS}(Q)$ the assignment that applicant $a$ receives. Given a profile of strict applicant preferences $R$, the *game induced by the ZVS mechanism* is denoted by $\Gamma^{ZVS}(R) = (\mathcal{A}, \mathcal{Q}^{|\mathcal{A}|}, f^{ZVS}, \tilde{R})$. A strategy profile $Q$ is a *Nash equilibrium* of $\Gamma^{ZVS}(R)$ if there is no profitable unilateral deviation, that is, for all $a \in \mathcal{A}$, $f_a^{ZVS}(Q) \tilde{R}_a f_a^{ZVS}(Q'_a, Q_{-a})$, for all $Q'_a \in \mathcal{Q}$. We now define a notion of stability for matchings that is adapted to the specific nature of the German admission environment and that will turn out to be crucial in characterizing equilibrium outcomes. First of all, a *university admissions problem* is given by an admissions environment, a profile of universities' preferences, and a profile of applicants' preferences. Since everything else is assumed to be fixed, we will think of an admissions problem as being given by a profile of student preferences.

**Definition 2.** *Let $R$ be a profile of strict applicant preferences. A matching $\mu = (\mu^A, \mu^W, \mu^H)$ is* ***stable with respect to the university admissions problem*** *$R$ if the following conditions are satisfied.*

*(i) No university is assigned an unacceptable applicant for one of its quotas, that is, $a \succ_u^i u$ for all $i \in \{E, W, U\}$, all $u \in \mathcal{U}$, and all $a \in \mu^i(u)$.*

*(ii) No applicant is matched to an unacceptable university, that is, $\mu(a) R_a a$ for all $a \in \mathcal{A}$,.*

*(iii) Applicants are matched as early as possible, that is, if $\mu^U(a) = u$ for some university $u$ then there is no $i \in \{E, W\}$ such that either ($a \succ_u^i u$ and $|\mu^i(u)| < q_u^i$), or ($a \succ_u^i \tilde{a}$ for some applicant $\tilde{a} \in \mu^i(u)$).*

*(iv) There is no blocking pair within and across quotas, that is, there is no applicant-university pair $(a, u)$ such that $uP_a\mu(a)$ and, for some $i \in \{E, W, U\}$, either ($a \succ_u^i u$ and $|\mu^i(u)| < q_u^i$) or ($a \succ_u^i \tilde{a}$ for some $\tilde{a} \in \mu^i(u)$).*

This definition of stability takes into account that different criteria are used to regulate admissions in the three different steps of the assignment procedure. Part (iii) of this definition ensures that in case of multiple possibilities of admission at a university applicants take the place that was intended for them. The following is the main result of this section.

**Theorem 1.** *Let $R$ be a profile of strict applicant preferences. The set of pure strategy Nash equilibrium outcomes of $\Gamma^{ZVS}(R)$ coincides with the set of stable matchings for the university admissions problem $R$.*

The proof of this result is deferred to Appendix A.1. We now briefly sketch why $f^{ZVS}(Q)$ has to be stable in the sense of Definition 2 if $Q$ is a Nash equilibrium. The intuition for the stability of $f^{ZVSE}(Q^E)$ with respect to the preferences of all applicants matched in step E of the ZVS procedure is very similar to the intuition for Ergin and Sönmez (2006)'s characterization for the Boston mechanism: A top-grade applicant takes into account that she loses her priority for a university in step E to other top-grade applicants unless she ranks it first and thus strategically overreports her preference for certain universities if necessary. However, here it is not necessarily optimal to be matched in step E by all means since there is another chance to be assigned a more preferred university in step U. If a top-grade applicant $a$ is matched in step E but blocks $f^{ZVSU}(Q^U, \succ^U)$ together with some university $u$ (according to Definition 2), she could profitably deviate by ranking only $u$ for all steps of the procedure: If $a$ did not receive a place at $u$ in step E for this alternative report $\tilde{Q}_a$, only a subset of applicants would have had to wait for step U (compared to $Q$). This implies that all applicants apart from $a$ receive weakly more offers in step U of the ZVS procedure under $\tilde{Q} = (\tilde{Q}_a, Q_{-a})$. This implies in particular that all applicants who rejected an offer by $u$ in the CDA under $Q$ will continue to do so when the profile of preferences is $\tilde{Q}$. But then if $a$ and $u$ blocked $f^{ZVSU}(Q^U, \succ^U)$, $a$ must obtain a place at $u$ in step U of the ZVS procedure under $\tilde{Q}$; contradiction. At this point it is illustrative to calculate the set of stable matchings and thus the set of equilibrium outcomes for a simple example.

**Example 1.** *Consider the example of section 1.2.1. To calculate the set of stable matchings, we first need to complete the preferences of universities for step U to a ranking of all applicants. We assume that*

$$\succ^U_{u_1}: \; a_1, \; a_4, \; a_5, \; a_2, \; a_3, \; a_6, \; a_7, \; a_8, \; a_9$$
$$\succ^U_{u_2}: \; a_1, \; a_6, \; a_2, \; a_3, \; a_5, \; a_4, \; a_7, \; a_8, \; a_9$$
$$\succ^U_{u_3}: \; a_1, \; a_4, \; a_5, \; a_2, \; a_3, \; a_6, \; a_7, \; a_8, \; a_9$$

*Note that restricted to $a_4, a_5, a_6$ these orderings coincide with the preferences of universities given in section 3.1. Preferences of applicants and the priority structures for steps E and W are as given in section 3.1. For this specification, there are exactly two stable matchings.*

$$\mu_1^E = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_1 & \emptyset & a_3 \end{array}, \qquad \mu_1^W = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_8 & a_9 & a_7 \end{array}, \qquad \mu_1^U = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_2 & \{a_4, a_5\} & a_6 \end{array}, \text{ and}$$

$$\mu_2^E = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_1 & a_3 & a_2 \end{array}, \qquad \mu_2^W = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_8 & a_9 & a_7 \end{array}, \qquad \mu_2^U = \begin{array}{ccc} u_1 & u_2 & u_3 \\ a_4 & a_6 & a_5 \end{array}.$$

*The proof that these are the only stable matchings can be found in Appendix A.1. By Theorem 1, there are thus two pure strategy equilibrium outcomes of the revelation game among applicants. Note that all applicants weakly prefer $\mu_1$ over $\mu_2$. One strategy profile that implements the first matching is the following: All top-grade applicants rank only their most preferred university for step E and submit their true ranking for step U. Wait-time applicants rank only their assignment under the stable matching $\mu_1$ for step W. The remaining three applicants rank only their most preferred university for step U. For $a_2$ this means that she* truncates *her true preferences so that she will stay in the procedure until step U, where she can be assigned a place at her most preferred university $u_1$ given the reports of the others. Applicants with a long waiting-time on the other hand, overreport their preferences in fear of falling through the cracks in the Boston mechanism of step W and knowing that their chances of obtaining a preferred university in step U are slim.*[18]

*However, note that $a_2$ is guaranteed to obtain a place at $u_3$ in step E if she ranks this university first - irrespective of the reports of the other applicants. On the other hand, $a_2$ has to rely on others to follow the right equilibrium strategy in order to reach the Pareto dominant equilibrium. In this sense, the Pareto dominant equilibrium is more risky for $a_2$ so that she might be inclined to use the safe strategy of overreporting her preference for $u_3$ in step E.*

[18]This is roughly in accordance with the empirical analysis in Braun, Dwenger, and Kübler (2008) who note that (i) about 28 percent of the places available to top-grade applicants are not filled in step E because selected applicants submit very short preference lists, and that (ii) there are almost no places reserved for but not filled in step W.

Theorem 1 shows that the potential instabilities of ZVS procedure we saw in the example of section 1.2.1 are "corrected" by the strategic behavior of applicants. It can be viewed as an extension of Ergin and Sönmez (2006)'s result about the equilibrium outcomes of the Boston mechanism to the more complicated ZVS procedure. The characterization of complete information outcomes of the ZVS procedure is a useful benchmark for comparing the current procedure with our proposal for a redesign in the next section.

We will later see that the set of stable matchings for the university admissions problem coincides with the set of stable matchings for a related college admissions problem with substitutable preferences. It is known (Roth (1984b)) that for such problems there always exists a student optimal stable matching that all students weakly prefer to any other stable matching. This result implies that the ZVS procedure supports Pareto dominated outcomes, a fact that we have already seen in the above example.

An interesting corollary of the result is that neither the constraint that applicants can rank at most six universities for steps E and U, nor universities' use of their rank in applications has any effect on the set of matchings that are attainable as equilibrium outcomes. To see this note that in order to achieve an arbitrary stable matching, applicants rank just their assigned university (if any) for each part of the procedure. Conversely, if a matching fails to be stable, an applicant involved in a blocking pair with some university $u$ can profitably deviate by ranking only this university for each step of the procedure. Thus, the ranking constraints set by universities never come into play in a complete information equilibrium. The just mentioned strategies are, of course, very risky since they entail a potentially high probability of being left unassigned by the end of the procedure if applicants are only slightly mistaken about the preferences of universities and other applicants. The discussion is meant to point out that *if* applicants had a very reliable estimate of their chances of admission at each university *then* constraints would be irrelevant.

### 1.3.2 Incomplete Information

The last section considered the case of complete information. This informational environment is a reasonable approximation if universities mostly rely on objective measures such as average grades to evaluate applicants. However, calculating chances of admission is significantly harder for universities using subjective criteria such as interviews: First, there is usually very little to no data available on past admission decisions (for step U). Secondly, applicants may find it hard to estimate their potential performance in an interview. Assuming complete information is thus more appropriate when relatively few universities use subjective criteria. In particular, the applicability of the results in the last section may depend on the course of study under consideration: In the assignment procedure for pharmacy, only 2 out of 22 universities used subjective criteria to assign at least part of their capacity. On the other hand, 10 out of 34 universities in the assignment procedure for medicine gathered additional information about applicants by conducting interviews (see Appendix A.1). In this case, a full answer to the question of which outcomes can be expected from the ZVS procedure would require modeling applicant expectations about the interviewing process, about other applicants' preferences, and so on. Given the complexities of the ZVS procedure the problem of characterizing incomplete information equilibria can quickly become very difficult or even intractable. Instead of aiming for general results, we use two very simple examples to point out the possible complications caused by (i) the sequential allocation of places, and (ii) the possibility that universities use ranking constraints for step U. These features will be abolished in our proposal for a redesign in the next section.

As before, we assume that applicants know the preferences and average grades of their peers. Applicant $a$'s average grade is denoted by $g(a)$. In contrast to the last section, however, applicants are uncertain about their performance in interviews and thus their chances of admission at a university in step U: Applicant $a$'s performance in interviews is the same across universities and is summarized by a one-dimensional random variable $\theta_a$ with realizations in some finite set $\Theta \subset \mathbb{R}$. Interview performance is identically and independently distributed across applicants. All universities rank applicants on basis of the sum of average grade and performance in the interview: Applicant $a$ ranks higher in $u$'s preferences than applicant $a'$ if and only if $g(a) + \tilde{\theta}_a < g(a') + \tilde{\theta}_{a'}$. Note that high realized values of $\theta$ are thus associated with bad performance in an interview. The first example outlines the problems of the sequential nature of the ZVS procedure.

**Example 2 (The Problems of Sequential Allocation).** *Suppose there are seven applicants $a_1, \ldots, a_7$ and three universities $u_1, u_2, u_3$. Applicants are ordered in increasing order of their average grades, that is, $g(a_i) < g(a_j)$ for $i < j$. For simplicity we concentrate on the assignment procedures for steps E and U in this example and assume that each university has only one seat to allocate in each part of the procedure. Hence, $a_1, a_2, a_3$ are the top-grade applicants. For this example, we assume that $\Theta = \{-x, 0, x\}$, where $x > g(a_5) - g(a_2)$.*[19] *Let $p := Prob(\theta_a = -x)$,*

[19]This is not an unrealistic assumption when $g(a_5) - g(a_2)$ is not too large. In the ZVS procedure universities often use weighted averages of average grades and performance in interviews. It is not uncommon that a grade difference of, say, 0.3 points is reversed by performances in interviews.

$q := Prob(\theta_a = 0)$, *and* $r := Prob(\theta_a = x)$. *Ordinal preferences of applicants are given by (only acceptable universities are listed)*

| $R$ | $R_{a_1}$ | $R_{a_2}$ | $R_{a_3}$ | $R_{a_4}$ | $R_{a_5}$ | $R_{a_6}$ | $R_{a_7}$ |
|---|---|---|---|---|---|---|---|
| | $u_1$ | $u_1$ | $u_2$ | $u_1$ | $u_2$ | $u_3$ | $u_3$ |
| | | $u_2$ | | | | | |

*For all applicants except* $a_2$ *it is optimal to rank their only acceptable university for each step of the procedure.*[20] *Applicant* $a_2$*'s utilities for obtaining a place at* $u_1$ *and* $u_2$ *are given by* $A$ *and* $B$ *with* $A > B > 0$, *respectively. The utility of not being assigned a university is normalized to zero and the utility of being assigned to* $u_3$ *is negative. It is easy to see that* $a_2$ *should either try to secure a match in step E by ranking* $u_2$ *as her first choice for step E, or should rank only* $u_1$ *as acceptable for step E and then submit her true ranking for step U. The last statement follows since interview performance is the same across universities so that for any realization* $u_1$ *and* $u_2$ *have the same preferences over applicants.*

*Suppose first that* $a_2$ *decides to wait for step U and ranks only* $u_1$ *for step E. Then the probability of obtaining a place at* $u_1$ *is* $p + q(1 - p) + r^2$ *and the probability of obtaining a place at* $u_2$, *conditional on not obtaining a place at* $u_1$, *is* $qp(1 - p) + r^2(1 - r)$ *(remember that interview performance was assumed to be the same across universities). Hence,* $a_2$*'s expected value of waiting for step U is* $[p+q(1-p)+r^2]A+[qp(1-p)+r^2(1-r)]B$. *On the other hand, she will be matched to* $u_2$ *for sure if she ranks it as her first choice for step E. Hence,* $a_2$ *will wait for step U if and only if* $[p + q(1 - p) + r^2]A + [qp(1 - p) + r^2(1 - r)]B > B$. *This can be rearranged to yield*

$$\frac{A}{B} > \frac{1 - [qp(1-p) + r^2(1-r)]}{p + q(1-p) + r^2}. \tag{1.1}$$

*Note first that if the utility difference between obtaining first and second choice university is big,* $a_2$ *is willing to accept a non-negligible risk of being left unassigned in step U. For example, if* $A = 3B$ *and* $p = q = r$, $a_2$ *prefers to wait for assignment in step U even though there is a probability of* $qp^2 + r(1-r)^2 \approx 0.2$ *that she ends up unassigned. Thus, the sequential assignment procedure undermines the idea that top-grade applicants should not suffer too much from having a bad interview since they sometimes have to accept risky gambles in order to maximize their expected payoff from participating in the ZVS procedure.*[21]

[20] The other applicants can also be interpreted as representing $a_2$'s beliefs about her competitors.

[21] Braun, Dwenger, and Kübler (2008) report the case of an applicant with an average grade of 1.1 who did not receive a place in step E and who was subsequently rejected by all four universities he listed for step U. An applicant with an average grade of 1.1 is usually among the top 2 percent of high school graduates. Thus, even

*Secondly, if the utility difference between obtaining first and second choice university is small, $a_2$ will often prefer the safe option of taking a place at $u_2$ in step E. Note that if $a_2$ refrains from taking part in step U, $p + q(1-p) + r^2$ is the probability that $a_2$ would have been $u_1$'s top candidate had she not been assigned in step E. This can be interpreted as the probability that $u_1$ ends up with the "wrong" applicant in step U. Now suppose that $A/B = 5/4$, $p = 0.5$, and $r = q = 0.25$. From (1.1) we see that $a_2$ (weakly) prefers the safe option even though there is a chance of almost* 70 *percent that she receives a place at her most preferred university in step U. Put differently, the probability that $u_1$ ends up with the wrong applicant in step U is close to* 70 *percent! This shows that from an* ex-ante *perspective the current ZVS procedure may produce outcomes that are very undesirable for universities.*

Note that the problems outlined in example 2 are not peculiar to the precise form of the assignment procedure used in the sequential steps of the ZVS procedure. For example, the same problems would occur if instead the SDA procedure would be used in each step. In this sense, the above problems are direct consequences of the decision to allocate places in the three different quotas sequentially. The next example illustrates the problems associated with allowing universities to use *ranking constraints* which force applicants to rank a university sufficiently high if they want to be considered.

**Example 3** (**The problems of Ranking Constraints**)**.** *There are three applicants $a_1, a_2, a_3$ and two universities $u_1$ and $u_2$. As above, applicants are ordered in increasing order of average grades. For simplicity we consider only step U and assume that both universities have only one place to assign among applicants. We assume that $\Theta = \{-y, 0, y\}$, where $y > g(a_3) - g(a_1)$. Let $p := Prob(\theta = -y)$, $q := Prob(\theta = 0)$, and $r = Prob(\theta = y)$. Ordinal preferences of applicants are given by (as above only acceptable universities are listed)*

| $R$ | $R_{a_1}$ | $R_{a_2}$ | $R_{a_3}$ |
|---|---|---|---|
| | $u_1$ | $u_1$ | $u_2$ . |
| | $u_2$ | | |

*Applicant $a_1$'s utilities for $u_1$ and $u_2$ are $A$ and $B$ with $A > B > 0$. Note that if no university uses ranking constraints then it is optimal for $a_1$ to submit her true ordinal ranking of universities for step U. As above this follows from the assumption that interview performance is the same at all universities. This implies that if no university uses its rank in applications the outcome of step U must be stable with respect to the true preferences of participants. In this*

for top-grade applicants there is a non-negligible risk of being left unassigned by the procedure if they fail to secure a place early in the procedure.

*case the probability that $u_2$ is assigned its most preferred applicant (among the two applicants $a_1$ and $a_2$ it interviews) is $q(1-p)p + r^2(1-r)$, while $u_1$ always obtains its most preferred applicant (among $a_1$ and $a_3$). To see this note that (i) if $\theta_{a_1} = -y$ then $a_1$ is the top candidate of both universities with probability one so that $u_2$ will not get its most preferred applicant, (ii) if $\theta_{a_1} = 0$ then $a_1$ is the top candidate of $u_2$ if $\theta_{a_3} \geq 0$ but $u_2$ is assigned $a_1$ only if $\theta_{a_2} = -y$, and (iii) if $\theta_{a_1} = y$ then $a_1$ is the top candidate of $u_2$ if $\theta_{a_3} = y$ but $u_2$ is assigned $a_1$ only if $\theta_{a_2} \leq 0$. Note that the expected utility of applicant $a_1$ is given $[p+q(1-p)+r^2]A + [qp(1-p)+r^2(1-r)]B$ as in Example 2.*

*Now suppose $u_2$ declares that it only considers applicants who ranked it first. In this case $a_2$ has to decide whether to report her first choice truthfully and thus forsake chances of admission at $u_2$, or to rank $u_2$ as her first choice university. Note that $a_2$'s expected utility of (truthfully) ranking $u_1$ first is now $[p+q(1-p)+r^2]A$, since she would not be considered by $u_2$ if she fails to secure a place at $u_1$. On the other hand her expected utility for misrepresenting her first choice and ranking $u_2$ first is $[p+q(1-p)+r^2]B + [qp(1-p)+r^2(1-r)]A$. She will thus misrepresent her first choice if and only if $[p+q(1-p)+r^2]A < [p+q(1-p)+r^2]B + [qp(1-p)+r^2(1-r)]A$. This is satisfied for example if $p = q = r$ and $A/B = 5/4$. For this parameter constellation, $u_2$ would increase the probability that it receives its most preferred applicant from $4/27$ to one by forcing $a_1$ to rank it first. Hence, $u_2$ has a strict incentive to employ this tool if we assume that $u_2$ has a strictly higher utility for obtaining its most preferred applicant. Note that for the above parameter constellation this means that* ex-post *there is a blocking pair ($a_1$ and $u_2$) for the ZVS outcome with probability $p + q(1-p)^2 + r^3 \approx 0.5$.*

There is one major benefit of allowing universities to use ranking constraints: If a university only considers applicants who ranked it first, its evaluation efforts are never wasted since any offer it makes must be accepted. Without ranking constraints, universities would potentially have to evaluate many more candidates to fill their places. This is problematic if the marginal cost of evaluating an additional candidate is not negligible as in the case of rankings that are based on interviews. Nevertheless, ranking constraints can hardly be seen as a satisfactory solution since they force applicants to forsake valuable chances of admission. For example, six of the universities offering medicine only consider applicants who ranked them first (see Appendix A.1). Thus, an applicant interested in only these universities effectively has to decide on *one* university to rank for step U of the ZVS procedure. Furthermore, note that there were no explicit interviewing costs in the above example and yet $u_2$ had an incentive to restrict attention to applicants who ranked it first. This shows that it is not necessarily a cost saving motive that drives universities' incentives to employ ranking constraints. There is empirical evidence that universities make excessive use of such constraints: Many universities use *mechanical* evaluation procedures where the ranking of applicants can be easily computed from characteristics such as average grade. The marginal cost of "evaluating" an additional candidate is thus negligible. Still, many of the universities with mechanical evaluation procedures use ranking constraints. In medicine, for example, 12 out of the 34 universities have a mechanical evaluation procedure and accept only applicants who ranked them "sufficiently" high (see Appendix A.1). Here, the ability to use ranking constraints lets universities take unduly advantage of the centralized procedure in the sense that they are able to elicit binding commitments from applicants that would not be possible in a decentralized procedure.

## 1.4 Towards a New Design

The main goal of this section is to propose a redesign of the German university admissions system. Our approach is to keep the university admissions environment as defined in section 3 fixed and to construct an alternative mechanism <u>within</u> this environment. More specifically, we maintain the following features of the current assignment procedure.

(i) Each university's capacity is divided into three parts: One for top-grade applicants, one for wait-time applicants, and one for which universities can evaluate candidates.

(ii) Places not taken by top-grade and wait-time applicants can be allocated on basis of university criteria among all applicants.

(iii) Places for top-grade and wait-time applicants are allocated on basis of $\succ^E$ and $\succ^W$, respectively.

The division of capacities into three parts makes sense since it represents the political will that some places should be reserved for top-grade and wait-time applicants. Without places reserved for wait-time applicants, these applicants would typically not have a chance to be admitted at popular universities due to their high average grades. Braun, Dwenger, and Kübler (2008) mention that the quota for wait-time students is generally seen to be necessary to fulfill the constitutional requirement that any student with the appropriate qualification should have a chance of being admitted at any university. One motivation for reserving some places for top-grade applicants is that they should not suffer too much from having a couple of bad interviews which would dampen admission chances in step U. The second requirement ensures that seats are not wasted because of insufficient demand from top-grade and wait-time applicants. Finally, the admission criteria used in steps E and W are taken to represent the preferences of policymakers that are not subject to discussion.

In our redesign we focus on the welfare and incentives of applicants and continue to assume that universities do not act strategically. The goal will be to design a stable, in the sense of Definition 2, and strategy-proof procedure that is as favorable as possible to applicants. Stability ensures that top-grade and wait-time applicants never lose their priority for a place in their quota, and that universities' preferences are respected. The requirement that applicants should be matched as early as possible can be understood as the desire to reach a 1:1:3 distribution of students admitted through the top-grade, wait-time, and the university quota, respectively. An important argument in favor of stability is that it offers a clear explanation for rejections: An applicant rejected by $u$ can be told that this was because there were other applicants who ranked higher with respect to (all of) the admission criteria at $u$. This feature could significantly increase public acceptance of the centralized procedure, which is important given that it is subject to immense public scrutiny. A practical reason for requiring the assignment to be as favorable as possible to applicants is that some universities allow applicants to exchange places assigned by the ZVS. These exchanges are not conducted by a centralized procedure[22] and there is a chance that due to lack of information some mutually beneficial exchanges are not carried out. It thus makes sense to minimize the number of mutually advantageous exchanges in advance (subject to stability constraints) and this can be achieved by implementing an applicant optimal stable matching. Apart from these arguments, Theorem 1 shows that the currently employed ZVS procedure achieves a stable matching in equilibrium. Hence, even in the current procedure any deviation from stability is either due to suboptimal strategizing or informational frictions.

In the current ZVS procedure there are three major impediments to non-manipulability that will need to be abolished:

1. The sequential allocation of places in the three quotas,
2. universities' ability to impose ranking constraints, and
3. the limited length of preference lists for steps E and U.

The problems associated with 1. and 2. were discussed in the last section. Similar to ranking constraints, the restriction to rank no more than six universities for steps E and U forces some applicants to forsake valuable chances of admission. Dropping 3. should not lead to a significant increase in the complexity of the assignment procedure. Furthermore, the ZVS already handles lists of arbitrary length in step W, where applicants are allowed to rank as many universities

[22]Regional newspapers often publish student requests and there are a number of websites, such as www.studienplatztausch.de, that offer to organize exchanges.

as they want.[23]

Given this discussion, we propose that the evaluation process should take place before the application deadline. We present some ideas on how the evaluation process can be organized efficiently in section 1.5. After evaluations have been completed, universities and applicants submit their rankings simultaneously to the centralized procedure. Since we are interested in a procedure that is non-manipulable by applicants, we consider assignment procedures that elicit a single ranking from applicants. The redesign is based on the student proposing deferred acceptance algorithm adapted to handle the three quota system of the German university admissions environment. For this application, we require a more general domain of preferences than the domain of responsive preferences that we now introduce.

### 1.4.1 The College Admissions Problem with Substitutable Preferences

In the general college admissions problem each college $c$ has a strict ranking $R_c$ of subsets of $I$. Given $R_c$ and a subset $J \subset I$, we can define $c$'s *choice from* $J$ as the subset $Ch_c(J)$ of $J$ such that for all subsets $J' \subseteq J$ with $Ch_c(J) \neq J'$, $Ch_c(J) P_c J'$. A matching $\mu$ is *pairwise stable*, if

(i) no student is matched to an unacceptable college, that is, $\mu(i) R_i i$ for all $i \in I$,

(ii) no college prefers to reject some of the students assigned to it, that is, $Ch_c(\mu(c)) = \mu(c)$ for all $c \in C$, and

(iii) there is no student-college pair $(i, c)$ such that $c P_i \mu(i)$ and $i \in Ch_c(\mu(c) \cup \{i\})$.

Note that this reduces to the definition of stability given in Chapter I.1. when college preferences are responsive.

We assume that college $c$'s preferences are *substitutable*: If $i \in Ch_c(J)$ for some $i \in J \subset I$ then $i \in Ch_c(J')$ for all $J' \subseteq J$ containing $i$. Kelso and Crawford (1982) show that the set of stable matchings is non-empty if all colleges have substitutable preferences. Under the same assumption, Roth (1984b) shows that there is a *student optimal stable matching* that all students weakly prefer to any other stable matching. As in the case of responsive preferences, this matching can be found by applying the *student proposing deferred acceptance algorithm* (SDA). We briefly describe the appropriate reformulation of the algorithm for the case of substitutable preferences.

[23] If there are binding constraints on the length of preference lists, the SOSM is not strategy-proof for students (Romero-Medina (1998)). Haeringer and Klijn (2008) show that there can even be unstable Nash equilibrium outcomes of the preference revelation game induced by the SOSM in this case.

**The Student Proposing Deferred Acceptance Algorithm**

Given a profile of student preferences over colleges and a profile of college preferences over groups of students, the *student proposing deferred acceptance algorithm* proceeds as follows

> In the first round each student applies to her most preferred college. Each college temporarily admits its choice from the set of students who applied to it.
>
> In the $t$th round each student who was rejected in round $t-1$ proposes to her best acceptable college among those that have not rejected her in any of the earlier rounds. Each college temporarily admits its choice from the set of students who applied to it in this step and the set of students it had temporarily admitted by the end of step $t-1$.

The algorithm ends when all unmatched students have proposed to all acceptable colleges. Let the direct mechanism that associates the outcome of this algorithm to each preference profile be denoted by $f^I$.

We now consider students' strategic incentives when face with this mechanism. Assume that, in addition to substitutability, college preferences satisfy the *law of aggregate demand*: If $J' \subseteq J \subseteq I$ then $|Ch_c(J')| \leq |Ch_c(J)|$. Hatfield and Milgrom (2005) show that if this assumption is satisfied and college preferences are substitutable, the SDA as a direct mechanism is *strategy-proof for students.* In fact, Hatfield and Kojima (2009) show that the SDA as a direct mechanism is even *group strategy-proof for students* in the sense that there is no group of students that can make all of its members strictly better off by a joint misrepresentation of preferences. These results show that the SDA is a very desirable allocation procedure in the college admissions problem with substitutable preferences if the law of aggregate demand is satisfied and the goal is to achieve a stable matching.

### 1.4.2 Proposal for a Redesign

In the following, the university admissions environment is taken as given and we continue to assume that universities are non-strategic. For each university $u \in \mathcal{U}$, we fix a ranking $\succ_u^U$ of $\mathcal{A} \cup \{u\}$ which is meant to represent the outcome of the evaluation process of university $u$. In order to compare the outcome of my proposed alternative with the current ZVS procedure, we assume that this is the same ranking as the one that would be used in the current procedure if universities ignored their rank in applications (see the discussion in section 1.3.1).

As in the analysis of the current assignment procedure, a university admissions problem is given by a profile of applicant preferences. The main idea is to construct an *associated college admissions problem* such that the student optimal stable matching for this problem is the applicant optimal matching among those that are stable for the university admissions problem. In the associated college admissions problem, the set of students is identified with the set of applicants and we associate to each university $u$ a college $c(u)$ whose preferences over groups of students are determined by applying the three different admission criteria of the university. Since only the choice function of a college is important in order to apply the deferred acceptance algorithm, we directly construct the choice function of $c(u)$ without explicitly specifying the preferences over groups of applicants. Given a set of applicants $A \subset \mathcal{A}$, university $u$'s *choice from* $A$, denoted by $Ch_{c(u)}(A)$, is constructed in three steps which use the three different admission criteria.

**Step E:** Let $Ch^E_{c(u)}(A)$ be the set of the $q^E_u$ top-grade applicants in $A$ who have the highest priorities according to $\succ^E_u$, or the set of all top-grade applicants if there are no more than $q^E_u$ top-grade applicants in $A$.

**Step W:** Let $Ch^W_{c(u)}(A)$ be the set of the $q^W_u$ wait-time applicants in $A$ who have the highest priorities according to $\succ^W_u$, or the set of all wait-time applicants if there are no more than $q^W_u$ wait-time applicants in $A$.

**Step U:** Let $q^U_u(A) = q_u - |Ch^E_{c(u)}(A)| - |Ch^W_{c(u)}(A)|$ denote the remaining capacity after steps E and W. Let $Ch^U_{c(u)}(A)$ be the set of the $q^U_u(A)$ highest ranking acceptable applicants in $A \setminus (Ch^E_{c(u)}(A) \cup Ch^W_{c(u)}(A))$ according to $\succ^U_u$, or all acceptable applicants if there are no more than $q^U_u(A)$ acceptable applicants in $A \setminus (Ch^E_{c(u)}(A) \cup Ch^W_{c(u)}(A))$.

Let $Ch_{c(u)}(A) = Ch^E_{c(u)}(A) \cup Ch^W_{c(u)}(A) \cup Ch^U_{c(u)}(A)$ denote the resulting choice of $c(u)$ from $A \subseteq \mathcal{A}$. This construction mimics the current ZVS procedure in the sense that we first check whether an applicant can be admitted as a top-grade applicant, then check whether she can be admitted as a wait-time applicant, and, finally, consider the applicant's chances of being admitted on basis of the criteria chosen by universities. Note that these choice functions cannot in general be rationalized by responsive preferences since different admission criteria are used to order top-grade, wait-time, and other applicants, respectively.

The idea is to determine an assignment for the university admissions problem by applying the SDA to the associated college admissions problem. The above construction of choice functions ensures that in any round in which a university $u$ receives only few proposals by top-grade applicants, remaining places can be allocated according to the admission criteria set by $u$. If a top-grade applicant proposes to $u$ in some later round of the SDA procedure, she can still claim one of the places that are reserved for exceptional applicants. This *floating quota* system is in sharp contrast to the current ZVS procedure, where a top-grade who does not claim a place early in the procedure loses her top-grade priority since places are irreversibly converted to places that are allocated on basis of universities' preferences. Note that through the course of the algorithm the number of places a university allocates according to its own criteria decreases: If at some point of the algorithm a top-grade applicant takes one of the places reserved for her, she will keep this place unless another top-grade applicant with higher priority applies in one of the later rounds. In the following, $f^{\mathcal{A}}(R)$ will denote the matching chosen by the SDA for the associated college admissions problem when applicant preferences are $R$ (remember that we assume everything but applicants' preferences to be fixed). The following example illustrates the construction of college preferences and the SDA procedure.

**Example 4.** *Consider again the university admissions problem of section 1.2.1 and Example 1. For the convenience of the reader, we briefly summarize priorities and preferences. Remember that each university was assumed to have one place to allocate for each step of the ZVS procedure. Preferences of students are given by*

| $R$ | $R_{a_1}$ | $R_{a_2}$ | $R_{a_3}$ | $R_{a_4}$ | $R_{a_5}$ | $R_{a_6}$ | $R_{a_7}$ | $R_{a_8}$ | $R_{a_9}$ |
|---|---|---|---|---|---|---|---|---|---|
| | $u_1$ | $u_1$ | $u_3$ | $u_2$ | $u_2$ | $u_3$ | $u_2$ | $u_2$ | $u_1$ |
| | $u_2$ | $u_3$ | $u_2$ | $u_1$ | $u_3$ | $u_2$ | $u_1$ | $u_1$ | $u_2$ |
| | $u_3$ | $u_2$ | $u_1$ | $u_3$ | $u_1$ | $u_1$ | $u_3$ | $u_3$ | $u_3$ |

*Applicants are indexed in increasing order of average grades so that $a_1, a_2, a_3$ are top-grade applicants. We assumed that applicants $a_7, a_8, a_9$ were wait-time applicants and that the priority structure for step $W$ was (we only list wait-time applicants here)*

$$\succ_{u_1}^{W}: \; a_8, \; a_7, \; a_9$$
$$\succ_{u_2}^{W}: \; a_9, \; a_7, \; a_8$$
$$\succ_{u_3}^{W}: \; a_7, \; a_8, \; a_9$$

*The preferences of universities for step $U$ were given by the following*

$$\succ_{u_1}^{U}: \; a_1, \; a_4, \; a_5, \; a_2, \; a_3, \; a_6, \; a_7, \; a_8, \; a_9$$
$$\succ_{u_2}^{U}: \; a_1, \; a_6, \; a_2, \; a_3, \; a_5, \; a_4, \; a_7, \; a_8, \; a_9$$
$$\succ_{u_3}^{U}: \; a_1, \; a_4, \; a_5, \; a_2, \; a_3, \; a_6, \; a_7, \; a_8, \; a_9$$

*For this university admissions problem we have $Ch_{c(u_2)}(\{a_4, a_5, a_7, a_8\}) = \{a_4, a_5, a_7\}$ since (i) no applicant in $\{a_4, a_5, a_7, a_8\}$ is a top-grade applicant, (ii) $a_7, a_8$ are wait-time applicants and $a_7 \succ^W_{u_2} a_8$, and (iii) $\{a_4, a_5\} \succ^U_{u_2} a_8$ so that $a_4$ and $a_5$ admitted in step U. Similarly, one can establish that $Ch_{c(u_1)}(\{a_1, a_2, a_8, a_9\}) = \{a_1, a_2, a_8\}$ and $Ch_{c(u_2)}(\{a_4, a_5, a_7, a_9\}) = \{a_4, a_5, a_9\}$. We now calculate the outcome of the SDA.*

*In the first round, $a_1, a_2, a_9$ apply to $u_1$, $a_4, a_5, a_7, a_8$ apply to $u_2$, and $a_3, a_6$ apply to $u_3$. Given that $Ch_{c(u_2)}(\{a_4, a_5, a_7, a_8\}) = \{a_4, a_5, a_7\}$ so that in the second round of the SDA $a_8$ applies to her second choice university $u_1$. Since $Ch_{c(u_1)}(\{a_1, a_2, a_8, a_9\}) = \{a_1, a_2, a_8\}$ this causes $a_9$ to apply to $u_2$ in the third round. This in turn leads to the rejection of $a_7$ since $Ch_{c(u_2)}(\{a_4, a_5, a_7, a_9\}) = \{a_4, a_5, a_9\}$. In the fourth round $a_7$ applies to $u_1$, where she is immediately rejected since $Ch_{c(u_1)}(\{a_1, a_2, a_7, a_8\}) = \{a_1, a_2, a_8\}$. In the fifth, and final, round she applies to $a_3$. Since she is accepted and no other applicant is unmatched the procedure ends and the resulting matching is thus*

$$f^{\mathcal{A}}(R) = \begin{array}{ccc} u_1 & u_2 & u_3 \\ \{a_1, a_2, a_8\} & \{a_4, a_5, a_9\} & \{a_3, a_6, a_7\} \end{array}$$

*Note that this coincides with the Pareto dominant equilibrium outcome $\mu_1$ in Example 1. We will see shortly that this is not peculiar to the above example.*

We now show that $f^{\mathcal{A}}$ is a desirable mechanism for the associated college admissions problem. For this purpose, the following proposition establishes the most important properties of the choice functions constructed above.

**Proposition 1.** *For any university $u$, $Ch_{c(u)}$ satisfies substitutability and the law of aggregate demand.*

**Proof:** Consider some university $u \in \mathcal{U}$ and a set of applicants $A \subset \mathcal{A}$. Suppose that $a \in Ch_{c(u)}(A)$ and consider a set $B \subset A$ with $a \in B$. We have to show that $a \in Ch_{c(u)}(B)$. If $a \in Ch^E_{c(u)}(A) \cap B$, we must have $a \in Ch^E_{c(u)}(B)$ since $B$ contains only a subset of top-grade applicants from $A$. Similarly, if $a \in Ch^W_{c(u)}(A) \cap B$ then $a \in Ch^W_{c(u)}(B)$. Hence, the only case left to consider is $a \in Ch^U_{c(u)}(A) \cap B$. Suppose that $a \notin Ch^E_{c(u)}(B) \cup Ch^W_{c(u)}(B)$. By the above arguments we must have that $B \setminus [Ch^E_{c(u)}(B) \cup Ch^W_{c(u)}(B)] \subseteq A \setminus [Ch^E_{c(u)}(A) \cup Ch^W_{c(u)}(A)]$ and $q^U_u(B) \geq q^U_u(A)$. Given that $a \in Ch^U_{c(u)}(A)$, $a$ must be one of the $q^U_u(A)$ highest ranking applicants according to $\succ^U_u$ within $A \setminus (Ch^E_{c(u)}(A) \cup Ch^W_{c(u)}(A))$. But then given the above, she must be among the $q^U_u(B)$ highest ranking applicants according to $\succ^U_u$ within $B \setminus (Ch^E_{c(u)}(B) \cup Ch^W_{c(u)}(B))$ and thus $a \in Ch^U_{c(u)}(B)$. This completes the proof of substitutability.

To see that the law of aggregate demand is also satisfied, note that if an applicant is accepted when the set of applicants is $A$, then for any set of applicants $B \supseteq A$, she will either still be chosen or some other applicant takes her place. Hence, the cardinality of the choice set of each university is weakly increasing in the superset order. $\square$

By the properties of the SDA for college admission problems with substitutable preferences (see section 1.4.1), we thus obtain the following.

**Proposition 2.** *(i) For any profile of strict applicant preferences $R$, $f^{\mathcal{A}}(R)$ is the applicant optimal stable matching in the associated college admissions problem.*

*(ii) $f^{\mathcal{A}}$ is group strategy-proof for applicants.*

Given the allocative goals laid out in the beginning of this section, this is a very encouraging result if we can show that stability in the associated college admissions problem is equivalent to stability in the university admissions problem. Since a matching in the university admissions problem is a triple of matchings, we need to define how to transform matchings between the two problems. First, let $\mu = (\mu^E, \mu^W, \mu^U)$ be a matching in the university admissions problem. For all $u \in \mathcal{U}$, set $\mu_1[c(u)] = \mu^E(u) \cup \mu^W(u) \cup \mu^U(u)$ and let $\mu_1$ be the resulting matching for the associated college admissions problem. Similarly, given a matching $\tilde{\mu}$ for the associated college admissions problem set $\tilde{\mu}_2^i(u) = Ch^i_{c(u)}(\tilde{\mu}[c(u)])$ for all $i \in \{E, W, U\}$ and all $u \in \mathcal{U}$, and let $\tilde{\mu}_2$ be the resulting matching for the university admissions problem. We have the following.

**Proposition 3.** *Fix a profile of strict applicant preferences $R$.*

*(i) If $\mu$ is a stable matching for the university admissions problem $R$, $\mu_1$ is stable for the associated college admissions problem with applicant preferences $R$.*

*(ii) If $\tilde{\mu}$ is stable for the associated college admissions problem with applicant preferences $R$, $\tilde{\mu}_2$ is a stable matching for the university admissions problem $R$.*

**Proof :**

(i) It is immediate that $\mu_1$ is individually rational for the associated college admissions problem. So suppose to the contrary that there is an applicant-college pair $(a, c(u))$ that blocks $\mu_1$ in the associated college admissions problem, that is, $uP_a\mu_1(a)$ and $a \in Ch_{c(u)}(\mu_1[c(u)] \cup \{a\})$. But then the construction of $Ch_{c(u)}$ implies that, for some $i \in \{E, W, U\}$, $a \in Ch^i_{c(u)}(\mu_1[c(u)] \cup \{a\})$. By Definition 2, if $\mu$ is a stable matching for the university admissions problem it has to match applicants as early as possible. This implies that, for all universities $u \in \mathcal{U}$, $\mu^E(u)$ ($\mu^W(u)$) contains the $q_u^E$ ($q_u^W$) highest priority top-grade (wait-time) applicants w.r.t. $\succ_u^E$ ($\succ_u^W$) within $\mu^E(u) \cup \mu^W(u) \cup \mu^U(u)$. Hence, $Ch^E_{c(u)}(\mu_1[c(u)]) = \mu^E(u)$, $Ch^W_{c(u)}(\mu_1[c(u)]) = \mu^W(u)$, and $Ch^U_{c(u)}(\mu_1[c(u)]) = \mu^U(u)$. If $a \in Ch^E_{c(u)}(\mu_1[c(u)] \cup \{a\})$ we would thus have that either ($|\mu^E(u)| < q_u^E$ and $a \succ_u^E u$), or ($a \succ_u^E \tilde{a}$, for some $\tilde{a} \in \mu^E(u)$). In both cases we obtain a contradiction to point (iv) of Definition 2 for $i = E$. The other cases are handled similarly and we obtain that $\mu_1$ is stable in the associated college admissions problem.

(ii) Let $\tilde{\mu}$ be a stable matching in the associated college admissions problem with applicant preferences $R$. Since, for all $i \in \{E, W, U\}$, $\tilde{\mu}_2^i(u) = Ch^i_{c(u)}(\tilde{\mu}[c(u)])$, $\tilde{\mu}_2^E(u)$ ($\tilde{\mu}_2^W(u)$) contains the $q_u^E$ ($q_u^W$) highest priority top-grade (wait-time) applicants w.r.t. $\succ_u^E$ ($\succ_u^W$) within $\tilde{\mu}[c(u)]$. This implies in particular that $\tilde{\mu}_2$ matches applicants as early as possible. Furthermore, if $uP_a\tilde{\mu}_2(a)$ and either ($|\tilde{\mu}_2^E(u)| < q_u^E$ and $a \succ_u^E u$), or ($a \succ_u^E \tilde{a}$, for some $\tilde{a} \in \tilde{\mu}_2(u)$), we would also have that $a \in Ch^E_{c(u)}(\tilde{\mu}[c(u)] \cup \{a\})$. Thus, $\tilde{\mu}_2$ must satisfy (iv) with respect to $i = E$. The argument that $\tilde{\mu}_2$ satisfies (iv) with respect to $i \in \{W, U\}$ is similar and omitted here.

□

This implies that for all profiles of applicant preferences $R$, $f^{\mathcal{A}}(R)$ is the applicant optimal stable matching for the university admissions problem $R$. This follows directly from Proposition 3 and the corresponding property of the set of stable matchings in a college admissions problem with substitutable preferences (see section 1.4.1). Given the characterization of complete information equilibrium outcomes in Theorem 1, we obtain the following.[24]

**Proposition 4.** *Let $R$ be a university admissions problem and let $\Gamma^{ZVS}(R)$ denote the set of Nash equilibrium outcomes of the game induced by the ZVS mechanism. Then for all $Q \in \Gamma^{ZVS}(R)$, all applicants weakly prefer the outcome chosen by the SDA with floating quotas, that is, $f_a^{\mathcal{A}}(R)\tilde{R}_a f_a^{ZVS}(Q)$ for all $a \in \mathcal{A}$.*

In case there are multiple stable matchings under the current procedure, our redesign eliminates the possibility that applicants coordinate on welfare inferior equilibria. Furthermore,

[24]This result parallels Ergin and Sönmez (2006)'s result comparing equilibrium outcomes of the Boston mechanism and the student optimal stable mechanism. The main differences to their result are the sequential structure of the ZVS procedure and the special construction of college preferences in the associated college admissions problem.

given the strong incentive properties of the new procedure, an applicant does not need any information about the preferences of universities or other applicants since truthfully reporting her preferences is always optimal.

The algorithm works equally well with other quota systems or different quotas across universities. This is important if the above is also taken to be a proposal for reform of the decentralized system that runs parallel to the ZVS procedure and in which the division of capacities can be quite different. For example, universities may decide to reserve significantly less places for wait-time applicants and this can easily be adapted by modifying the quotas in the SDA procedure.

## 1.5 Conclusion and Discussion

This paper analyzed the assignment procedure that is used to allocated places for some courses of study at public universities in Germany. The procedure uses two algorithms, the Boston algorithm and the college proposing deferred acceptance algorithm, that have been studied extensively in the matching literature. The major difference to previous studies is that the German system combines these two algorithms into a complicated sequential assignment procedure. Assuming universities to be non-strategic, we derived a full characterization of complete information equilibria of the revelation game between applicants induced by the admissions procedure. Outside of the complete information setup, two examples demonstrated the problems created by the sequential allocation of seats and the use of ranking constraints by universities.

A modified version of the deferred acceptance algorithm was introduced that can accommodate most of the specific constraints of the German university admissions system and in particular dispenses of the sequential allocation of seats. The main idea was to treat places in quotas for which seats are assigned early in the current procedure as options that remain open throughout the procedure. It was shown that the alternative procedure produces stable matchings as favorable as possible to applicants and provides applicants with (dominant strategy) incentives to submit their true ranking of universities. If applicants use their dominant strategy and universities are not strategic, the outcome of the new assignment procedure (weakly) Pareto dominates any complete information equilibrium outcome of the currently employed procedure with respect to applicant preferences.

Our hope is that these results direct public attention towards the actual problems of the university admissions system in Germany and how these might be overcome. This is particularly important given that there have recently been significant problems for those courses of study that are not part of the centralized procedure.[25] A big problem of these decentralized procedure that the market often fails to clear since some applicants hold multiple offers and this leads to congestion. We think that the widespread refusal of a centralized procedure is rooted in a wrong assessment of the benefits and disadvantages of a decentralized system and that a well designed centralized procedure, possibly along the lines of what has been proposed in this chapter, is capable of addressing many of the concerns about the current system.

We now discuss a number of important issues with the alternative design proposed in the last section. In some instances, these questions present important avenues for future research.

### 1.5.1 Efficiency, Stability, and the Welfare of Universities

Given that the allocation of places among top-grade and wait-time applicants is based on a priority structure that is exogenously assigned to universities, one may also be interested in achieving a matching that is efficient with respect to applicant preferences for this part of total capacity. This is not in general guaranteed by the procedure proposed in the last section: In Example 4 the procedure assigns applicants $a_8$ and $a_9$ places reserved for wait-time applicants which they would prefer to exchange with each other.[26] The problem with this approach is that there is no matching mechanism that is strategy-proof for applicants and satisfies efficiency for the priority based as well as stability for the two-sided part of the procedure.[27] Since stability is important in the sense that it ensures that universities' preferences play a role in the assignment process, we chose to impose stability as the main allocative criterion.

A problem of the new design is that political authorities might perceive it as favoring applicants too much since it chooses the worst stable matching for universities. However, it is easy to show that for any stable matching mechanism there is always an equilibrium which yields the applicant optimal stable matching. Hence, welfare gains of universities relative to the applicant optimal stable matching cannot be guaranteed. [28]

[25] See e.g. http://www.spiegel.de/unispiegel/studium/0,1518,610971,00.html.

[26] It is easy to see that stability uniquely pins down the allocation for places in the wait-time quota so that for this example there is a conflict between stability and efficiency. This problem is well known and Ergin (2002) has characterized the class of priority structures for which there is no such conflict between efficiency and stability.

[27] This is an adaption of a result by Papai (2009), who studies a hybrid model of school choice and college admission problems.

[28] Furthermore, some papers in the literature have argued that the differences between student and college optimal stable matching are "small" if the matching market is "large". This is found as an empirical fact in studies of the matching programs for American doctors (Roth and Peranson (1999)) and the Boston school choice program (Pathak and Sönmez (2008)). Kojima and Pathak (2009) show analytically that the set of stable matchings converges to a singleton if the number of participants becomes very large.

### 1.5.2 The Evaluation Process

For the redesign we assumed that the evaluation of applicants takes place before the application deadline. If a university has a non-mechanical evaluation procedure, for which the marginal cost of considering an additional applicant is not negligible, there are two problems with this approach:

(i) The university may evaluate an applicant who would have received a place at the university anyway through the top-grade or the wait-time quota.
(ii) The university may interview an applicant who ends up ranking the university low.

As a consequence, universities' incentives to invest in a costly evaluation of applicants might decrease. On the other hand, the cost of the evaluation process might increase since more applicants need to be considered to fill available places. While a theoretical investigation of universities' investment incentives is an interesting avenue for future research, it is beyond the scope of this study and we now informally discuss some ideas about the evaluation process. For concreteness, we will talk about *interviews* whenever we mean non-mechanical evaluation procedures.

The new assignment procedure could be implemented as a two-stage mechanism where in the first stage applicants are evaluated and the second stage determines assignments using the SDA with floating quotas. For the evaluation stage, applicants report a set (not a ranking) of acceptable universities. Similar to the current ZVS procedure, universities could set upper bounds on the number of applicants they are willing to interview and criteria according to which interviewees are selected. It might also be sensible to limit the number of interviews an applicant can have. In this case, an applicant would be required to report not only a set of acceptable universities, but would also have to specify at which universities she would like to interview. Prospective students may still have a chance of being assigned to a university at which they did not interview through the top-grade/wait-time quota, or because the university allocates only part of its places on basis of interviews. Once all interviews have taken place, universities and applicants submit their rankings to the centralized clearinghouse. While choosing where to apply for interviews is a difficult decision, submitting the true ranking of universities is still a dominant strategy once the evaluation process of universities is completed (provided that universities are not allowed to use ranking constraints).

### 1.5.3 Floating Quotas and Affirmative Action Constraints

The ideas behind my proposal for a redesign might be useful outside the specific context of the German university admissions system. For example, many school districts in the United States are interested in controlling the distribution of student characteristics, or types, within schools. Abdulkadiroglu and Sönmez (2003) and Abdulkadiroglu (2005) consider the case of *type specific quotas* where schools set upper bounds on the numbers of students of each type.[29] These upper bounds are rigid in the sense that if, say, 60 female and 40 male students apply to a school desiring a 1:1 ratio of male to female students, 10 places would be left unassigned. Our proposal for a redesign outlines one way to rule out a wasteful allocation of places: Assuming that student types are one-dimensional, the total capacity of each school could be split up into several parts according to the type distribution desired by the school choice authority. One could then use the SDA with floating quotas, where female students take the role of top-grade students and male students take the role of wait-time students (the assignment of roles is of course irrelevant here). If in a round of the algorithm the capacities for some types are not exhausted, remaining places are released for allocation among all students irrespective of their type (this corresponds to step U in our construction of choice sets). In the above case of binding constraints for female students, the algorithm would admit all female students. If additional male students apply in a later round, some or all of the 10 female students admitted in excess of capacity would be rejected. Note that the modified SDA ensures that once the desired distribution of types is reached, the distribution is not changed in any of the later rounds. Furthermore, the analysis shows that the criteria for admission can be allowed to differ across quotas. For example, social criteria could be used to allocate places in the reserved quotas, while students in excess of type specific capacity are admitted according to tests and the like. Although we have only considered three quota systems in the paper, it is easy to see that our approach is applicable for any finite number of type categories provided that type categories induce a partition of the set of students.

[29] Abdulkadiroglu and Ehlers (2007) consider a different form of controlled choice where each school has a lower and an upper bound on the number of students with a given type it has to/can admit. They show that in general there is no strategy-proof mechanism satisfying a weak stability criterion and non-wastefulness. Our approach circumvents this problem by allowing controlled choice constraints to be violated if necessary.

# Chapter 2

# Breaking Ties in School Choice: (Non-)Specialized Schools

## 2.1 Introduction

Recently, the school choice problem has received a lot of attention in the theoretical and applied matching literature starting with Abdulkadiroglu and Sönmez (2003). In this problem, a set of students has to be assigned among a set of public schools. Each school has an exogenously given priority ordering of students. A central allocative criterion in the literature is stability, which requires that no student should envy another student for a school that she has *strictly* higher priority for. If students cannot have equal priority at schools, the student proposing deferred acceptance algorithm (SDA) produces a student optimal stable matching and provides students with dominant strategy incentives to submit their preferences over schools truthfully. This is not only of theoretical interest, as school choice authorities in Boston and New York have recently decided to adopt a variant of this mechanism (Abdulkadiroglu, Pathak, Roth, and Sönmez (2006) and Abdulkadiroglu, Pathak, and Roth (2009)). A problem that has not received much attention in the theoretical literature until the recent work by Erdil and Ergin (2008) is that in most real-life applications students may have the same priority at a given school. In the Boston public school choice system, for example, a major determinant for the priority of a student is whether she lives in the *walk zone* of a school, that is, not further away from the school than some fixed distance. Of course, a walk zone inherits (much) more than one student in a densely populated city so that schools' priority orderings have large indifference classes. This seemingly small change in the model changes results dramatically. The major problem is that ties between equal priority students have to be broken in order to determine an assignment. This induces additional stability constraints that can lead to a substantial decrease in student welfare (Abdulkadiroglu, Pathak, and Roth (2009)). We call a mechanism *constrained efficient*, if it is stable with respect to the original weak priority structure and never incurs welfare loss due to tie-breaking. Unfortunately, Erdil and Ergin (2008) show that there are priority structures for which a constrained efficient and strategy-proof mechanism fails to exist.

A natural question that is at the heart of the present study is whether this is an exception or the rule in school choice problems with weak priority orders. We call a priority structure *solvable*, if there exists a strategy-proof and constrained efficient mechanism. In this paper we make important initial progress in characterizing the class of solvable priority structures. We introduce a model of (non-)specialized schools in which a school is either *specialized* and has a strict priority ranking of students, or a school is *non-specialized* and all students have the same priority. While it cannot be expected that this assumption is exactly satisfied in real-life applications, we view the analysis of this model as a useful first step since it provides important insights into how we can deal with large indifference classes in the priority structure. Furthermore, this model is interesting in its own right since it unifies the school choice problem with strict priority orders of Abdulkadiroglu and Sönmez (2003) and the *house allocation problem* of Hylland and Zeckhauser (1979).

For the case that no school can admit more than one student, we fully characterize solvable priority structures by two simple and intuitive conditions. These conditions ensure that most of the ties can be broken exogenously, that is, without referring to student preferences. Since the conditions required for solvability are very restrictive, our results for the unit capacity case have the flavor of an impossibility result. This is in line with Gibbard (1973) and Satterthwaite (1979)'s classical negative results concerning dominant strategy implementation. However, our negative results critically depend on the assumption that no school can admit more than one student. In a second step we then consider general capacity vectors and show that significantly weaker conditions are sufficient for solvability. As for the unit capacity case, our conditions connect the capacity vector with the amount of allowable variability across the priority orderings of specialized schools. Most importantly, our results show that there is some scope for breaking ties according to student preferences and we introduce a new version of the SDA with endogenous tie-breaking (SDA-ETB). For solvable priority structures the associated matching mechanism is strategy-proof for students even though tie-breaking is (partly) based on elicited preferences. Interestingly, increasing capacities substantially enlarges the scope for preference based tie-breaking.

This chapter is organized as follows: After discussing the related literature we introduce the school choice problem with weak priorities and relevant existing results in section 2.2. In section 2.3 we motivate the need for preference based tie-breaking by means of a simple example. In section 2.4 we introduce the (non-)specialized schools model. This section contains the main results of this chapter. In section 2.5 we conclude and discuss our results as well as possible extensions. All proofs are relegated to Appendix A.2.

### Related Literature

Ehlers (2007) was the first to study the problem of indifferences in priority orders. In particular he considered a simple three student example (that does not belong to our (non-)specialized schools environment) for which no exogenous tie-breaking rule guaranteed the constrained efficiency of the SDA. Nevertheless, he showed by construction that a strategy-proof and constrained efficient matching mechanism existed. We are the first to systematically study the possibility of preference based tie-breaking in school choice problems with indifferences in priority orders.

Apart from the above paper, the literature on the school choice problem with indifferences has mainly focused on exogenous tie-breaking. Here, a central question has been whether there should be a single lottery that is used to break ties at all schools, or whether there should be a separate lottery for each school. Pathak (2008) considers a random assignment problem in which all students initially have the same priority for each school. He shows that a market based approach, in which a priority structure for each school is randomly selected and students are then allowed to trade their priorities, is equivalent to the random serial dictatorship, in which a single lottery is conducted and students then choose schools in the order determined by the lottery, in the sense that both produce exactly the same lottery over outcomes.[1] In a similar vein, Abdulkadiroglu, Pathak, and Roth (2009) show that for any school choice problem, any constrained efficient matching can be reached by first using a single lottery to break all ties and then running the deferred acceptance algorithm. The focus of both of these papers is to give a rationale for using a single lottery to break all ties for all schools instead of multiple lotteries. They do not discuss how one can elicit the information about student preferences that is necessary to break ties in a way that avoids additional welfare loss, which is the main focus of our study.

[1]This is an extension of a classical result by Abdulkadiroglu and Sonmez (1998) who show that in the house allocation problem the random serial dictatorship is equivalent to the core from random endowments, which conducts a single lottery to determine an initial allocation of indivisible objects and then lets agents trade towards a core outcome.

More related in focus is the main theoretical result in Abdulkadiroglu, Pathak, and Roth (2009), which shows that no strategy-proof matching mechanism can dominate student optimal stable matching mechanism with *any* fixed tie-breaking rule. The dominance relation they consider is very strong since it requires that *all* students weakly prefer the outcome of the dominating mechanism to the outcome of the dominate mechanism for *all* preference profiles, with at least one strict preference for at least one profile. This already suggests that the class of mechanisms that is not dominated by another strategy-proof mechanism is quite large. Our results show that it is not sufficient to restrict attention to the class of SDAs resulting from fixed tie-breaking rules if one is interested in strategy-proof and constrained efficient mechanisms. Of course, our SDA-ETB does not dominate the SDA with an arbitrary fixed tie-breaking rule for all preference profiles in the above sense. However, in contrast to the latter mechanism it guarantees that there is never additional welfare loss due to tie-breaking if the priority structure is solvable.

Another closely related paper is Erdil and Ergin (2008). They show that a matching is constrained efficient if and only if there is no *stable improvement cycle*, that is, no cyclical sequence of trades that respects stability constraints and makes all students involved better off. This motivates a simple constrained efficient procedure: Calculate the SDA outcome using an arbitrary tie-breaking rule. If the outcome is inefficient, successively eliminate stable improvement cycles until a constrained efficient outcome is reached. We will see that there exist solvable priority structures for which the stable improvement cycles procedure is not strategy-proof no matter how ties are broken initially and no matter how stable improvement cycles are selected. In particular, it is not sufficient to restrict attention to the stable improvement cycles procedure if one is interested in strategy-proof and constrained efficient mechanisms.

For the case of strict priorities, a number of papers have studied the relation between properties of the priority structure and the existence of mechanisms with certain desirable properties. Most prominently, Ergin (2002) studies the relationship between efficiency (with respect to student preferences) and stability. He introduces a simple but restrictive acyclicity condition that is shown to be necessary and sufficient for the compatibility of efficiency and stability.[2] Note that for the problem with strict priorities, the compatibility of strategy-proofness and constrained efficiency follows from the strategy-proofness of the SDA. At least for the unit capacity case we can formally show that Ergin's conditions are more restrictive than the conditions required for

[2]Another example is Kesten (2006) who derives conditions under which the SDA coincides with the *top trading cycles* algorithm, which has been one of SDA's main competitors in applications to the school choice problem.

solvability.[3]

Finally, we mention the recent paper by Abdulkadiroglu, Che, and Yasuda (2008) who study the school choice problem with equal priorities from an ex-ante cardinal welfare perspective. They introduce a "choice augmented deferred acceptance algorithm" (CADA) in which students submit an ordinal ranking of schools and also specify a target school. The auxiliary message is used as a tie-breaking device and can be interpreted as allowing a student to signal the intensity of her preference for the target school. In a model with a continuum of students, the CADA is shown to improve upon the SDA with fixed tie-breaking from an ex-ante perspective. The approach of Abdulkadiroglu, Che, and Yasuda (2008) differs from ours as we concentrate on the classical ex-post welfare perspective in a model with a finite number of students.

## 2.2 The School Choice Problem with Weak Priorities

A *School Choice Problem with Weak Priorities* is given by

- a finite set of students $I$,
- a finite set of schools $S$,
- a vector of capacities $q = (q_s)_{s \in S}$,
- a profile of weak priority orders of schools $\succeq = (\succeq_s)_{s \in S}$, and
- a profile of strict student preferences $R = (R_i)_{i \in I}$.

The only difference to the school choice problem with strict priorities introduced in Chapter I.2 is that two distinct students $i$ and $i'$ can now have *equal priority* for a school $s$, denoted by $i \sim_s i'$. Remember that $i \succ_s i'$ means that $i$ has strictly higher priority for school $s$ than student $i'$. For two subsets $J, J' \subset I$, we denote by $J \succ_s J'$ that $i \succ_s i'$ for all $i \in J$ and $i' \in J'$. Note that we continue to assume students can never be indifferent between two distinct schools. As everything else is fixed, we will think of a (school choice) problem as being given by a profile $R$ of student preferences. A *rule*, or *matching mechanism*, is a function that assigns a matching to each problem. A *correspondence* is a function that assigns a non-empty set of

[3]Although Ergin (2002)'s conditions are for the case of strict priorities, it is easy to see that they guarantee the compatibility of efficiency and stability when imposed on the priority structure of specialized schools in our model. If one demands that all constrained efficient matchings should be constrained efficient, stronger conditions are required (Ehlers and Erdil (2009)).

matchings to each problem and rule $f$ is a *selection* from correspondence $F$, if $f(R) \in F(R)$ for all problems $R$.

Remember that a matching $\mu$ is *stable* (or *fair*) for the school choice problem $R$, , if it

(i) is *individually rational*, that is, $\mu(i) R_i i$ for all students $i \in I$,

(ii) *eliminates justified envy*, that is, there is no student school pair $(i, s)$ such that $s P_i \mu(i)$ and $i \succ_s i'$ for some $i' \in \mu(s)$, and

(iii) is *non-wasteful*, that is, there is no student school pair $(i, s)$ such that $s P_i \mu(i)$ and $|\mu(s)| < q_s$.

At this point it is important to note that stability only depends on strict rankings in the priority structure. It is known (cf example 2.15 in Roth and Sotomayor (1991)) that in the presence of ties in the priority structure there may not exist a stable matching $\mu$ that all students weakly prefer to any other stable matching. However, given the finiteness of the problem there always exists (at least one) stable assignment which is not Pareto dominated by any other stable matching with respect to student welfare. We call a matching with this property *constrained efficient* and given some profile of strict student preferences $R$ we denote by $OS^{\succeq}(R)$ the set of constrained efficient matchings.

If priorities are strict, $OS^{\succeq}(R)$ contains exactly one matching which can be found by applying the SDA. However, if there are ties in the priority structure the SDA cannot be applied unless we specify some rule for breaking ties. Formally, a *fixed tie-breaking rule*, or *strict transformation*, of $\succeq$ is a strict priority structure $\succ'$ that preservers the strict ranking of $\succeq$, that is, $i \succ'_s j$ whenever $i \succ_s j$. Let $ST(\succeq)$ denote the set of all strict transformations of $\succeq$. Given some $\succ' \in ST(\succeq)$, let $f^{\succ'}$ denote the matching mechanism that associates the outcome of the SDA with strict priority structure $\succ'$ to each problem. It is known (Dubins and Freedman (1981), Roth (1982)) that for all $\succ' \in ST(\succeq)$, $f^{\succ'}$ is strategy-proof and stable with respect to $\succeq$. Erdil and Ergin (2008) note that $f^{\succ'}$ may, however, fail to be constrained efficient. Abdulkadiroglu, Pathak, and Roth (2009) aim to provide a rationale for the SDA with a fixed tie-breaking rule and show that no strategy-proof mechanism can dominate $f^{\succ'}$ for any $\succ' \in ST(\succeq)$, that is, there is no strategy-proof mechanism $g$ such that for all problems $R$, $g_i(R) R_i f^{\succ'}(R)$ for all $i \in I$, with at least one strict preference for at least one problem and at least one student. Of course, this dominance relation is very strong so that the set of mechanisms that are undominated in this sense is very large.

Recently, Erdil and Ergin (2008) introduced an algorithm that always produces a constrained efficient matching for weak priority structures. The main idea is that whenever a stable matching is not constrained efficient, then it is possible to increase student welfare via a cyclical exchange that respects stability constraints. More formally, let $\mu$ be a stable matching for some $R$. Then student $i$ *desires school* $s$ *at* $\mu$ if $sP_i\mu(i)$. For each school $s$, let $D_s(\mu)$ denote the set of highest $\succeq_s$-priority students among those who desire $s$ at $\mu$. A *stable improvement cycle (SIC)* at $\mu$ and $R$ consists of $m$ distinct students $i_1, \ldots, i_m$ such that for all $l = 1, \ldots, m$, $i_l \in D_{\mu(i_{l+1})}(\mu)$ (where $m + 1 := 1$). Erdil and Ergin (2008) show that $\mu \in OS^{\succeq}(R)$ if and only if $\mu$ admits no stable improvement cycle (SIC) at $\mu$ and $P_I$. This leads them to suggest the following procedure to achieve a constrained efficient outcome.

**The Stable Improvement Cycles Algorithm**

Select a fixed single tie-breaking rule and compute the associated SDA outcome given the submitted preferences of students.

If the outcome is not constrained efficient, allow students involved in a SIC to realize the corresponding cyclical exchange. Continue with this procedure until we arrive at a constrained efficient matching.

As shown in Erdil and Ergin (2008), this procedure is not, in general, strategy-proof. This is not necessarily a fault of the stable improvement cycles as the same authors show that there exist weak priority structures $\succeq$ which do not admit *any* strategy-proof selection from $OS^{\succeq}$.

Motivated by this result, we call a priority structure $\succeq$ *solvable*, if there exists a strategy-proof and constrained efficient selection from $OS^{\succeq}$. Our main goal is to characterize the class of solvable priority structures. In the next section we start with a motivating example.

## 2.3 Motivating Preference Based Tie-Breaking

We consider the following school choice environment with three students $1, 2, 3$: There are six schools $s_1, \ldots, s_6$ at which all three students have different priorities and one school $s_7$ at which all students have equal priority. All schools can admit at most one student. The priority orderings of the various schools are summarized in the following table.[4]

[4]This notation means that e.g. at $s_1$, 1 has the highest, 2 has the second highest, and 3 has the lowest priority.

| $\succsim_{s_1}$ | $\succsim_{s_2}$ | $\succsim_{s_3}$ | $\succsim_{s_4}$ | $\succsim_{s_5}$ | $\succsim_{s_6}$ | $\succsim_{s_7}$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 3 | 3 | 1, 2, 3 |
| 2 | 3 | 1 | 3 | 1 | 2 | |
| 3 | 2 | 3 | 1 | 2 | 1 | |

To implement the Erdil-Ergin procedure, we have to choose a fixed tie-breaking rule. Due to the symmetries of the example, we may assume without loss of generality that the strict transformation $\succ'_{s_7}$: 1, 2, 3 was chosen. Consider the following preference profile:[5]

| $R$ | $R_1$ | $R_2$ | $R_3$ |
|---|---|---|---|
| | $s_4$ | $s_7$ | $s_7$ |
| | $s_7$ | | |

For this school choice problem the outcome of the SDA with the above fixed tie-breaking rule is constrained efficient: 1 obtains her first choice school $s_4$ and 2 obtains her first choice school $s_7$. Student 3 does not receive a place at some school since she did not rank enough schools. Now suppose that instead of ranking only $s_7$, 3 declares $s_4$ to be her second choice, i.e. claims that her preferences are $R'_3$ : $s_7, s_4$. If all other students submit the same ranking as before, the SDA outcome is not constrained efficient for the above fixed tie-breaking rule: 1 obtains $s_7$, since she was randomly given the highest priority for this school, and 3 obtains $s_4$, since she has higher priority for this school than 1 but (randomly chosen) lower priority for $s_7$ than 2. There is a unique stable improvement cycle in this example since 1 and 3 would prefer to trade places, which cannot be vetoed by student 2. But this means that the stable improvement cycles procedure would now assign 3 a place at $s_7$, her true top choice school (under $R_3$). Hence, 3 has a strict incentive to manipulate the procedure at the original preference profile $R$. Since the ordering was chosen at random, this shows that no matter which fixed tie-breaking rule is used, there does not exist a strategy proof rule for selecting stable improvement cycles. The following procedure, however, is strategy-proof and constrained efficient:

Calculate the SDA outcome assuming that $s_7$ has unlimited capacity and let $\mu^1$ be the resulting temporary assignment. If $\mu^1(s_7) = \{1, 2, 3\}$, reject 3. If $\mu^1(s_7) = \{i_1, i_2\}$ with $i_1 < i_2$, reject $i_1$ if and only if the third student, $i_3$, is temporarily matched to a school $s$ such that $i_2 \succ_s i_1$ and $i_3 \succ_s i_1$. Now continue with the SDA in which students propose down their lists starting with their most preferred school that has not rejected them yet. Should another

[5]Remember that the above notation means that agent 1 strictly prefers school $s_3$ to school $s_7$, and that $s_4$ and $s_7$ are the only schools which agent 1 prefers to not being assigned to any school.

tie-breaking become necessary, apply the same rules as above.

An exact proof of strategy-proofness and constrained efficiency of this rule is deferred to the next sections. Note that the above tie-breaking rule takes the indexing of students as a baseline, which is only modified if two students are matched to $s_7$ while the third student is matched to a school in $s_1, \ldots, s_6$ at which she does not have *lowest priority*. This rule ensures that a student can affect the tie-breaking decision only if she changes her own temporary assignment prior to tie-breaking. By the strategy-proofness of the SDA procedure with fixed tie-breaking, it is clear that such a manipulation can, by itself, not be profitable. But the rule for tie-breaking ensures that if a student affects the tie-breaking decision, she will be matched to the new temporary assignment. The intuition for constrained efficiency is similar. This shows - a formal proof is deferred to later sections - that if we are interested in identifying the class of solvable priority structures, it is not sufficient to restrict attention to the stable improvement cycles procedure.

## 2.4 The (Non-)Specialized Schools Model

In this paper we consider a restricted class of school choice environments with two types of schools: *Specialized schools* have a strict priority ranking, while *non-specialized* schools assign the same priority to every student. More formally, we have the following.

**Definition 3.** *The priority structure $\succeq$ is a* ***(non-)specialized schools environment****, if there exists a partition of $S$ into two non-empty sets $S^0$ and $S^1$ such that*

*(i) $S^0$ comprises the set of non-specialized schools, that is, for all $s \in S^0$ and all $i, j \in I$, $i \sim_s j$, and*

*(ii) $S^1$ comprises the set of specialized schools, that is, for all $s \in S^1$ and all $i, j \in I$ such that $i \neq j$, $i \succ_s j$ or $j \succ_s i$.*

In this language schools $s_1, \ldots, s_6$ in the example of section 3 were specialized, while $s_7$ was the only non-specialized school. One interpretation of this model is that a specialized school's priority ordering result from subject test(s) in the discipline(s) relevant for this school, e.g. a sports oriented school makes admission contingent on sports trials. Non-specialized schools on the other hand offer general educational training and therefore do not discriminate between students. Apart from this interpretation, our motivation for studying the (non-)specialized schools model is twofold.

First, our analysis is an important initial step towards a characterization of solvable priority structures in the school choice problem with indifferences in priority orders. We consider the extreme situation where a school's priority order has either one or $|I|$ indifference classes. However, it will become clear that our results also have important implications for the general school choice problem whenever some or all schools assign equal priority to a group of students which is sufficiently large, but at the same time potentially much smaller than $|I|$.

Secondly, our model bridges the gap between two important environments, which have been studied extensively in the literature.

(i) If $S^1 = \emptyset$, then all students have equal priority at all schools. This case is known as the *house allocation problem*.[6] Among others, Svensson (1999), Papai (2000), and Pycia and Unver (2009), are interested in identifying rules which satisfy strategy-proofness and efficiency.[7] Since all students have equal priority at all schools, stability is vacuously satisfied by any rule and constrained efficiency is equivalent to efficiency. The class of strategy-proof and efficient rules is very large and has not been characterized in the literature.

(ii) If $S^0 = \emptyset$, then no two students have equal priority at a school and we are back to the school choice problem with strict priorities from Chapter I.2. For this problem the student optimal stable rule is the only strategy-proof and constrained efficient rule.[8]

In the presence of both specialized and non-specialized schools a strategy-proof and constrained efficient mechanism does not always exist and we derive tractable conditions under which a priority structure is solvable.

For the remainder of this paper we restrict attention to the (non-)specialized schools environment. It is important to keep in mind that stability constraints only come from the priority orders of specialized schools. In the following, we will denote the priority ordering of a specialized school $s \in S^1$ by $\succ_s$ instead of $\succeq_s$ to emphasize that no two students can have equal priority. Let $\succ^1 = (\succ_s)_{s \in S^1}$ be the priority structure of specialized schools. It is easy to see that the conditions for solvability can only concern $\succ^1$. We assume throughout that there are at least two specialized schools and at least two non-specialized schools.[9] We first consider the case of unit capacity at all schools to develop intuition for the requirements imposed by solvability.

[6] This problem was first studied by Hylland and Zeckhauser (1979).

[7] In fact, all these papers derive characterizations of rules that satisfy strategy-proofness, efficiency, and different sets of additional axioms.

[8] In fact, the student optimal stable rule is also the only strategy-proof and stable rule (Alcalde and Barbera (1994)).

[9] If there is only one specialized school, solvability is trivial. If there is only one non-specialized school, the sufficient conditions for solvability are slightly different as we discuss below.

### 2.4.1 Unit capacities - Necessary Conditions

Throughout this section we consider the case where all schools can admit at most one student, i.e. where $q_s = 1$ for all $s \in S$. Of course, schools can usually admit more than one student and the reader may prefer to think of the allocation of tasks in a society rather than the school choice problem for this section. Society has a strict preference over who takes on specialized tasks, while it is indifferent as to who takes on a non-specialized task.

The example in Section 2.3 suggests that strategy-proofness and constrained efficiency are always compatible if there are at most three students (a formal proof can be found below). Not surprisingly, this positive result does not extend to the case of four or more students as we will shortly see. In this section we identify two related sources for the incompatibility between strategy-proofness and constrained efficiency. The first source is introduced in the following definition.

**Definition 4.** *Let $\succeq$ be a non-specialized schools environment with unit capacities. Then $\succ^1$ contains **an ambiguous** $1$**-tie** if there exist two specialized schools $s_1, s_2 \in S^1$ and four distinct students $i_1, i_2, i_3, i_4 \in I$ such that both $i_1 \succ_{s_1} i_3 \succ_{s_1} i_2$ and $i_2 \succ_{s_2} i_4 \succ_{s_2} i_1$.*

To see the problems associated with ambiguous 1-ties, consider the smallest example where it can be violated: There are four students $1, \ldots, 4$, two specialized schools $s_1, s_2$, and one non-specialized school $s_3$. The priority structure $\succ^1$ is such that (the remaining rankings are irrelevant)

$$1 \succ_{s_1} 3 \succ_{s_1} 2 \text{ and } 2 \succ_{s_2} 4 \succ_{s_2} 1.$$

Now consider the preference profile

| $R$ | $R_1$ | $R_2$ | $R_3$ | $R_4$ |
|---|---|---|---|---|
| | $s_2$ | $s_1$ | $s_3$ | $s_3$ |
| | $s_3$ | $s_3$ | | |
| | $s_1$ | $s_2$ | | |

For this profile, 1 and 2 would prefer to *exchange* their priorities for $s_1$ and $s_2$. Here, this is not problematic since neither 3 nor 4 are interested in these schools. However, either 3 or 4 will have to remain unassigned since $s_3$ cannot admit more than one student. A strategy-proof procedure has to ensure in particular that 3 and 4 cannot profit by claiming that $s_1$ and/or $s_2$ are acceptable, respectively. We show in Appendix A.2 that this cannot be achieved by *any* constrained efficient mechanism, thus proving the following result.

**Proposition 5.** *Let $\succeq$ be a (non-)specialized schools environment with unit capacities. Then $\succeq$ is solvable only if $\succ^1$ does not contain an ambiguous* 1*-tie.*

The absence of ambiguous 1-ties is a strong restriction. Suppose for example that $s_1$ is a music oriented school while $s_2$ is a sports oriented school. Both schools assign priorities according to performance in auditions. Typically, there will be *allrounders* who do relatively well in both specializations. At the same time, there will also be *specialists* who have a musical talent but are not very sportive (and the other way around). If there are at least two allrounders and two specialists, the priority structure is not solvable since it contains an ambiguous 1-tie. However, there is still some scope for different priority orderings across specialized schools as the next example demonstrates.

**Example 5.** *There are four students* $1, \ldots, 4$*, six specialized schools* $s_1, \ldots, s_6$*, and one non-specialized school* $s_7$*. The priority structure at specialized schools is as follows.*

| $\succ_{s_1}$ | $\succ_{s_2}$ | $\succ_{s_3}$ | $\succ_{s_4}$ | $\succ_{s_5}$ | $\succ_{s_6}$ |
|---|---|---|---|---|---|
| 1 | 2 | 1 | 1 | 2 | 2 |
| 3 | 3 | 2 | 2 | 1 | 1 |
| 2 | 1 | 4 | 3 | 4 | 3 |
| 4 | 4 | 3 | 4 | 3 | 4 |

*Note that $\succ^1$ does not contain an ambiguous 1-tie. Thus, in principle the door remains open for possibility results.*

*However, it is easy to see that any fixed tie breaking rule leads to violations of constrained efficiency for some preference profiles so that some or all of the ties have to broken preference based. A natural candidate for a constrained efficient assignment procedure is the following: Set* $1 \sim^0 2 \succ^0 3 \succ^0 4$ *and break ties at the non-specialized school $s_7$ according to this ordering. Thus, only the tie between 1 and 2 remains to be broken endogenously. Now if 1 and 2 apply to $s_7$ in some round of the SDA procedure, temporarily ignore the capacity constraint at $s_7$. If 3 is temporarily matched to $s_2$ by the end of the SDA procedure, 1 is rejected by $s_7$. In any other case 2 is rejected. While this certainly guarantees a constrained efficient allocation, 4 can manipulate the tie breaking decision to her benefit. To see this consider the profile*

| $R$ | $R_1$ | $R_2$ | $R_3$ | $R_4$ |
|---|---|---|---|---|
| | $s_7$ | $s_7$ | $s_3$ | $s_1$ |
| | $s_3$ | $s_1$ | $s_2$ | |

*Here, 4 would be left unmatched while 3 obtains a place at $s_3$. However, if she claims that $R'_4 : s_3, s_1$, 3 would be rejected by $s_3$ and would subsequently apply to $s_2$. But then 1 would be rejected by $s_7$, causes 4 to be rejected at $s_3$, and 4 ultimately obtains a place at her true top choice (under $R_4$) school $s_1$. Hence, the above procedure is not strategy-proof.*

The following definition formalizes one problematic feature of the priority structure in this example.

**Definition 5.** *Let $\succeq$ be a (non-)specialized schools environment with unit capacities. Then $\succ^1$ contains* ***ambiguity at the top*** *if there are four distinct students $i_1, i_2, i_3, i_4$ and three distinct specialized schools $s_1, s_2, s_3 \in S^1$ such that $i_1 \succ_{s_1} i_3 \succ_{s_1} i_2 \succ_{s_1} i_4$, $i_2 \succ_{s_2} i_3 \succ_{s_2} i_1 \succ_{s_2} i_4$, and $\{i_1, i_2\} \succ_{s_3} i_4 \succ_{s_3} i_3$.*

In the above example there was ambiguity at the top concerning $\succ_{s_1}, \succ_{s_2}$, and $\succ_{s_3}$. In order to avoid ambiguity at the top, at least one of the schools' priority orderings needs to be changed. For example, we could set $\tilde{\succ}_{s_2}$ equal to any of the priority orderings of the other specialized schools to obtain a priority structure that contains no ambiguous 1-ties and no ambiguity at the top. The following shows that ambiguity at the top is the second source for the incompatibility of strategy-proofness and constrained efficiency.

**Proposition 6.** *Let $\succeq$ be a (non-)specialized schools environment with unit capacities. Then $\succeq$ is solvable only if $\succ^1$ does not contain no ambiguity at the top.*

Above we showed that our intuitive idea for achieving a constrained efficient matching does not provide students with the right incentives. Note that the statement of Proposition 6 is much stronger since it says that *any* assignment procedure has to sacrifice either strategy-proofness or constrained efficiency.

### 2.4.2 General Capacities - Sufficient Conditions

The results in the last section support a pessimistic view about the possibilities of obtaining strategy-proof and constrained efficient mechanisms. It is important to keep in mind that we assumed that all schools could admit at most one student. In this section we turn to the case of general capacities. We derive a precise connection between the capacity vector and the *amount of variability* in school rankings allowed by a solvable priority structure. We first consider the case of identical capacities at all specialized schools.

### Symmetric Capacities at Specialized Schools

In this subsection we concentrate on the case of identical capacities at all specialized schools. We assume for now that the set of students is *connected* in the sense that there is no strict subset $J \subset I$ such that $J \succ_s I \setminus J$ for all $s \in S^1$. We discuss below how our results translate to the case where this assumption is not satisfied.

As a first step we derive an equivalent formulation of the two necessary conditions for the unit capacity case of the last section. The idea is to then adapt these conditions to the case of general symmetric capacities, where a full characterization seems to be out of reach as we discuss in section 2.5. We require a bit of additional notation and terminology: For all $s \in S^1$ and $k \in \{1, \dots, |I|\}$, $r_k(\succ_s)$ denotes the student who has $k$th highest priority for $s$, i.e. $|\{i \in I : i \succ_s r_k(\succ_s)\}| = k-1$. For $k \in \{1, \dots, |I|\}$, let $L_k = (\cup_{s \in S^1}\{r_k(\succ_s)\}) \backslash (L_1 \cup \dots \cup L_{k-1})$ denote the set of students who have $k$th highest priority at some specialized school but never rank higher. Let $K$ be the smallest integer such that $N = L_1 \cup \ldots \cup L_K$, so that in particular $L_k = \emptyset$ for all $k > K$. We have the following.

**Proposition 7.** *If $|I| > 3$ and $\succ^1$ does not contain ambiguous 1-ties or ambiguity at the top then*

***O1*** *$L_k \subseteq \{r_k(\succ_s), r_{k+1}(\succ_s), r_{k+2}(\succ_s)\}$ for all $s \in S^1$ and $k \in \{1, \dots, K\}$, and*

***O2*** *there is exactly one student in $L_1$ who has third highest priority at some specialized school.*

*Conversely, if $|I| > 3$ and $\succ^1$ satisfies **O1** and **O2** then $\succ^1$ does not contain ambiguous 1-ties or ambiguity at the top.*

Propositions 5 - 7 imply that a student's rank in priority orderings can differ by at most two across specialized schools if all schools can admit at most one student and the priority structure is solvable. This allows us to define a global ordering $\succeq^0$ on $I$ by setting $i_1 \succ^0 i_2$ if $i_1 \in L_k$ and $i_2 \in L_{k'}$ for some $k < k'$. The key property here is that if $i_1 \succ^0 i_2$ for two students $i_1, i_2$, there cannot be a third student $i_3$ and a specialized school $s \in S^1$ such that $i_2 \succ_s i_3 \succ_s i_1$. As we show below this implies that ties between two students who are strictly ordered according to $\succeq^0$ can be broken exogenously, i.e. without conditioning on student preferences. Thus, only ties between students in the same indifference set of $\succeq^0$ remain to be broken according to student preferences. Condition **O2** implies that if $|I| \geq 4$, we never have to consider the preferences of lower priority students in $L_2 \cup \ldots \cup L_K$ in order to break the tie between the two students in $L_1$. Note that since $|L_1| = 2$ this basically means that the tie between the two students at the top can be broken exogenously if $|I| \geq 4$.

Intuitively, it is clear that increasing capacities should enlarge the scope for preference based tie breaking and should increase the allowable variability in priority orderings of specialized schools. The important task here is to identify the exact form of this relationship. We now show how the conditions for solvability from the unit capacity case can be adapted to the capacity vector. For the following, we fix a capacity vector for schools with the property that all specialized schools have the same capacity. Let $q^1$ be the common capacity of all specialized schools, let $q^0_{(1)}$ be the lowest capacity of any non-specialized school, and $q^0_{(2)} \geq q^0_{(1)}$ be the second lowest capacity.[10]

First of all, the priority structure is solvable if the number of students is sufficiently small compared to available capacities. Here, the critical value turns out to be $p = q^1 + q^0_{(1)} + \min\{q^1, q^0_{(2)}\}$. To see that any priority structure is solvable if $|I| \leq p$ note that if tie-breaking becomes necessary, i.e. at least $q^0_{(1)} + 1$ students are interested in the same non-specialized school $s_1 \in S^0$, at most one specialized school can have filled its capacity. Furthermore, if some specialized school $s_2 \in S^1$ has filled its capacity, there cannot be a third school $s_3 \in S \setminus \{s_1, s_2\}$ that has to reject any student. We show below how the priority ordering of $s_2$ can be used to determine who should be rejected by $s_1$. Secondly, if $|I| > p$ the variability of priority orders across specialized schools has to be restricted. The next definition formally summarizes our requirements in this case.

**Definition 6.** *Suppose that $I$ is connected and that $|I| > p$. Then $\succ^1$ satisfies **limited** $p$-**variability** if*

***O1(p)*** *$L_k \subset \{r_k(\succ_s), \ldots, r_p(\succ_s)\}$ for all $k \leq p-2$ and all $s \in S^1$,*

***O2(p)*** *$L_k \subset \{r_k(\succ_s), r_{k+1}(\succ_s), r_{k+2}(\succ_s)\}$ for all $p-2 < k$ and all $s \in S^1$, and*

***O3(p)*** *there is exactly one student in $L_1 \cup \ldots \cup L_{p-2}$ who has pth highest priority at some specialized school.*

The idea behind this condition is that we want to assign high priority for non-specialized schools to students who have *high*, i.e. at least $(p-2)$nd highest, priority for specialized schools. Note that the amount of allowable variability declines as we move down the rankings of specialized schools. This is because the demands of students with higher priority could effectively lead to a reduction of the number of seats at some schools. Eventually, everything

[10] Remember that we assumed $|S^0| \geq 2$.

reduces to the unit capacity case and a student's priority can vary by at most two. This is illustrated by the following example.

**Example 6.** *There are four specialized schools* $s_1, \ldots, s_4$*, two non-specialized schools* $s_5, s_6$*, and six students* $1, \ldots, 6$*. Capacities are* $q^1 = 2$ *and* $q_{s_5} = 1$ *and* $q_{s_6} = 2$*. Priorities of specialized schools are given by*

$$\begin{array}{llllll} \succ_{s_1}: & 1 & 2 & 3 & 4 & 5 & 6 \\ \succ_{s_2}: & 4 & 3 & 2 & 1 & 5 & 6 \\ \succ_{s_3}: & 2 & 1 & 3 & 6 & 4 & 5 \\ \succ_{s_4}: & 3 & 2 & 1 & 5 & 4 & 6 \end{array}$$

*Note that for this priority structure the set of all students is connected and* $|I|$ *exceeds the critical value of* $p = 5$*. We have* $L_1 \cup L_2 \cup L_3 = \{1, 2, 3, 4\}$ *and* $L_4 = \{5, 6\}$*. Since no student in* $L_1 \cup L_2 \cup L_3$ *is ranked lower than fifth and only* 4 *is ranked fifth (at schools* $s_3$ *and* $s_4$*)* $\succ^1$ *satisfies limited p-variability.*

*Consider again the interpretation of priorities at specialized schools as being determined by test scores. If all schools had unit capacity, the priority structure would not be solvable: 1 and 4 would then be specialists for schools* $s_1$ *and* $s_2$*, respectively, while 2 and 3 would be allrounders. However, given the above capacity vector 1 and 4 are not* too *specialized and we will see below that the above priority structure is solvable.*

Intuitively, assigning high priority to students in the *upper segment* of students who rank at least $(p-2)$nd at some specialized school minimizes the number of rejections following a rejection at a non-specialized school. It is important to note that limited p-variability is a joint condition on $\succ^1$ and the capacity vector. In particular, the capacity vector determines the size of the upper segment. In case $I$ exceeds the critical value of $p$, limited p-variability ensures that ties between two students in the upper segment can always be broken conditional only on the preferences of other upper segment students. Since the size of the upper segment cannot exceed $p$, this is always possible as argued above and formally proven below.[11] At this point a few remarks about limited p-variability are in order.

**Remark 1:**

(i) Conditions **O1(p)** and **O2(p)** imply that $|L_1| \leq p$ and that $I = L_1$ if $|L_1| = p$. Now let

[11] Actually, limited p-variability ensures that the upper segment contains exactly $p-1$ students if $|I| > p$. The reasoning behind restricting the upper segment of students to those ranking no lower than $(p-2)$nd (and not $(p-1)$st) is a bit subtle and will become clear in the proof Theorem 2 in Appendix A.2.

$K$ be the minimal integer such that $N = L_1 \cup \ldots \cup L_K$ so that in particular $L_k = \emptyset$ for all $k > K$. If $|I| > p$, **O1(p)** and **O2(p)** imply that

· $|L_1 \cup \ldots \cup L_{p-2}| = p - 1$,

· $|L_k| = 1$ for all $p - 2 < k \leq K - 1$, and

· $|L_K| \in \{1, 2\}$.

In particular we must have $|I| \in \{K + 1, K + 1\}$ if $|I| > p$.

(ii) A major benefit of limited p-variability is that it is tractable and very easy to verify. If $|I| > p$, we first need to check that $L_1 \subset \{r_1(\succ_s), \ldots, r_p(\succ_s)\}$ for all specialized schools $s \in S^1$. This can be implemented as follows: Take an arbitrary specialized school $s \in S^1$. Then check whether $r_1(\succ_s) \in \{r_1(\succ_{s'}), \ldots, r_p(\succ_{s'})\}$ for all $s' \in S^1 \setminus \{s\}$. This requires at most $(|S^1| - 1)p$ steps. Proceeding in this fashion, we can test whether no student in $L_1$ is ever ranked lower than $p$th in at most $|S^1|(|S^1| - 1)p < (|S^1|)^2|I|$ steps.

Now, the conditions for the remaining $L_k$ sets can be verified completely analogously so that checking **O1(p)** and **O2(p)** requires at most $K|S^1|(|S^1| - 1)p < (|I|)^2(|S^1|)^2$ steps. Note that **O3(p)** can be tested at (almost) no additional computational cost: As soon as we find a student in $L_1 \cup \ldots \cup L_{p-2}$ who is ranked $p$th at some specialized school we have to check that all other students in this segment of the priority structure rank no lower than $(p - 1)$st.

We now design an assignment procedure that is strategy-proof and constrained efficient provided that limited p-variability is satisfied. For the following we fix a capacity vector as well as the priority structure $\succ^1$ of specialized schools and assume that $\succ^1$ satisfies limited p-variability. The procedure consists of two steps: In the first step we define an ordering $\succeq^0$ as in the unit capacity case. In the second step, we introduce a new version of the SDA algorithm which uses this ordering as the *common* priority ordering of all non-specialized schools. The procedure breaks ties between students in the same indifference class of $\succeq^0$ endogenously on basis of temporary assignments.

**Step 1: Ordering Students**

- If $|I| \leq p$, set $i \sim^0 j$ for all $i, j \in I$.

- If $|I| > p$, set

(i) $i \sim^0 j$ for all $i, j \in L_1 \cup \ldots \cup L_{p-2}$

(ii) $i \succ^0 j$ if $i \in L_k$ and $j \in L_{k'}$ with $k < k' \leq K$ and $k' \geq p$

(iii) $i \sim^0 j$ if $i, j \in L_K$

As in the unit capacity case this ordering has the property that if $i_1 \succ^0 i_2$ there cannot be a third student $i_3$ and a specialized school $s \in S^1$ such that $i_2 \succ_s i_3 \succ_s i_1$. We show below that this implies that the tie between $i_1$ and $i_2$ can be broken exogenously without violating constrained efficiency. All remaining ties can be broken endogenously. By Remark 1.(i) there are at most two non-singleton indifference sets of $\succeq^0$ if $|I| > p$ (and one indifference set if $|I| \leq p$): An *upper segment* consisting of $p-1$ students who have at least $(p-2)$nd highest priority for some specialized school and, possibly, a *lower segment* consisting of two students in $L_K$. For the purpose of breaking ties in these two segments endogenously we label students according to their position in $\succeq^0$. Within an indifference class, the label is arbitrary with the exception that if $|I| > p$ we assign the highest label $p-1$ (remember Remark 1.(i)) in the upper indifference set of $\succeq^0$ to the only student in $L_1 \cup \ldots \cup L_{p-2}$ who has $p$th highest priority at some specialized school. [12] Labels will be used as a baseline for endogenous tie breaking. This baseline is modified only if a specialized school has filled its capacity. In the following we abuse notation slightly and identify a student with her label. Thus, if $|I| > p$ we write $I = \{1, \ldots, K+1\}$ if $|L_K| = 1$ and $I = \{1, \ldots, K+2\}$ if $|L_K| = 2$, where the labeling adheres to the rules above. We are now ready to describe the SDA procedure with endogenous tie breaking (SDA-ETB).

**Step 2: The SDA with Endogenous Tie Breaking**

The algorithm takes as inputs the (relevant portion of the) capacity vector $(q^1, q^0_{(1)}, q^0_{(2)})$, the priority structure of specialized schools $\succ^1$, the ordering $\succeq^0$ calculated in Step 1, and a profile of student preferences.

**Round 1:** Each student applies to her most preferred school. Each specialized school $s \in S^1$ admits the $q^1$ highest priority students according to $\succ_s$. Each non-specialized school $s \in S^0$ admits the $q_s$ students with the lowest labels among those who aply to it. If necessary,

[12] More formally, the labeling can be described as follows: If $|I| \leq p$ choose a permutation $\pi_I : I \to \{1, \ldots, |I|\}$ at random. If $|I| > p$ let $\tilde{i}$ be the unique student in $L_1 \cup \ldots \cup L_{p-2}$ who can have $p$th highest priority at some specialized school and set $\pi_I(\tilde{i}) = p-1$.

(i) Choose a permutation $\pi_I : (L_1 \cup \ldots \cup L_{p-2}) \setminus \{\tilde{i}\} \to \{1, \ldots, p-2\}$ at random.

(ii) For $k \in \{p-1, \ldots, K-1\}$ and $i \in L_k$ set $\pi_I(i) = k+1$.

(iii) If $|L_K| = 2$ randomly pick a student $i \in L_K$ and set $\pi_I(i) = K+1$. Set $\pi_I(i') = K+2$ for the other student $i'$ in $L_K$.

it admits all students in the same indifference class of $\succeq^0$ as the $q_s$th highest labeled student who was admitted in addition. Let $\mu^1$ be the resulting temporary assignment.

If one of the rejected students has not yet applied to all acceptable schools, go to **Round 2**. If all rejected students have applied to all acceptable schools and there is a non-specialized $s \in S^0$ such that $|\mu^1(s)| > q_s$, use subroutine **TB($\mu^1$)** to determine a rejection and go to **Round 2**. Else, stop.

$\vdots$

**Round $t$:** Each student rejected in Round $t-1$ applies to her most preferred school among those that have not yet rejected one of her proposals. Each specialized school $s \in S^1$ admits the $q^1$ highest priority students according to $\succ_s$. Each non-specialized school $s \in S^0$ admits the $q_s$ students with the lowest labels among those who aply to it. If necessary, it admits all students in the same indifference class of $\succeq^0$ as the $q_s$th highest labeled student who was admitted in addition. Let $\mu^t$ be the resulting temporary assignment.

If one of the rejected students has not yet applied to all acceptable schools, go to **Round** $t+1$. If all rejected students have applied to all acceptable schools and there is a non-specialized school $s \in S^0$ such that $|\mu^t(s)| > q_s$, use subroutine **TB($\mu^t$)** to determine a rejection and go to **Round** $t+1$. Else, stop.

The crucial ingredient of this algorithm is the tie-breaking subroutine which is applied to determine a rejection at non-specialized schools. The subroutine is applied only if *nothing else moves* in the sense that there is no other way for the algorithm to proceed than to break a tie within an indifference class of $\succeq^0$.

**Subroutine TB($\mu^t$):** If there is a non-specialized school $s \in S^0$ such that $|\mu^t(s)| > q_s$ and $i \sim^0 j$ for all $i, j \in \mu^t(s)$, set $s_0 := s$ and go to **Step TB($\mu^t$).1**. Else, let $s_0 \in S^0$ be the non-specialized school such that $L_K \subset \mu^t(s_0)$ and go to **Step TB($\mu^t$).2**.

**Step TB($\mu^t$).1:** If there is a specialized school $s_1 \in S^1$ s.t. $|\mu^t(s_1)| = q^1$ and $i \sim^0 j$ for all $i, j \in \mu^t(s_1) \cup \mu^t(s_0)$ let $i_1$ be the student with the lowest priority according to $\succ_{s_1}$ among students in $\mu^t(s_0)$. School $s_0$ rejects $i_1$ if $\mu^t(s_1) \succ_{s_1} i_1$.

In any other case $s_0$ rejects student with the highest label among students in $\mu^t(s_0)$.

**Step TB($\mu^t$).2:** Let $s_1 := \mu^t(K)$. If $K+2 \succ_{s_1} K+1$, $s_0$ rejects $K+1$.

In any other case, $s_0$ rejects $K+2$.

The intuition for the tie-breaking subroutine is as follows: **Step TB($\mu^t$).1** covers tie-breaking in the upper indifference class of $\succeq^0$. It ensures that following a tie breaking decision in the upper segment there is a further rejection of a student in this segment only if it is unavoidable. **Step TB($\mu^t$).2** covers tie-breaking in the lower indifference class of $\succeq^0$. It ensures that there is no further rejection following a tie-breaking decision in the lower segment. Note that this step of the tie-breaking subroutine is reached only if $|I| > p$ and $|L_K| = 2$ since otherwise there can never be a non-specialized school that temporarily admits students from different indifference classes of $\succeq^0$ and violates its capacity constraint.

Note that of the inputs required by the mechanism everything but students' preferences are assumed to be exogenously given. In the following we supress the dependency of the outcome of the SDA-ETB on the exogenous factors for notational simplicity. Given a problem $R$ let $f^{ETB}(R)$ thus denote the associated outcome of the SDA-ETB procedure. We have the following.

**Theorem 2.** *Suppose that either $|I| \leq p$ or $\succ^{ss}$ satisfies limited p-variability. Then the following statements are true.*

*(i) $f^{ETB}(R)$ is constrained efficient for all problems $R$.*

*(ii) $f^{ETB}$ is strategy-proof.*

*In particular, $\succeq$ is solvable if either $|I| \leq p$, or $|I| > p$ and $\succ^1$ satisfies limited p-variability.*

At this point it makes sense to illustrate the SDA-ETB by means of an example.

**Example 7.** *Consider again the environment of example 6. Note that the labels of students have been chosen in accordance with our rules since* 4 *is the only student in $L_1 \cup \ldots \cup L_3$ who ranks* 5*th (at schools $s_3$ and $s_4$). Consider the following problem:*

$$
\begin{array}{c|cccccc}
R & R_1 & R_2 & R_3 & R_4 & R_5 & R_6 \\
\hline
 & s_6 & s_6 & s_3 & s_6 & s_5 & s_5 \\
 & & & & s_3 & s_1 & s_3 \\
 & & & & s_5 & & 
\end{array}.
$$

*For this problem we obtain*

$$
f^{ETB}(R) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ s_6 & s_6 & s_3 & s_3 & s_1 & s_5 \end{pmatrix}.
$$

*This example illustrates that it is important to break ties in the* upper segment $L_1 \cup L_2 \cup L_3$ *before breaking ties in the* lower segment $L_4$*: If we would have broken the tie at $s_5$ first (according* *to our rules for tie breaking), 6 would have been rejected. In subsequent rounds of SDA-ETB, student 4 would then have been rejected by* $s_6$ *and* $s_3$*. Since* $4 \succ^0 5$*, 4 would have subsequently obtained a place at* $s_5$*. But then there would be a stable improvement cycle consisting of* 4 *and* 6*. The main reason for breaking ties in the lower segment last is that this way we can ensure that there are no further rejections after tie-breaking. A similar example can be used to show that it is important to wait with endogenous tie-breaking until* nothing else moves.

**Asymmetric Capacities**

In this section we turn to the case of general capacity vectors. In the following, let $q^1_{(1)}$ be the minimal capacity of specialized schools, $q^0_{(1)}$ and $q^0_{(2)}$ be defined as in the last section, and $p = q^1_{(1)} + q^0_{(1)} + \min\{q^1_{(1)}, q^0_{(2)}\}$ be the modified critical value. It is easy to see that our previous results imply that if the set of students is connected and $\succ^1$ satisfies limited p-variability whenever $|I| > p$ then $\succeq$ is solvable.[13]

We now discuss how our results extend to the case where $I$ is not connected. Since we are dealing with a finite problem there has to exist a minimal set $J_1$ such that for any $s \in S^1$, $J_1 \succ_s I \setminus J_1$. We call $J_1$ the *minimal top set* of $I$ with respect to $\succ^1$. Proceeding inductively, let $J_t$ be the minimal top set of $I \setminus (J_1 \cup \ldots \cup J_{t-1})$ with respect to $\succ^1$. We call $(J_t)_{t\geq 1}$ the *minimal top set partition* of $I$ with respect to $\succ^1$. Suppose for the sake of clarity that $I = J_1 \cup J_2$. Let $f^{ETB}|_{J_1}$ denote the SDA-ETB mechanism when we make all places at all schools available to students in $J_1$. Since $J_1$ is connected, our previous analysis implies that $f^{ETB}|_{J_1}$ is strategy-proof and constrained efficient provided that $\succ^1|_{J_1}$ satisfies limited p-variability. In principle, there are two ways to guarantee that there is a strategy-proof and constrained efficient procedure for students in $J_2$ which we now discuss. In both cases, we assign all students in $J_1$ with higher priority for non-specialized schools than all students in $J_2$. This ensures in particular that a student in $J_1$ can never envy a student in $J_2$.

(i) In some instances it might be feasible to elicit reports from students in $J_2$ after assignments for students in $J_1$ have been determined. In this case given a profile $R_{J_1}$ elicited from students in $J_1$, we can reduce capacities at schools according to $f^{ETB}|_{J_1}(R_{J_1})$. Let $p^1$ be the resulting modified critical value and let $f^{ETB}|_{J_2}$ denote the SDA-ETB that allocates remaining places among students in $J_2$. Again, our analysis from the connected case implies that $f^{ETB}|_{J_2}$ is strategy-proof and constrained efficient provided that $\succ^1|_{J_2}$ satisfies limited $p^1$-variability.

[13] One just needs to replace $q^1$ with the actual capacities of specialized schools in the formulation of the SDA-ETB. Everything else remains exactly the same.

(ii) If we restrict attention to assignment procedures that simultaneously elicit a report from all students, the restrictions for solvability in $J_2$ become more restrictive. We have to require solvability for the *lowest* possible critical value that could be induced by the demands of students in $J_1$. For example, consider the case $q^1_{(1)} = 4$, $q^0_{(1)} = q^0_{(2)} = 2$, and $|J_1| = 3$. Here, the worst case would be if all students in $J_1$ were interested in the minimal capacity specialized school leading to a new critical value of 4.

From the above discussion it is clear that even when all specialized schools initially have identical capacities, we have to consider the case of asymmetric capacities if $I$ is not connected since the demands of students in $J_1$ may lead to a problem with asymmetric capacities for the remaining student population.

However, note that in the unit capacity case the critical value is always 3. This implies that the same conditions guaranteeing solvability for the connected case also guarantee solvability for the general case. Hence, we obtain the following theorem as a corollary to Propositions 5 -7 and Theorem 2.

**Theorem 3.** *Suppose $\succeq$ is a (non-)specialized schools environment with unit capacities. Then $\succeq$ is solvable if and only if $\succ^1$ does not contain ambiguous 1-ties or ambiguity at the top.*

To conclude this section note that it could be the case that even though $\succeq$ is not solvable, there is a strategy-proof and constrained efficient procedure for a subpopulation of students. To see this consider again the case of unit capacities and suppose that $I = J_1 \cup J_2$. If $\succ^1 |_{J_1}$ satisfies limited 3-variability, but $\succ^1 |_{J_2}$ does not, there is a strategy-proof and constrained efficient mechanism for students in $J_1$ but not for students in $J_2$.

## 2.5 Conclusion and Discussion

This chapter derived a full characterization of solvable priority structures in (non-)specialized schools environments with unit capacity. Significantly weaker sufficient conditions were introduced for the case of general capacity vectors. Our conditions show precisely how much variability in priority orderings across specialized schools can be allowed in order to guarantee existence of a constrained efficient and strategy-proof mechanism. The proof of sufficiency was constructive and used a modified deferred acceptance procedure with (potentially) preference based tie-breaking. The results show that it is not sufficient to concentrate on fixed tie-breaking rules if one is interested in strategy-proof and constrained efficient school choice systems. Furthermore, the scope for preference based tie-breaking increases in the number of slots available at schools. We now discuss several important open questions.

### 2.5.1 Uniqueness of the Tie-Breaking Rule

In this chapter we introduced tractable conditions that guarantee solvability of a priority structure. Given that these conditions are satisfied we introduced the strategy-proof and constrained efficient SDA-ETB procedure. One important idea of this mechanism was to assign those students who have high priority at specialized school also high priority for all non-specialized schools. This could be considered problematic from an equity perspective and school choice authorities might be interested in knowing whether there are other strategy-proof and constrained efficient mechanisms. In the following we discuss whether there could be other ways to break ties. In this section we concentrate on the case of unit capacities at all schools, assume that $I$ is connected, and fix a solvable environment $\succeq$.

First, suppose we have set $i_1 \succ^0 i_2$ for two students $i_1, i_2 \in I$ so that $i_2$ can never obtain a non-specialized school desired by $i_1$. Can there be strategy-proof and constrained efficient procedure that exogenously breaks ties at non-specialized schools in favor of $i_2$? To see that this is impossible, note that the construction of $\succeq^0$ ensures that there exists a specialized school $s \in S^1$ and a third student $i_3$ such that $i_1 \succ_s i_3 \succ_s i_2$. Let $\tilde{s} \in S^0$ be one of the non-specialized schools.[14] Now let $f$ be a strategy-proof and constrained efficient mechanism such that $i_2$ always has higher priority for $\tilde{s}$ than $i_1$. Consider first the problem

$$\begin{array}{c|ccc} R^1 & R^1_{i_1} & R^1_{i_2} & R^1_{i_3} \\ \hline & \tilde{s} & s & \tilde{s} \\ & & \tilde{s} & s \end{array}.$$

We must have $f_{i_1}(R^1) = 1$, $f_{i_2}(R^1) = s$, and $f_{i_3} = \tilde{s}$. Otherwise there would be a stable improvement cycle given that $i_2$ can never envy $i_1$ for $\tilde{s}$. Now suppose that $R^2_{i_3} : \tilde{s}$ and consider $R^2 = (R^1_{i_1}, R^1_{i_2}, R^2_{i_3})$. By strategy-proofness we must have $f(R^2) = f(R^1)$. Next, let $R^2_{i_1} : \tilde{s}, s$ and $R^3 = (R^2_{i_1}, R^1_{i_2}, R^2_{i_3})$. By strategy-proofness and stability, we must have $f_{i_1}(R^3) = s$. Constrained efficiency then implies $f_{i_2}(R^3) = i_2$ and $f_{i_3}(R^3) = \tilde{s}$. Finally, let $R^3_{i_3} = s, \tilde{s}$ and consider $R^4 = (R^2_{i_1}, R^1_{i_2}, R^3_{i_3})$. Since $i_2$ cannot envy $i_1$ for a place at $\tilde{s}$ and $i_2$ cannot obtain a place at $s$ given $R^3_{i_3}$, it is not possible that $f_{i_3}(R^4) = s$. But if $f_{i_3}(R^4) = \tilde{s}$, we must have $f_{i_1}(R^4) = s$ so that $i_1$ and $i_3$ form a stable improvement cycle. Hence, we must have

[14] Note that we must have $|I| \geq 4$ if we exogenously break any tie. For the following we assume that no student in $I \setminus \{i_1, i_2, i_3\}$ is interested in $s$ or $\tilde{s}$.

$f_{i_3}(R^4) = i_3$. But then $i_3$ has an incentive to submit $R^2_{i_3}$ when the other students submit $R^2_{i_1}$ and $R^1_{i_2}$!

More generally, we would like to know whether strategy-proofness and constrained efficiency require us to *always* follow the ordering $\succeq^0$ for the case of solvable priority structures. That is, if $f$ is a strategy-proof and constrained efficient mechanism can it be the case that for some problem $R$ we have $\tilde{s} P_i f_i(R)$, $i \succ^0 j$, and $f_j(R) = \tilde{s}$ for some non-specialized school $\tilde{s} \in S^0$? To see that this is possible let $\hat{\mathcal{R}} = \{R : |\{i \in I : A(R_i) \neq \emptyset\}| \leq 2\}$ denote the set of profiles where at most two students $i \in I$ have a non-empty set of acceptable schools $A(R_i)$. Let $\succ' \in ST(\succeq)$ be an arbitrary strict transformation. Now we can modify the rule $f^{ETB}$ as follows: for any profile $R$, (i) if $R \notin \hat{\mathcal{R}}$, then $\hat{f}(R) = f^{ETB}(R)$; and (ii) if $R \in \hat{\mathcal{R}}$, then $\hat{f}(R) = f^{\succ'}(R)$. It is easy to see that $\hat{f}$ is strategy-proof and constrained efficient. An important open question is whether we can allow such *violations* of $\succeq^0$ on more interesting domains of preferences.

### 2.5.2 Full Characterization for General Capacities

Beyond the case of unit capacities we have only derived sufficient conditions for solvability. An important question is whether these conditions can be weakened further. We first illustrate the additional problems for designing strategy-proof and constrained efficient mechanisms when our conditions are not satisfied using a simple example.

There are two specialized schools $s_1, s_2$ and two non-specialized schools $s_3, s_4$. Both specialized schools can admit three students while the two non-specialized schools can only admit one student. There are six students $1, \ldots, 6$ and the priority ordering is given by

$$\begin{array}{lcccccc} \succ_{s_1}: & 1 & 2 & 3 & 4 & 5 & 6 \\ \succ_{s_2}: & 6 & 5 & 4 & 3 & 2 & 1 \end{array}$$

Note that in this example the critical value is $p = 5$ and that the priority structure does not satisfy limited $p$-variability. Now consider the preference profile

| $R$ | $R_1$ | $R_2$ | $R_3$ | $R_4$ | $R_5$ | $R_6$ |
|---|---|---|---|---|---|---|
| | $s_3$ | $s_2$ | $s_4$ | $s_4$ | $s_2$ | $s_3$ |
| | $s_7$ | $s_3$ | | $s_2$ | | $s_2$ |

.

Now suppose we were to use the SDA-ETB with the tie-breaking procedure we defined above for this example assuming that $i \sim^0 j$ for all $i, j \in \{1, \ldots, 6\}$. Then $\mu^1(s_1) = \emptyset$, $\mu^1(s_2) = \{2, 5\}$, $\mu^1(s_3) = \{1, 6\}$, and $\mu^1(s_4) = \{3, 4\}$. Now 6 would be rejected in the first round and 4 would be rejected by $s_4$ in the second round. In the third round, 4 applies to $s_2$ so that 2 would be rejected and applies to $s_3$. Since $s_2$ has filled its capacity and 1 is the lowest priority student at $s_2$, SDA-ETB would break the resulting tie in favor of 2. But then 2 and 6 form a stable improvement cycle. The problem in this example is that there are two non-specialized schools ($s_3$ and $s_4$) that have to reject students. If $\succ^1$ had satisfied limited p-variability, it would have been irrelevant which student is rejected by $s_3$ or $s_4$ since there could not have been a subsequent rejection at some specialized school. Here, in contrast it is important to condition tie-breaking on the priority ranking of school $s_2$ in the first place even though this school has not filled its capacity in round 1 of SDA-ETB. Nevertheless, we do not have a counterexample showing that the above priority structure is not solvable so that the door remains open for further possibility results despite the just mentioned complications.[15]

Secondly, consider the case of identical capacity $q \geq 2$ at all schools so that $p = 3q$ (and the above issue does not arise). Is limited p-variability necessary for solvability here? While we do believe that this is true, potential counterexamples needed to show necessity quickly become intractable. The main problem here is that it is very hard to pin down assignments in case non-specialized schools can admit more than one student. We view the weakening of sufficient conditions for solvability to be substantially more important than extending our impossibility results and have thus not worked towards obtaining a full characterization for the case of general capacities at non-specialized schools.

### 2.5.3 Beyond Non-specialized schools environments

In this paper we have concentrated on the (non-)specialized schools environment. To see that there is room for positive results outside this environment, we now consider an easy example.

There are three schools $s_1, s_2, s_3$ and three students $1, 2, 3$. All schools have a capacity of one and the priority structure is as follows

| $\succ_{s_1}$ | $\succ_{s_2}$ | $\succ_{s_3}$ |
|---|---|---|
| 1 | 2 | 3 |
| $\{2,3\}$ | $\{1,3\}$ | $\{1,2\}$ |

To see that this priority structure is solvable, note that the above environment is isomor-

[15]We do expect the tie-breaking rule to be somewhat more complicated to describe. In general, we conjecture that a critical value of $2q^1 + q^0_{(1)}$ could work in case of identical capacities at specialized schools. This is easily seen to be true when there is just one non-specialized school. However, we have not (yet) been able to prove that this critical value works in the general case. A similar remark applies to the case of asymmetric capacities, where there also seems to be some additional room for solvability.

phic to a *house allocation with existing tenants* problem as introduced by Abdulkadiroglu and Sonmez (1999): Student $i$ is an existing tenant for school $s_i$. Their version of the top-trading cycles algorithm is strategy-proof and constrained efficient for such problems.[16]

In general, it will not be possible to rely on the top-trading cycles algorithm since it is known that it may lead to unstable allocations. Furthermore, the above approach is not applicable, for example, when the set of students with top priority for some school is larger than the school's capacity (as in the (non-)specialized schools environment). However, the above shows that the door in principal remains open for possibility results and a (partial) characterization of solvable priority structures in the general case is an important question for future research.

[16]This is a special case of the hierarchical exchange rules introduced by Papai (2000). She shows that this class of rules exhaust the class of rules that are group strategy-proof, efficient, and satisfy a notion of reallocation-proofness. Recently, Pycia and Unver (2009) have characterized the (slightly) larger class of group strategy-proof and efficient rules.

# Chapter 3

# Market Structure and Matching with Contracts

## 3.1 Introduction

Theoretical models of network formation and matching markets are concerned with predicting which outcomes are likely to emerge when self-interested agents interact. An important strand of this literature belongs to the area of cooperative game theory and "likely outcomes" are not defined by writing down an explicit negotiation protocol, but rather by postulating a set of stability constraints that one perceives to be relevant in the problem under study. In the previous chapters we discussed several examples where such constraints have been an important guideline for the design of real-life mechanisms for *two-sided matching problems* in which a group of *workers* (students) has to be assigned among a set of *firms* (universities/schools). This literature has focused on the *pairwise stability* concept, which only considers the possibility of coordinated deviations by pairs of players. One might worry that a pairwise stable matching could be susceptible to deviations by larger coalitions. However, as long as workers can take at most one job and firms have substitutable preferences, a pairwise stable matching not only exists (Kelso and Crawford (1982)), but is also *group stable* (Roth and Sotomayor (1991)): There is no group of agents who can obtain a strictly preferred matching by forming new partnerships only among themselves, possibly dropping some previously held partnerships. In particular, a pairwise stable matching is *efficient*.

While these are encouraging results for a restricted class of assignment problems, many interesting applications do not fit the assumptions above: Some workers may demand multiple jobs in a labor market,[1] firms may not view workers as substitutes,[2] and markets are often not two-sided.[3]

Recently, Ostrovsky (2008) introduced a model of matching in vertically ordered networks which allows for some of these features. The location of an agent in the (directed) network is exogenously given and agents have preferences over sets of trading relationships, or contracts, with their neighbors. A set of contracts is *chain stable* if (i) no agent would prefer to drop some of her contracts, and (ii) there is no downstream sequence of agents who can obtain a strictly preferred set of contracts by forming new contracts only with their direct neighbors in the sequence, possibly dropping some of their previously held contracts. Ostrovsky shows that chain stable outcomes exist as long as agents' preferences satisfy *same side substitutability* and *cross side complementarity*. However, unlike pairwise stable matchings in the simple matching models above, chain stable allocations may not be group stable. In fact, chain stable outcomes may even be inefficient and thus fail to be in the *core*.[4] Thus, even when deviating agents are not allowed to maintain relationships with outsiders there can be profitable deviations from a chain stable allocation. We characterize the conditions under which these problems cannot occur. Furthermore, we show how the use of the chain stability concept can be justified even though the positive results from two-sided matching markets have no direct extension to the more general supply chain model.

A main contribution of this study is methodological. Unlike most of the literature we will not try to guarantee efficiency and group stability of chain stable outcomes by (directly) restricting the class of allowed preferences. Rather, we take the domain of preferences introduced by Ostrovsky (2008) as given and develop restrictions on the exogenously given network structure. The main structural restriction in the paper is an acyclicity notion which rules out certain kinds of trading cycles. We show that this condition is necessary and sufficient for (i) the equivalence of group and chain stability, (ii) the core stability of chain stable outcomes, (iii) the efficiency of chain stable outcomes, (iv) the existence of a group stable outcome, and (v) the existence of an efficient and *individually stable* outcome.[5] The equivalences provide two justifications for the use

[1]For example, Echenique and Oviedo (2006) mention that 35 % of teachers in Argentina work for more than one school.

[2]Many tasks can only be accomplished by the combined workforce of a set of specialized workers. The construction of a building, for example, requires a structural engineer, a carpenter, and so on, so that complementarities between individual workers are likely.

[3]Brokers act as intermediaries between owners and potential tenants in housing markets, temporary employment agencies supply firms with short-term labor, some stores (e.g. Gamestop) allow customers to trade in used goods which they then sell to other customers, and so on.

[4]An outcome is in the *core* (defined by weak domination), if no group of agents can obtain a weakly preferred outcome for all involved by trading only among themselves.

[5]An outcome is *individually stable*, if no agent refrains from taking her part of the outcome.

of chain stability in the unrestricted model: First, whenever the minimal stability requirement of individual stability can be reconciled with efficiency, chain stable outcomes are also efficient. Thus, imposing the stronger chain stability concept does not lead to any additional efficiency loss. Second, if chain stable networks fail to be group stable, the very existence of a group stable outcome cannot be guaranteed. In this sense chain stable allocations are *as stable as it gets*.

The above acyclicity condition is not sufficient for the equivalence of chain stability and the core, one of the most important solution concepts in cooperative game theory. In the second part of the paper we introduce a stronger acyclicity condition which is then shown to characterize the class of supply chain models for which the equivalence obtains. The characterization subsumes a number of important existing results.

This chapter is organized as follows: After discussing the related literature we motivate the main ideas of this chapter by means of a simple example in section 3.2. In section 3.3 we introduce Ostrovsky's supply chain model. In section 3.4 we present the main results of this chapter in detail. In section 3.5 we present some further results on the core. In section 3.6 we conclude. All proofs, a discussion of the main results, and an extension are relegated to Appendix A.3.

## Related Literature

From a methodological perspective, the paper most closely related to the present study is Abeledo and Isaak (1992). They start from a fixed structure of potential partnerships in a simple one-to-one matching model represented by an undirected graph that contains edges between mutually acceptable pairs of agents. Their main result is that a pairwise stable matching will exist for all preference profiles if and only if the market is two-sided. In this chapter we restrict attention to the model introduced by Ostrovsky (2008) for which the existence of a chain stable allocation is guaranteed. In contrast to Abeledo and Isaak (1992) our focus is the relationship between cooperative solution concepts. Furthermore, their methodology has no direct extension to the model we consider since the acceptability of a trade depends on the whole set of available trades. In particular the set of acceptable allocations cannot be summarized by a simple undirected graph.

More closely related in focus is a line of research that has been concerned with stability concepts for two-sided many-to-many matching markets. If all preferences are substitutable, this model is a special case of the supply chain model so that existence of a pairwise stable allocation follows from Ostrovsky (2008).[6] Blair (1988) was the first to note that in such markets the core can be empty. This implies in particular that (i) group stable allocations may fail to exist, and (ii) the set of pairwise stable allocations may be disjoint from both, the core and the set of group stable allocations. In light of these problems, most studies have focused on stability concepts which limit the set of allowable coalitional deviations. Roth (1984b) considers the restriction that all members of a deviating coalition should obtain a subset of their most preferred set of contracts out of previously held and newly formed contracts. He shows that if all preferences are substitutable, there is no such coalitional deviation from a pairwise stable allocation. In a matching model without contracts, Konishi and Ünver (2006) consider the restriction that deviations have to be pairwise stable themselves. A matching that is not susceptible to such a deviation is called credibly group stable. They show that if one side of the market has responsive preferences and the other side of the market has categorywise responsive preferences,[7] pairwise stability is equivalent to credible group stability. Echenique and Oviedo (2006) consider the restriction that deviations have to be individually stable in the sense that no deviating agent wants to drop some of her partners after the deviation. A matching that is not susceptible to such a deviation is called setwise stable.[8] They show that if one side of the market has substitutable preferences and the other side has strongly substitutable preferences, pairwise and setwise stability are equivalent.[9] There are two important differences between this line of research and our study: First, all these papers study stability notions that restrict the set of allowable coalitional deviations. A problem with this approach is that none of these concepts guarantee efficiency of the outcome. This leaves open the question whether there are other natural stability concepts compatible with efficiency. The results of our paper apply equally well to many-to-many matching models and, to the best of our knowledge, this is the first systematic study of the relationship between pairwise stability, group stability, the core, and efficiency. A second difference is that most of the above papers introduce stronger restrictions on preferences than those needed to guarantee existence of a pairwise stable allocation. Thus, while they provide insightful foundations for pairwise stability in these restricted models, their results do not explain why this is a desirable property when all substitutable preferences are

[6]The existence of a pairwise stable allocation in this model had been established earlier by Roth (1984b).

[7]An agent has categorywise responsive preferences, if there is partition of the set of available partners such that her preferences restricted to any element of this partition are responsive. Konishi and Ünver (2006) show that this is a stronger restriction than substitutability.

[8]Setwise stability was introduced by Sotomayor (1999).

[9]Strong substitutability requires that if an agent $i$ chooses another agent $j$ when the set of available partners is $A \cup \{j\}$ and prefers $A$ to $B$, then $i$ must still choose $j$ when the set of available partners is $B \cup \{j\}$. The results of Echenique and Oviedo (2006) have recently been generalized to matching models with contracts by Klaus and Walzl (2008).

allowed.

Other papers that have studied the importance of the set of allowed potential interactions for cooperative solution concepts include Papai (2007), who analyzes how to restrict allowable trades in a general indivisible goods exchange market in order to guarantee a singleton core (see also Papai (2004) for a similar analysis in the context of coalition formation games), and Kalai, Postlewaite, and Roberts (1978), who compare core outcomes between an unrestricted market game and a game in which some players are not allowed to form coalitions.

Finally, our paper is also related to the literature on game theoretic network formation models initiated by Jackson and Wolinsky (1996). It is known that in these models stability is often incompatible with efficiency. By characterizing the class of supply chain models for which efficiency and even a minimal notion of stability are compatible, we show precisely when the compatibility results from the two sided matching literature break down and the more negative results from the network formation literature obtain.

## 3.2 An Example

An actor, $M$, is in negotiations about starring in a movie that studio $S$ has in development. In addition, $M$ has just finished work on her self-financed pet project and is seeking a distributor for it. There are two potential distributors, $D_1$ and $D_2$. The set of available trading relationships is fixed and exogenously given by $X = \{x(M, S), x(M, D1), x(M, D2), x(S, D1)\}$. Here, $x(i, j)$ represents a trading relationship in which agent $i$ sells something to agent $j$. Note that we assume there is at most one possible trading relationship between each pair of agents and that $D2$ cannot distribute $S$' project. Available trading relationships can be summarized by a directed graph that contains an edge from $i$ to $j$ if and only if $x(i, j) \in X$.

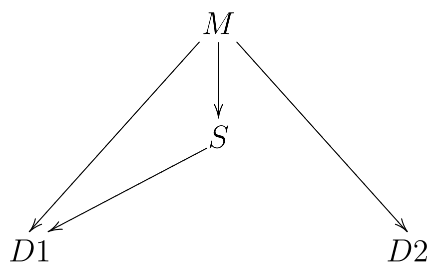


**Figure 1:** Graph $G_1$ of potential interactions

An allocation, or *network*, is a set of contracts. A network is *chain stable*, if (i) no agent wants to drop some of his or her contracts, and (ii) there is no downstream sequence of agents who, by signing new contracts only with immediate neighbors in the sequence while possibly dropping some previously held contracts, can obtain a strictly preferred outcome. Consider the following exemplary profile of preferences:

| $R$ | $R_M$ | $R_S$ | $R_{D1}$ | $R_{D2}$ |
|---|---|---|---|---|
| | $\{x(M,D1)\}$ | $\{x(M,S),x(S,D1)\}$ | $\{x(S,D_1)\}$ | $\{x(M,D2)\}$ |
| | $\{x(M,D1),x(M,S)\}$ | | | |
| | $\{x(M,D2)\}$ | | | |

This notation has the following interpretation: $\{x(M,D1)\}$ is the most preferred (set of) contract(s) for $M$, $\{x(M,D1),x(M,S)\}$ is the second most preferred set of contracts, $\{x(M,D2)\}$ is the third most preferred set of contracts, and no other set of contracts is acceptable to $M$. The other formulas have analogous interpretations. It is easy to verify that $\{x(M,D2)\}$ is the unique chain stable allocation. While this allocation is also efficient, it is vulnerable to a joint deviation by $M,S$, and $D1$ since they all strictly prefer the network $\{x(M,S),x(S,D1),x(M,D1)\}$. This deviation is not considered by chain stability since it would require $M$ to sign contracts with both $S$ and $D1$. What are the conditions under which chain stable networks are not susceptible to such a deviation?

Suppose that each agent is characterized by her location in the network of figure 1 and a *capacity vector* which contains an exogenously given upper bound on the number of contracts she can sign. In the example above, $D1$'s capacity would then have to be at least two since she would, in principle, be willing to sign the two contracts $x(S,D1)$ and $x(M,D1)$ (as $\{x(S,D1),x(M,D1)\}P_{D1}\emptyset$). Suppose now that $D1$ cannot distribute both projects. This obviously rules out the above example, but what about an arbitrary profile of preferences? We now show that if each agent views two contracts in which she is the seller (buyer) as substitutes, there cannot be a profitable deviation from a chain stable network.

Suppose to the contrary that for some substitutable preference profile agents $S$, $M$, $D1$ strictly prefer some network $\mu'$ over a chain stable network $\mu$, and can obtain $\mu'$ by signing new contracts only among themselves. Then either $x(M,D1)\in\mu'\setminus\mu$ or $x(S,D1)\in\mu'\setminus\mu$ since $D1$ can sign at most one contract. Suppose the former case applies (the latter case is analogous). Since $M$ strictly prefers $\mu'$ over $\mu$ her most preferred subset out of all her contracts in $\mu$ and $\mu'$ must include at least one contract in $\mu'\setminus\mu$ by revealed preference. If $x(M,D1)$ is in $M$'s most preferred subset, $M$ and $D1$ block $\mu$ using $x(M,D1)$ in the sense of chain stability since (i) $M$ would still want to sign $x(M,D1)$ even when she can only choose from contracts in $\mu\cup\{x(M,D1)\}$ by substitutability, and (ii) $D1$ must prefer $x(M,D1)$ over her contract (if any) in $\mu$ given that she can sign at most one contract. If $x(M, D1)$ is not in $M$'s most preferred subset of $\mu \cup \mu'$, we must have $x(M, S) \in \mu' \setminus \mu$. Since $D1$ can sign at most one contract, $x(M, S)$ must be $S$' only contract in $\mu'$. But then $M$ and $S$ block $\mu$ in the sense of chain stability and we again obtain a contradiction.

Since the above argument is valid for any potential block of a chain stable allocation, we see that if $D1$ can sign at most one contract, chain stable allocations have to be immune to any coalitional deviation. In the next section we show that this is a special case of a much more general result.

## 3.3 The Supply Chain Model

This section briefly describes Ostrovsky (2008)'s supply chain model. The presentation here is self contained, but readers interested in a more detailed introduction may want to consult the original article.

Consider a market consisting of a finite set of agents $V$. Agents trade discrete units of indivisible goods and trading relationships are represented by bilateral *contracts*. Each contract is of the form $(s, b, a, p)$ and represents a sale of one (unit of a) *good* $a \in \mathbb{N}$ from seller $s \in V$ to buyer $b \in V$ at a *price* $p \in \mathbb{R}$.[10] The set of all possible contracts, denoted by $X$, is assumed to be exogenously given and finite. For $x \in X$, let $s_x$ denote the seller in contract $x$ and let $b_x$ denote the buyer in contract $x$. It is assumed that there are no *directed trading cycles in* $X$, that is, there is no sequence of agents $v_1, \ldots, v_n$ such that, for all $i \in \{1, \ldots, n\}$, there exists a contract $x_i$ such that $s_{x_i} = v_i$ and $b_{x_i} = v_{i+1}$ (where $n + 1 := 1$).[11] A *supply chain model* is given by the pair $(V, X)$, with the assumption that there are no directed trading cycles in $X$. For future reference we now introduce some more terminology and notation: A contract in which agent $v$ is the seller is called a *downstream contract* for $v$. A contract in which agent $v$ is the buyer is called an *upstream contract* for $v$. Given a set of contracts $Y \subseteq X$, let $D_v(Y)$ denote the set of contracts (in $Y$) in which $v$ is a seller, $U_v(Y)$ denote the set of contracts in which $v$ is a buyer, and $Y(v)$ denote the set of all contracts involving $v$. An agent $v \in V$ with $U_v(X) \neq \emptyset$ and $D_v(X) \neq \emptyset$ is an *intermediary*. For each pair $v, w \in V$, $X(v, w)$ denotes the set of all possible contracts between $v$ and $w$.

[10] This formulation of contracts follows Ostrovsky (2008) and is chosen for concreteness. Our results do not depend on the exact nature of the set of contracts apart from the assumption that each contract is bilateral. For example, a labor market contract could specify wage, days of leave, retirement plans, and so on.

[11] This assumption corresponds to Ostrovsky (2008)'s assumption that agents are located in a vertically ordered network.

### 3.3.1 Preferences

In a supply chain model only agents' preferences over sets of contracts are allowed to vary. Preferences are subject to four restrictions.

(i) An agent cares only about the contracts that she is involved in as a buyer or as a seller (**No Direct Externalities**)

(ii) No agent is ever indifferent between two distinct sets of contracts (**Strict Preferences**)

The first two assumptions imply that the preferences of an agent $v \in V$ can be described by a linear order $R_v$ on the set of all subsets of contracts involving $v$, $2^{X(v)}$. For two sets of contracts $Y, Z \subseteq X$, we denote by $Y R_v Z$ that $v$ weakly prefers $Y$ over $Z$ and by $Y P_v Z$ that $v$ strictly prefers $Y$ over $Z$, that is, $Y R_v Z$ and $Y \neq Z$. Given a set of contracts $Y \subseteq X$ and a strict preference relation $R_v$, $Ch_v(Y)$ denotes $v$'s most preferred subset of $Y$, that is, $Ch_v(Y) P_v Z$ for any $Z \subseteq Y$ with $Z \neq Ch_v(Y)$. A nonempty set of contracts $Y$ is *acceptable* (according to $R_v$) if $Y P_v \emptyset$. The next two restrictions concern the choices of agents from various sets of contracts.

(iii) Whenever a downstream contract becomes unavailable, $v$ does not reduce her demand for any other downstream contract.[12] More formally, let $Y \subseteq X$ and $x, x' \in D_v(Y)$. Then $x \in Ch_v(Y)$ implies that also $x \in Ch_v(Y \setminus \{x'\})$ (**Same Side Substitutability**)

(iv) If an additional upstream contract becomes available to an agent $v \in V$, she does not reduce her demand for any downstream contract.[13] More formally, let $Y \subseteq X$, $x \in D_v(Y)$, and $x' \in U_v(X) \setminus Y$ be arbitrary. Then $x \in Ch_v(Y)$ implies that also $x \in Ch_v(Y \cup \{x'\})$ (**Cross Side Complementarity**)

Let $\mathcal{R}$ denote the set of all preference profiles satisfying assumptions (i) to (iv). Note that in a supply chain model without intermediaries CSC is vacuously satisfied so that such a model reduces to a many-to-many two-sided matching model with *substitutable* preferences as studied in e.g. Roth (1984b).

### 3.3.2 Networks and Solution Concepts

Given a preference profile in the domain introduced above, the aim of a supply chain model is to predict which contracts will be signed by the agents. In the supply chain model the relevant

[12] An analogous condition is required to hold for upstream contracts.

[13] An analogous condition is required to hold for upstream contracts in case a new downstream contract becomes available.

outcomes are sets of contracts, or *networks*.[14] Networks will usually be denoted by $\mu$ and agent $v$'s set of contracts under $\mu$ will be denoted by $\mu(v)$. Predictions take the form of (cooperative) solution concepts which require a network to be robust against certain deviations of individuals or groups.

A network is *individually rational* if no agent is assigned an unacceptable set of contracts. This assumes that if an individual wanted to deviate she has to discontinue all of her existing relationships. A network $\mu$ is *individually stable* if no agent $v$ wants to drop some of her contracts in $\mu(v)$, that is, $Ch_v(\mu(v)) = \mu(v)$ for all $v \in V$. In contrast to individual rationality, individual stability thus allows an individual to delete some but also to keep other contracts. Next, we consider stability notions that rule out coordinated deviations by groups of agents.

The *core* (defined by weak domination) considers deviations by arbitrary groups of players, but deviating agents are not allowed to maintain existing relationships with "outsiders". More formally, a network $\mu'$ *weakly dominates network* $\mu$ *via coalition* $A$ if (i) no member of $A$ trades with an outsider under $\mu'$, that is, $x \in \mu'$ implies that either $\{s_x, b_x\} \subseteq A$ or $\{s_x, b_x\} \cap A = \emptyset$, (ii) $\mu'(a) R_a \mu(a)$ for all $a \in A$, and (iii) $\mu'(a) P_a \mu(a)$ for at least one agent $a \in A$. A network $\mu$ is in the *core (defined by weak domination)* (or *core stable*), if it is not weakly dominated by any other network.[15]

*Group stability*, on the other hand, considers any deviation that a coalition can implement by forming new contracts only among themselves while possibly dropping some previously held contracts. Coalition $A$ can *obtain* $\mu'$ *from* $\mu$ if (i) additional contracts are formed only between members of $A$, that is, $x \in \mu' \setminus \mu$ implies that $\{s_x, b_x\} \subseteq A$, and (ii) only agents in $A$ drop some of their contracts, that is, $x \in \mu \setminus \mu'$ implies that $\{s_x, b_x\} \cap A \neq \emptyset$. Network $\mu$ is *blocked by coalition* $A$ via network $\mu'$ if (i) $A$ can obtain $\mu$ from $\mu'$, and (ii) $\mu'(a) P_a \mu(a)$ for all $a \in A$. A network is *group stable* if it is not blocked by any coalition. Note that group stability is a stronger solution concept than the core.

A major problem is that group and core stable networks can fail to exist even under quite restrictive assumptions about preferences.[16] Thus, in order to guarantee existence the set of coalitional deviations has to be restricted. Ostrovsky (2008) introduces a new stability criterion which generalizes the idea of pairwise stability in the sense that it considers (some)

[14] Here, networks take the role of matchings from the other two chapters of this thesis. The above formulation facilitates the analysis in this chapter.

[15] Analogously, one can define the core by strong domination by requiring that all members of the coalition $A$ have to be strictly better off. Roth and Sotomayor (1991) (Chapter 5) show that even in many-to-one matching markets with responsive preferences, this core concept allows for matchings that are not pairwise stable.

[16] See e.g. Klaus and Walzl (2008) and Konishi and Ünver (2006), who show that even when preferences are strongly substitutable and responsive, respectively, there may not exist a group or even a core stable network/matching.

coordinated deviations by downstream sequences of agents instead of only deviations by pairs. A *chain* is a downstream sequence of contracts $\{x_1, \ldots, x_n\} \subseteq X$ such that for all $i < n$, $b_{x_i} = s_{x_{i+1}}$. Network $\mu$ is *blocked by the chain* $x_1, \ldots, x_n \notin \mu$ if (i) $x_1 \in Ch_{s_{x_1}}(\mu(s_{x_1}) \cup \{x_1\})$, (ii) $\{x_i, x_{i+1}\} \subseteq Ch_{s_{x_{i+1}}}(\mu(s_{x_{i+1}}) \cup \{x_i, x_{i+1}\})$ for all $i < n$, and (iii) $x_n \in Ch_{b_{x_n}}(\mu(b_{x_n}) \cup \{x_n\})$. A network $\mu$ is *chain stable*, if it is individually stable and if it is not blocked by any chain. One of the main results in Ostrovsky (2008) is that chain stable networks exist for all profiles in the domain $\mathcal{R}$ introduced in 3.3.1. Note that if there are no intermediaries chain stability reduces to *pairwise stability* so that his result generalizes the existence results from two-sided matching models with substitutable preferences.

To justify the use of chain stability, Ostrovsky (2008) argues that it is reasonable to assume that a downstream sequence of agents is able to coordinate a deviation since *a customer need only pick up the phone and call a potential supplier* [...]; *the potential supplier* [...] *calls one of his potential suppliers, and so on.* (p. 911). On the other hand he argues that *coalitions involving several competing firms require much more coordination and information exchange between the agents* (p. 911). There are two caveats to this justification: First, as we saw in the example of section 3.2, chain stability does not generally rule out any profitable manipulation by downstream sequences of agents. So it is not necessarily true that any deviation from a chain stable network has to involve competing agents or firms. Second, requiring robustness against any possible chain block in complex supply chain models means that coordinated deviations by large groups of agents are thought to be possible. Given that, as we saw in the example, coalitions consisting of as few as three agents might have a profitable joint deviation from a chain stable network it is not clear why robustness against chain blocks is important while robustness against the latter type of deviation is not. Finally, requiring robustness against chain blocks may come at the expense of efficiency. If some of the deviations considered by chain stability are implausible due to e.g. the size of the coalitions involved, less demanding stability concepts may reduce the efficiency loss while still being satisfactorily robust. To summarize, Ostrovsky (2008)'s results do not provide a firm foundation for the concept of chain stability.

In many-to-one two-sided matching markets with substitutable preferences pairwise stable matchings are always efficient and both, core and group stability, reduce to pairwise stability (see Roth and Sotomayor (1991)). These results can be seen as a justification for the use of pairwise stability as the leading concept of stability in such models. The purpose of this paper is to (a) characterize the largest class of supply chain models for which chain stable networks are efficient and group or core stable, and (b) to derive a foundation for the use of chain stability in the unrestricted model.

## 3.4 Main Results

This section develops conditions under which chain stable networks are efficient as well as core and group stable. Instead of introducing further restrictions on preferences, we restrict potential interactions between agents. These restrictions concern *who can contract with whom* and *how many relationships an agent can form.* Such assumptions are common in two-sided matching models, where an agent cannot contract with another agent on her side of the market and it is often assumed that agents on one side of the market can all engage in at most one relationship. The restrictions that we develop can be interpreted as restricting the sets of acceptable contracts. However, unlike the usual preference restrictions in the matching literature, such as responsiveness or strong substitutability, our conditions do not place further restrictions on the ranking of acceptable (or unacceptable) sets of contracts.

For our analysis it is useful to work with a graphical representation of the supply chain model. Let $G_X$ be the simple directed graph that contains an edge from agent $v$ to agent $w$ if and only if $X$ contains some contract $x$ with $s_x = v$ and $b_x = w$. This is the *graph of potential interactions* that describes who can contract with whom. Note that even though there may be more than one possible contract between a pair of agents, $G_X$ contains at most one edge between each pair of agents. Given that the set of available contracts is assumed to be fixed, agents are not allowed to choose their location in the network. The graph is not necessarily complete so that some pairs of agents may not be able to trade with each other at all. For example, two agents may not know each other, or a trade embargo forbids them to engage in a contractual relationship. If there are no directed trading cycles in $X$, $G_X$ contains no *directed cycles*,[17] and vice versa.

We restrict attention to supply chain models in which agents face fixed upper bounds on the number of contracts they can sign. First of all, for each pair of agents $v, w \in V$ with $(v, w) \in G_X$, there is an integer $c(v, w) \in \{1, \ldots, |X(v, w)|\}$ representing the maximum number of contracts $v$ and $w$ can sign with each other. If $X(v, w) = \emptyset$, so that $v$ and $w$ cannot directly trade with each other, we set $c(v, w) = 0$. One interpretation is that $c(v, w)$ represents the capacity of the unique distribution channel between $v$ and $w$. Furthermore, each agent $v \in V$ has an upper bound on the number of agents she can trade with: $v$ can sign contracts with at most $q_v^D$ downstream agents and at most $q_v^U$ upstream agents. For example, an agent who owns $k$ indivisible goods and whose only interest is to sell these goods can sign contracts with at most $k$ agents. A set of contracts $Y \subseteq X$ *violates v's capacity constraints*, if either $|\{w \in V \setminus \{v\} :$

[17]A directed cycle in $G_X$ is a sequence of agents $v_1, \ldots, v_n$ such that, for all $i \in \{1, \ldots, n\}$, $(v_i, v_{i+1}) \in G_X$ (where we set $n + 1 := 1$).

$s_x = w$ for some $x \in Y(v)\}| > q_v^U$, $|\{w \in V \setminus \{v\} : b_x = w$ for some $x \in Y(v)\}| > q_v^D$, or if there is an agent $w \in V \setminus \{v\}$ such that $|Y \cap X(v,w)| > c(v,w)$. We assume that preferences *conform to capacities*: No agent finds a set of contracts acceptable that violates (one of) her capacity constraints.[18] In the following, $\mathcal{R}_{(c,q)} \subseteq \mathcal{R}$ denotes the set of all preference profiles conforming to capacities and satisfying the assumptions of the supply chain model. Given $V$ and $X$, the *unrestricted model* is obtained by setting $c(v,w) = q_v^U = q_v^D = |X|$, for all $v, w \in V$.

The triple $(G_X, c, q)$ is the *market structure* induced by the supply chain model $(V, X)$ and the capacity constraints $(c, q)$. The market structure is taken to be fixed throughout so that, in particular, we do not consider location or capacity choice.

The example in Section 3.2 suggests that the presence or absence of certain cycles in the market structure is key to the relationship between the solution concepts studied in this paper. Note that the assumptions made so far ruled out only directed, but not *undirected cycles* in $G_X$.[19] The market structure in the example contained exactly one undirected cycle: $M$, $S$, $D1$. Such cycles by themselves cannot be the problem: Even a marriage market permits cycles (of even length) and yet most stability concepts collapse to pairwise stability. However, a crucial difference to the unrestricted supply chain model is that in a marriage market cycles cannot actually be realized since each agent can have at most one partner. The following definition introduces the notion of capacity constraints on cycles.

**Definition 7.** *Let $v_1, \ldots, v_n$ be an undirected cycle in $G_X$. Agent $v_i$ is* ***capacity constrained on cycle*** *$v_1, \ldots, v_n$ if*

*(i) $v_i$ is a* ***source****, that is, $\{(v_i, v_{i+1}), (v_i, v_{i-1})\} \subset G_X$, and $q_{v_i}^D \leq 1$,*

*(ii) $v_i$ is a* ***sink****, that is, $\{(v_{i-1}, v_i), (v_{i+1}, v_i)\} \subset G_X$, and $q_{v_i}^U \leq 1$, and*

*(iii) $v_i$ is a* ***passing node****, that is, either $\{(v_{i-1}, v_i), (v_i, v_{i+1})\} \subset G_X$ or $\{(v_{i+1}, v_i), (v_i, v_{i-1})\} \subset G_X$, and $\min\{q_{v_i}^D, q_{v_i}^U\} = 0$.*

*A* ***2 cycle*** *in $(G_X, c, q)$ is an undirected cycle such that no agent is capacity constrained on it.*

If $v$ is a capacity constrained passing node, either all incoming or all outgoing edges in $G_X$ relative to $v$ are irrelevant. For example, $v$ could be a firm that has signed an exclusive long-term contract with a supplier. The example of section 2 contained a unique cycle in which $M$ was a source, $S$ was a passing node, and $D1$ was a sink. According to the first preference profile $(R_M, R_S, R_{D1}, R_{D2})$ we considered in this example, no agent was capacity constrained

[18] Instead of requiring preferences to conform to the capacity vector, one could place feasibility restrictions on the set of networks. We discuss this approach in Appendix A.3.

[19] An undirected cycle in $G_X$ is a sequence of distinct agents $v_1, \ldots, v_n$ such that, for all $i \in \{1, \ldots, n\}$, either $(v_i, v_{i+1}) \in G_X$ or $(v_{i+1}, v_i) \in G_X$ (where $n + 1 := 1$).

so that $M$, $S$, $D_1$ was a 2 cycle. Note that by the definition of $G_X$ an undirected cycle has to contain at least three agents. Thus, a supply chain model with only two agents can never have a 2 cycle. The following example shows that even when there are only two agents the set of chain stable networks, the core, and the set of group stable networks can all be disjoint.

**Example 8.** *Consider a supply chain model with two agents $v$ and $w$ in which $v$ possesses two indivisible goods $A$ and $B$. The only available contracts are $x_A$ and $x_B$, where $x_i$ represents the sale of good $i \in \{A, B\}$ to agent $w$ at some fixed price. Selling both goods is efficient but agents differ in their evaluation of the most preferred outcome: $\{x_A\}P_w\{x_A, x_B\}P_w\emptyset P_w\{x_B\}$ and $\{x_B\}P_v\{x_A, x_B\}P_v\emptyset P_v\{x_A\}$. It is easy to see that this profile of preferences satisfies SSS and CSC. However, (i) the unique chain stable network is the empty network, (ii) there is no group stable network, and (iii) the unique core network is $\{x_A, x_B\}$.*

*Now suppose that a sale of both objects has to be implemented by a single contract $x_{A,B}$ instead of the two independent contracts $x_A$ and $x_B$. In this case, $\{x_{A,B}\}$ is the unique chain stable network. This network is also the unique core and group stable network.*

In order to guarantee that chain stable networks are efficient as well as core and group stable we thus need to assume that each pair of agents signs at most one contract with each other. As the example shows this does not mean that an agent can sell at most one good to any given neighbor. Rather, it requires that all trading relationships between a given pair of agents can be *bundled* into one contract. We now introduce the concept of acyclic market structures that will play a crucial role in our analysis.

**Definition 8.** *The market structure $(G_X, c, q)$ is* ***weakly acyclic*** *if it contains no 2 cycles and $c(v, w) \leq 1$ for all $v, w \in V$.*

Some further notation facilitates the statement of our results: Given a preference profile $R \in \mathcal{R}_{(c,q)}$, let $\mathcal{CS}(R)$ denote the set of all chain stable networks, $\mathcal{IS}(R)$ denote the set of all individually stable networks, $\mathcal{GS}(R)$ denote the set of all group stable networks, $\mathcal{C}(R)$ denote the core, and $\mathcal{E}(R)$ denote the set of all efficient networks. With these preparations the first main result reads as follows.

**Theorem 4.** *The following are equivalent:*

*(i) The market structure $(G_X, c, q)$ is weakly acyclic.*

*(ii) Chain stable networks are always group stable, that is, $\mathcal{CS}(R) = \mathcal{GS}(R)$ for all $R \in \mathcal{R}_{(c,q)}$.*

*(iii) Chain stable networks are always in the core, that is, $\mathcal{CS}(R) \subseteq \mathcal{C}(R)$ for all $R \in \mathcal{R}_{(c,q)}$.*

*(iv) Chain stable networks are always efficient, that is, $\mathcal{CS}(R) \subseteq \mathcal{E}(R)$ for all $R \in \mathcal{R}_{(c,q)}$.*

Theorem 4 characterizes the class of supply chain models for which chain stability possesses properties that are analogous to those of pairwise stability in two-sided markets. The most difficult part of the proof, which can be found in Appendix A.3, is to show that (i) implies (ii). The intuition here is that due to SSS and CSC any deviation which cannot be implemented as a sequence of blocking chains must contain a cyclical sequence of trades that makes all agents involved strictly better off compared to some baseline allocation. But if the market structure is weakly acyclic, a deviation that contains a cyclical sequence of trades can improve the welfare of all agents involved only if the baseline allocation was unacceptable to some of the agents. Since a chain stable network is individually stable, and thus in particular individually rational, it cannot be blocked by any coalition given weak acyclicity. Note that the result only says that for weakly acyclic supply chain models chain stable networks are always in the core. This leaves open the question whether the chain stable set can be a strict subset of the core for weakly acyclic models and we will return to this question in the next section. Before proceeding, we now consider an application.

**Application 1** (A market with a central intermediary)**.** *Consider a market consisting of a set $S$ of suppliers, a set $C$ of consumers, and one central intermediary $I$. We make the following assumptions:*

- *Suppliers can either sell directly to consumers or through the intermediary*
- *Suppliers and the intermediary can both sign contracts with an arbitrary number of agents*
- *Each supplier can sign at most one contract with the intermediary*
- *Each consumer can sign at most one contract*

*Figure 2 summarizes potential interactions for the case of $S = \{S_1, S_2\}$ and $C = \{C_1, C_2\}$ by the directed graph $G_2$.*

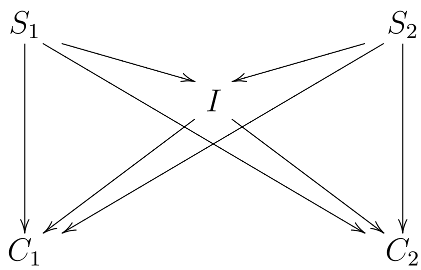


***Figure 2:*** *Graph $G_2$ of potential interactions.*

*The underlying market structure is weakly acyclic since any undirected cycle must contain at least one consumer, who, by assumption, signs at most one contract. If every agents' preferences satisfy SSS and CSC chain stable networks not only exist but are also efficient and group stable by Theorem 4. If there is more than one intermediary, chain stable networks may be neither group stable nor efficient. The reason is that there could be cycles consisting exclusively of suppliers and intermediaries. Without further restrictions on capacities or the pattern of connections, the market structure would fail to be weakly acyclic.*

Theorem 4 identifies weak acyclicity as a necessary and sufficient condition to guarantee efficiency and group stability of chain stable networks. This condition is quite restrictive and we now derive a foundation for chain stability in the unrestricted model. The following result is key to this foundation since it relates weak acyclicity to the existence of group stable networks and to the existence of efficient individually stable networks.

**Theorem 5.** *The following are equivalent:*

*(i) The market structure $(G_X, c, q)$ is weakly acyclic.*

*(ii) A group stable network always exists, that is, $\mathcal{GS}(R) \neq \emptyset$ for all $P \in \mathcal{R}_{(c,q)}$.*

*(iii) An efficient and individually stable network always exists, that is, $\mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$ for all $R \in \mathcal{R}_{(c,q)}$.*

Theorem 5 is related to the literature on network formation models since the supply chain model is a special case of the general network formation models studied in e.g. Jackson and Wolinsky (1996). For these models the incompatibility between efficiency and stability is well known. On the other hand, the supply chain model contains most of the matching models previously studied in the literature as a special case. As mentioned above efficiency and stability are compatible in these models. Theorem 5 thus identifies a point at which the positive results from the two sided matching literature break down and the general incompatibility results from the network formation literature obtain since even the minimal requirement of individual stability cannot in general be reconciled with efficiency.[20] The following is an immediate corollary of Theorems 4 and 5.

**Corollary 1.** *An efficient and individually stable network always exists if and only if chain stable networks are always efficient.*

[20] Individual stability is often viewed as a minimal stability requirement in the matching literature. For example, except for the core all stability concepts considered by Echenique and Oviedo (2006), Konishi and Ünver (2006), and Klaus and Walzl (2008) require individual stability.

This can be seen as a justification for chain stability from an efficiency perspective: The only reason for a chain stable network to fail the efficiency criterion is that even the minimal requirement of individual stability cannot, in general, be reconciled with efficiency. This implies that there is no additional efficiency loss from imposing the stronger chain stability concept. The following is another immediate corollary of Theorems 4 and 5.

**Corollary 2.** *A group stable network always exists if and only if chain stable networks are always group stable.*

This can be seen as a justification for chain stability from a robustness perspective: The only reason for a chain stable network to fail the group stability criterion is that the existence of a group stable network cannot, in general, be guaranteed. In this sense a chain stable network is *as stable to coordinated deviations as it gets.* The following corollary summarizes all equivalences derived in this section.

**Corollary 3.** *The following are equivalent:*

*(i) The market structure $(G_X, c, q)$ is weakly acyclic.*

*(ii) $\mathcal{CS}(R) = \mathcal{GS}(R)$ for all $R \in \mathcal{R}_{(c,q)}$*

*(iii) $\mathcal{CS}(R) \subseteq \mathcal{C}(R)$ for all $R \in \mathcal{R}_{(c,q)}$*

*(iv) $\mathcal{CS}(R) \subseteq \mathcal{E}(R)$ for all $R \in \mathcal{R}_{(c,q)}$*

*(v) $\mathcal{GS}(R) \neq \emptyset$ for all $R \in \mathcal{R}_{(c,q)}$*

*(vi) $\mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$ for all $R \in \mathcal{R}_{(c,q)}$.*

It is important to bear in mind that the above results are about solution concepts. As we show in Appendix A.3 the non-trivial of the above implications do not hold without the quantifiers. For example, it is not true that if for some profile $R \in \mathcal{R}_{(c,q)}$ an efficient and individually stable network exists, then all chain stable networks are efficient. In the appendix we construct an example in which all chain stable networks are strongly inefficient even though efficiency and individual stability are compatible. This may lead some readers to question the efficiency justification for chain stability given above. After all, we might achieve a better compromise between efficiency and stability considerations if we settle for individual stability whenever it is compatible with efficiency but chain stability is not, and otherwise require chain stability. Apart from the question whether such a concept is descriptively appealing, such a solution concept is (computationally) infeasible as one would need to check that given a particular problem, (a) there is an individually stable and efficient network, and (b) any chain stable network is inefficient. Ostrovsky (2008) provides a reasonably fast algorithm to compute chain stable allocations so that chain stability is immune to this type of criticism. A similar remark applies for the robustness justification of chain stability.

## 3.5 Further results on the Core

In this section we direct attention to the core, which has been one of the most important solution concepts in cooperative game theory.[21] The last section, and in particular Corollaries 1 and 2, showed how the chain stability concept can be motivated on basis of efficiency and stability considerations. The main questions of this section are:

(a) Is there a similarly strong justification for the core?

(b) When do core and chain stability coincide?

First, we focus on the relationship between individual stability and the core. It is easy to see that the core does not, in general, satisfy individual stability: The unique core allocation in the example of section 2 is $\{x(M,S), x(S,D1), x(M,D1)\}$. This is not individually stable since $M$ and $D1$ both want to drop one of their contracts.[22] The plausibility of the core allocation thus rests on the assumptions that all contracts are signed simultaneously and that renegotiation is impossible. If the agents would try to implement the allocation sequentially, one of the deviators will defect. The next result shows that the same conditions which guaranteed that chain stable networks are always in the core characterize the class of supply chain models for which core networks are always individually stable.

**Theorem 6.** *The following are equivalent:*

*(i) The market structure $(G_X, c, q)$ is weakly acyclic.*

*(ii) Core allocations are always individually stable, that is, $\mathcal{C}(R) \subseteq \mathcal{IS}(R)$ for all $R \in \mathcal{R}_{(c,q)}$.*

An immediate corollary of Theorems 5 and 6 is that the core sacrifices individual stability only if efficiency cannot in general be reconciled with individual stability. The following variation of the example in Section 3.2 shows that chain stability has an edge over the core in weakly acyclic supply chain models.

**Example 9.** *Suppose that capacities are $q^U_M = 0$, $q^D_M = 2$, $q^U_S = q^D_S = 1$, $q^U_{D1} = 1$, and $q^D_{D1} = 0$. Consider the following preference profile ($D2$'s preferences and capacities are not specified since they are irrelevant for the example)*

[21] A good, albeit technical, introduction to the cooperative game theory is Peleg and Sudhölter (2003).

[22] That core allocations can fail to be individually stable in many-to-many matching markets has previously been shown by Echenique and Oviedo (2006) and Konishi and Ünver (2006).

| $R$ | $R_M$ | $R_S$ | $R_{D1}$ |
|---|---|---|---|
| | $\{x(M,S), x(M,D1)\}$ | $\{x(M,S), x(S,D1)\}$ | $\{x(M,D_1)\}$ |
| | $\{x(M,S)\}$ | | $\{x(S,D1)\}$ |
| | $\{x(M,D1)\}$ | | |

*It is easy to check that the market structure is weakly acyclic and that preferences belong to the domain of preferences we consider. Note that the network $\{x(M,S), x(S,D1)\}$ is in the core. While this allocation is individually stable, it is not chain (or group) stable since (i) $M$ would prefer to sell her project to $D1$ in addition to signing $x(M,S)$, and (ii) $D1$ strictly prefers to sign $x(M,D1)$ instead of $x(S,D1)$. This deviation is not considered by the core since it would make $S$ worse off and $M$ can only be made better off if she is allowed to maintain her relationship with $S$. In particular, chain stability is a stronger requirement than the core.*

We now turn to the question of when core and chain stability are equivalent. By Theorem 4, we can concentrate our attention to the class of weakly acyclic supply chain models. Given the equivalence of chain and group stability for this class of supply chain models, the question can be rephrased as: When is it irrelevant whether a stability concept allows deviating agents to maintain relationships with outsiders or not? We now introduce a stronger structural condition which will later be seen to guarantee the equivalence of chain and core stability. The main idea is that it does not only matter that some agent on a cycle is capacity constrained but exactly which and how many agents are capacity constrained. More formally, we have the following.

**Definition 9.** *A **restricted 2 cycle** of $(G_X, c, q)$ is an undirected cycle such that either (i) no agent is capacity constrained, or (ii) there is exactly one capacity constrained agent and this agent is a source or a sink of the cycle. A market structure $(G_X, c, q)$ is **strongly acyclic** if it has no restricted 2-cycles and $c(v,w) = 1$, for all $v, w \in V$.*

The market structure in Example 9 would be strongly acyclic if either $M$ and $D1$ are both capacity constrained or if $S$, among potentially other agents, is capacity constrained. Note that strong implies weak acyclicity. One example of a weakly but not strongly acyclic market structure is Application 1: The market structure contains cycles consisting of one supplier, the intermediary, and one consumer. Of these three agents, only the consumer was assumed to be capacity constrained. We have the following.

**Theorem 7.** *The following are equivalent:*

*(i) The market structure $(G_X, c, q)$ is strongly acyclic.*

*(ii) Chain stability is equivalent to core stability, that is, $\mathcal{CS}(R) = \mathcal{C}(R)$ for all $R \in \mathcal{R}_{(c,q)}$.*

The intuition for this result is that in a weakly but not strongly acyclic market structure there could be an undirected cycle $v_1, \ldots, v_n$ such that (i) agents $v_1, \ldots, v_j$ can reach some core allocation $\mu$ by signing contracts only with their direct neighbor(s) in $\{v_1, \ldots, v_j\}$, and (ii) there is a blocking chain of $\mu$ involving $v_{j+1}, \ldots, v_n$ as well as $v_1$ and $v_j$. This is possible, if e.g. $v_1$ is a source and is the only capacity constrained agent on the cycle. The above is impossible if there are no restricted 2 cycles: If $v_1$ was the only capacity constrained agent on the cycle she would have to be a passing node. But then $v_1$ would never agree to sign a contract with $v_n$ if $\mu$ was individually rational. The following is an application of Theorems 4 and 7.

**Application 2** (Many-to-Many (One) Matching)**.** *Consider a labor market with at least two firms and at least two workers. All firms can hire arbitrarily many workers, while some or all of the workers are constrained to work for at most one firm. We assume that all workers are connected to all firms and that firms' as well as agents' preferences satisfy substitutability. This is a special case of a supply chain models in which the graph of potential interactions contains edges <u>from</u> all workers <u>to</u> all firms.*[23] *Note that chain stability reduces to pairwise stability here, since there are no intermediaries and a chain can thus involve at most two agents. How many workers can be allowed to demand multiple contracts if we want the market structure to be weakly and strongly acyclic, respectively?*

*In order to satisfy weak acyclicity we need to guarantee that every cycle contains at least one capacity constrained agent. Note that for any pair of workers there exists a cycle that contains only these two workers and a pair of firms. Since only workers can be capacity constrained, weak acyclicity thus requires us to assume that at most one worker can work for multiple firms. In order to satisfy strong acyclicity, both workers would have to be capacity constrained so the market has to be a many-to-one matching market. As a corollary to Theorem 7 we thus obtain that the core of a many-to-many matching market is equivalent to the set of pairwise stable matchings for all substitutable preference profiles if and only if the market is actually many-to-one.*[24]

To conclude, we now discuss three (independent) corollaries of Theorem 7.

[23]The direction of the edges is arbitrary and we could also direct the edges from firms to workers. A direction of edges has to be chosen to embed this two-sided matching model in a supply chain model.

[24]The interested reader may note that if there are no intermediaries, as in Application 2, the preferences in counterexample used to prove that $(ii) \Rightarrow (i)$ of Theorem 7 are responsive in the sense of Roth (1985). Hence, the statements apply equally well to the class of all responsive preferences. We thank Lars Ehlers for pointing this out.

As previously mentioned, the core coincides with the set of pairwise stable matchings in many-to-one matching models with substitutable preferences (Roth and Sotomayor (1991)). Note that this is a special case of our result since (i) chain reduces to pairwise stability if there are no intermediaries, (ii) agents on one side of the market can sign at most one contract, and (iii) each cycle must contain at least two agents from each of the two market sides and thus at least two capacity constrained agents.

Next, consider the *unit capacity model* in which each agent can sign at most one upstream and at most one downstream contract. Ostrovsky (2008) shows that in the unit capacity model the core coincides with the set of chain stable networks. To see that this is a direct consequence of Theorem 7 note that in the absence of directed cycles any undirected cycle must contain at least one source and at least one sink. In the unit capacity model sources and sinks of a cycle are always capacity constrained. Hence, any undirected cycle must have at least two capacity constrained agents and the market structure is strongly acyclic. Note that Ostrovsky's result does not in general imply core equivalence for the many-to-one matching model with substitutable preferences.

Finally, consider a many-to-one matching model with substitutes and complements. We assume that firms can hire an arbitrary number of workers, while workers can work for at most one firm. Assume that it is possible to decompose the set of workers into two sets $W_1$ and $W_2$ such that all firms view two workers from the same set as substitutes (in the sense of SSS) and workers from different sets as complements (in the sense of CSC). This is a variant of the model studied in Sun and Yang (2006) and Sun and Yang (2009). Ostrovsky (2008) discusses how this can be formulated as a supply chain model in which the set of sellers of basic inputs comprises $W_1$, the set of consumers of final products comprises $W_2$, and the set of intermediaries comprises all firms. Note that since there are no edges connecting two firms or two workers, any undirected cycle must contain at least two workers. Since all workers are capacity constrained the market structure is strongly acyclic so that the core coincides with the set of chain stable networks. Hence, we obtain the non-emptiness of a core network as a corollary to the existence of a chain stable network for this model.

## 3.6 Conclusion

This paper showed that the structural properties of supply chain models are important for the relationship between (cooperative) solution concepts. Weak acyclicity was shown to be necessary and sufficient for (i) the equivalence of chain and group stability, (ii) the core stability of chain stable networks, (iii) the efficiency of chain stable networks, (iv) the existence of group stable networks, and (iv) the existence of efficient individually stable networks. In the second part of the paper we derived some further results on the competing stability concept of the core. In particular, it was shown that strong acyclicity is necessary and sufficient for core equivalence to obtain. We have argued that our results can be interpreted as a justification of chain stability on basis of efficiency and robustness considerations. The cooperative foundation shows that this stability concept is a reasonable allocative goal for markets that fit the assumptions of the supply chain model. An important open question for future research is how such markets would have to be organized in order to reach this goal when agents act strategically.

# Bibliography


Abdulkadiroglu, A. (2005): "College Admissions with Affirmative Action," *International Journal of Game Theory*, 33(4), 535 – 549.

Abdulkadiroglu, A., Y.-K. Che, and Y. Yasuda (2008): "Expanding "Choice" in School Choice," Working Paper, Columbia University.

Abdulkadiroglu, A., and L. Ehlers (2007): "Controlled School Choice," Working Paper, Montreal University.

Abdulkadiroglu, A., P. Pathak, and A. Roth (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*.

Abdulkadiroglu, A., P. Pathak, A. Roth, and T. Sönmez (2006): "Changing the Boston School Choice Mechanism," Working Paper, Harvard University.

Abdulkadiroglu, A., and T. Sönmez (2003): "School Choice - A Mechanism Design Approach," *American Economic Review*, 93(3), 729 – 747.

Abdulkadiroglu, A., and T. Sonmez (1998): "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems," *Econometrica*, 66, 689 – 701.

——— (1999): "House Allocation with Existing Tenants," *Journal of Economic Theory*, 88, 233 – 260.

Abeledo, H., and G. Isaak (1992): "A characterization of graphs that ensure the existence of stable matchings," *Mathematical Social Sciences*, pp. 93 – 96.

Alcalde, J., and S. Barbera (1994): "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems," *Economic Theory*, 4, 417 – 435.

Alcalde, J., and A. Romero-Medina (2000): "Simple Mechanisms to Implement the Core of College Admissions Problems," *Games and Economic Behavior*, pp. 294 – 302.

Balinski, M., and T. Sönmez (1999): "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84(1), 73 – 94.

Blair, C. (1988): "The Lattice Structure of the Set of Stable Matchings with Multiple Partners," *Mathematics of Operations Research*, pp. 619 – 628.

Braun, S., N. Dwenger, and D. Kübler (2008): "Telling the Truth May Not Pay Off: An Empirical Study of Centralised University Admissions in Germany," Working Paper, Technical University of Berlin.

Chen, Y., and T. Sönmez (2006): "School Choice: An experimental study," *Journal of Economic Theory*, 127(1), 202 – 231.

Dubins, L., and D. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *American Mathematical Monthly*, 88, 485 494.

Dutta, B., and J. Masso (1997): "Stability of Matchings When Individuals Have Preferences over Colleagues," *Journal of Economic Theory*, 75, 464 – 475.

Echenique, F., and J. Oviedo (2006): "A theory of stability in many-to-many matching markets," *Theoretical Economics*, pp. 233 – 273.

Echenique, F., and B. Yenmez (2007): "A solution to matching with preferences over colleagues," *Games and Economic Behavior*, 59, 46 – 71.

Ehlers, L. (2007): "Respecting Priorities when Assigning Students to Schools," Working Paper, Universite de Montreal.

Ehlers, L., and A. Erdil (2009): "Efficient Assignment Respecting Priorities," Working Paper, Universite de Montreal.

Erdil, A., and H. Ergin (2008): "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice," *American Economic Review*, pp. 669 – 689.

Ergin, H. (2002): "Efficient resource allocation on the basis of priorities," *Econometrica*, 70(6), 2489 – 2497.

Ergin, H., and T. Sönmez (2006): "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 90, 215 – 237.

Featherstone, C., and M. Niederle (2008): "Ex Ante Efficiency in School Choice Mechanisms: An Experimental Investigation," Working Paper, Stanford University.

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

Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587 – 601.

Guillen, P., and O. Kesten (2008): "On-Campus Housing: Theory vs. Experiment," Working Paper, Tepper School of Business, Carnegie Mellon University.

Haeringer, G., and F. Klijn (2008): "Constrained School Choice," Working Paper, Universitat Autonoma de Barcelona.

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

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

Hylland, A., and R. Zeckhauser (1979): "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 87, 293 – 314.

Jackson, M., and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, pp. 44 – 74.

Kalai, E., A. Postlewaite, and J. Roberts (1978): "Barriers to Trade and Disadvantageous Middlemen: Nonmonotonicity of the core," *Journal of Economic Theory*, pp. 200 – 209.

Kara, T., and T. Sonmez (1996): "Nash Implementation of Matching Rules," *Journal of Economic Theory*, 68, 425 – 439.

——— (1997): "Implementation of College Admission Rules," *Economic Theory*, 9, 197 – 218.

Kelso, A., and V. Crawford (1982): "Job matching, coalition formation, and gross substitutes," *Econometrica*, pp. 1483 – 1503.

Kesten, O. (2006): "On two competing mechanisms for priority-based allocation problems," *Journal of Economic Theory*, 127(1), 155 – 171.

Klaus, B., and F. Klijn (2005): "Stable Matchings and the Preferences of Couples," *Journal of Economic Theory*, 121, 75 – 106.

Klaus, B., and M. Walzl (2008): "Stable Many-to-Many Matchings with Contracts," Harvard Business School Working Paper 09-046.

Kojima, F. (2008): "Games of school choice under the Boston mechanism with general priority structures," *Social Choice and Welfare*, pp. 357–365.

Kojima, F., and P. Pathak (2009): "Incentives and Stability in Large Matching Markets," *American Economic Review*.

Konishi, H., and M. U. Ünver (2006): "Credible group stability in many-to-many matching problems," *Journal of Economic Theory*, pp. 57 – 80.

Miralles, A. (2008): "School Choice: The Case for the Boston Mechanism," Working Paper, Boston University.

Ostrovsky, M. (2008): "Stability in Supply Chain Networks," *American Economic Review*, pp. 897 – 923.

Papai, S. (2000): "Strategy-proof Assignment by Hierarchical Exchange," *Econometrica*, 68, 1403 – 1433.

——— (2004): "Unique stability in simple coalition formation games," *Games and Economic Behavior*, pp. 337 – 354.

——— (2007): "Exchange in a general market with indivisible goods," *Journal of Economic Theory*, pp. 208 – 235.

——— (2009): "Mixed Priority Matching," Working Paper, Concordia University Montreal.

Pathak, P. (2008): "Lotteries in Student Assignment: The Equivalence of Queueing and a Market-Based Approach," Working Paper, MIT.

Pathak, P., and T. Sönmez (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, pp. 1636 – 1652.

Peleg, B., and P. Sudhölter (2003): *Introduction to Theory of Cooperative Games*. Kluwer Academic.

Pycia, M., and U. Unver (2009): "A Theory of House Allocation and Exchange Mechanisms," Working Paper, Boston College.

Romero-Medina (1998): "Implementation of Stable Solutions in a Restricted Matching Market," *Review of Economic Design*, 3, 137 – 147.

Roth, A. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7, 617 – 628.

——— (1984a): "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 92, 991 – 1016.

——— (1984b): "Stability and Polarization of Interests in Job Matching," *Econometrica*, pp. 47 – 57.

——— (1985): "The College Admissions Problem is not equivalent to the Marriage Problem," *Journal of Economic Theory*, 36, 277 – 288.

——— (2002): "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica*, 70, 1341 – 1378.

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

Roth, A., and E. Peranson (1999): "The redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economc Review*, 89(4), 748 – 780.

Roth, A., and M. Sotomayor (1989): "The College Admissions Problem Revisited," *Econometrica*, pp. 559–570.

Roth, A., and M. Sotomayor (1991): *Two Sided Markets - A Study in Game Theoretic modeling*. Cambridge University Press.

Satterthwaite, M. (1979): "Strategy-proofness and Arrow's conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, 187 – 217.

Scheer, e. (1999): *Zentralstelle für die Vergabe von Studienplätzen 1973 - 1998.* Informations- und Pressestelle der ZVS.

Shapley, L., and H. Scarf (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 23 – 37.

Sonmez, T. (1997): "Games of Manipulation in Marriage Problems," *Games and Economic Behavior*, 20, 169 – 176.

——— (1999): "Strategy-Proofness and Essentially Single-Valued Cores," *Econometrica*, 67, 677 – 689.

Sonmez, T., and U. Unver (2008): "Matching, Allocation, and Exchange of Discrete Resources," Working Paper, Boston College.

Sotomayor, M. (1999): "Three remarks on the many-to-many stable matching problem," *Mathematical Social Sciences*, pp. 55 – 70.

Sun, N., and Z. Yang (2006): "Equilibria and Indivisibilities," *Econometrica*, pp. 1385 – 1402.

——— (2009): "A Double-Track Adjustment Process for Discrete Markets with Substitutes and Complements," *Econometrica*.

Svensson, L.-G. (1999): "Strategy-proof allocation of indivisible goods," *Social Choice and Welfare*, 16, 557–567.

# Appendices

## A.1 Appendix to Chapter 1

### Proof of Theorem 1

We show first that for any profile of strict applicant preferences $R$, if $Q$ is a pure strategy Nash equilibrium of $\Gamma^{ZVS}(R)$ then $f^{ZVS}(Q)$ must be a stable matching for the university admissions problem $R$. For economy of notation let $f^{ZVS}(Q) = (\mu^E_{ZVS}, \mu^W_{ZVS}, \mu^U_{ZVS})$ and $q^U$ be the capacity vector in step U. Condition (i) is satisfied irrespective of whether $Q$ is a Nash equilibrium or not, since only top-grade applicants can receive a place in step E, only wait-time applicants can obtain a place in step W, and a university never makes an offer to an unacceptable student in the college proposing deferred acceptance algorithm of step U. If an applicant $a$ is matched to an unacceptable (w.r.t. $R$) university she could profitably deviate by ranking just her most preferred university (w.r.t. $R$) for each step of the procedure: The worst thing that could happen is to be left unassigned if she submits the alternative report. Since $aP_a f^{ZVS}_a(Q)$ this is profitable so that $Q$ could not have been a Nash equilibrium of $\Gamma^{ZVS}(R)$. This shows that $f^{ZVS}(Q)$ has to satisfy (ii).

Suppose to the contrary that $f^{ZVS}(Q)$ does not match applicants as early as possible and that there is a top-grade applicant $a$ such that $\mu^U(a) = u$ and either $|\mu^E(u)| < q^E_u$ or $\pi^E_u(a) < \pi^E_u(\tilde{a})$ for some $\tilde{a} \in \mu^E(u)$. If $a$ had ranked $u$ as her most preferred university for step E, she would have been matched to $u$ in step E since in the first case at most $q^E_u - 1$ top-grade applicants could have applied to $u$ in the course of the Boston mechanism under $Q$ and in the second case at most $q^E_u - 1$ other (top-grade) applicants with higher $\pi^E_u$ priority than $\tilde{a}$ could have listed $u$ as their top choice in $Q^E$ (and thus in $\tilde{Q}^E$). Since applicants prefer to be matched as early as possible, this would be a profitable deviation for $a$. The argument in case of a wait-time applicant matched too late is completely analogous. Hence, $f^{ZVS}(Q)$ has to satisfy condition (iii).[25]

[25]These parts of the proof that concern the Boston mechanisms of steps E and W are similar to the arguments

Next, suppose that $f^{ZVS}(Q)$ satisfies (i) to (iii) but that there is an applicant-university pair $(a, u)$ such that (iv) is violated for some $i \in \{E, W, U\}$. Let $\tilde{Q}_a$ be an alternative report for applicant $a$ that lists only $u$ for each step of the procedure. Let $\tilde{Q} = (\tilde{Q}_a, Q_{-a})$, $f^{ZVS}(\tilde{Q}) = (\tilde{\mu}^E, \tilde{\mu}^W, \tilde{\mu}^U)$, and $\tilde{q}^U$ be the corresponding quota vector in step U. For $i \in \{E, W\}$ the same argument used to show that $f^{ZVS}(Q)$ satisfies (iii) can be used to establish that $\tilde{\mu}^i(a) = u$. So suppose that $i = U$. Let $A^U$ and $\tilde{A}^U$ denote the sets of applicants apart from $a$ who remain in the procedure by the beginning of step U under $Q$ and $\tilde{Q}$, respectively. It is clear that unless $\tilde{\mu}(a) = a$, $\tilde{Q}_a$ is a profitable deviation for $a$. So suppose that $\tilde{\mu}(a) = a$. Note that this implies $\tilde{A}^U \subseteq A^U$, $\tilde{q}^U_u = q^U_u$, as well as $|\tilde{\mu}^U(u)| = \tilde{q}^U$: If $a$ does not receive a place at $u$ in steps E or W under $\tilde{Q}$, there are (weakly) less rejections in steps E and W under $\tilde{Q}$ than under $Q$. In particular the set of applicants who are assigned a place at $u$ in steps E and W must be the same under $Q$ and $\tilde{Q}$. Furthermore, since $(a, u)$ is a pair such that (iv) is violated for $i = U$ it has to be the case that $a \succ^U_u u$. But then $u$ would always make an offer to $a$ before leaving a place unassigned so that we must have $|\tilde{\mu}^U(u)| = \tilde{q}^U_u$. This implies that if $a$ has not received an offer by $u$ in the first $t$ rounds of the CDA under $\tilde{Q}$ then all applicants in $\tilde{A}^U$ must have received weakly more offers in rounds 1 through $t$ than they received in the first $t$ rounds of the CDA under $Q$. Hence, if $a$ does not receive an offer by $u$ in the course of the CDA under $\tilde{Q}$, all applicants in $\tilde{A}^U$ receive a weakly better assignment (wrt to $\tilde{Q}^U_{-a}$) than under $Q$. In particular, any applicant in $\tilde{A}^U$ who rejected an offer by $u$ in the CDA under $Q$ will also reject an offer by $u$ in the CDA under $\tilde{Q}$. But then we must have $|\mu^U(u)| = q^U_u$ and, since (iv) is violated, there has to be an applicant $\tilde{a} \in \mu^U(u)$ such that $a \succ^U_u \tilde{a}$. By the above this implies that if $a' \notin \mu^U(u)$ and $a' \succ^U_u a$ then $a'$ will reject an offer by $u$ in the CDA under $Q$. But then $\tilde{\mu}^U(u)$ has to contain at least one applicant $a''$ with $a \succ^U_u a''$. Hence, $u$ must have made an offer to $a$ in the CDA under $\tilde{Q}$ (which $a$ would have accepted). This contradicts the assumption that $\tilde{\mu}(a) = a$ and shows that $f^{ZVS}(Q)$ has to satisfy (iv).

Now let $\mu$ be a stable matching in the university admissions problem $R$. We construct a Nash equilibrium $Q$ such that $f^{ZVS}(Q) = \mu$. Let $a$ be an arbitrary applicant. If $\mu(a) = a$, let $a$ rank her six most preferred acceptable universities according to $R_a$ steps E and W, and all acceptable universities for step U. If $\mu(a) = u$, let $a$ rank only $u$ for all parts of the procedure. Let $Q$ be the resulting strategy profile.

We show first that $f^{ZVS}(Q) = \mu$. Let $f^{ZVS}(Q) = (\tilde{\mu}^E, \tilde{\mu}^W, \tilde{\mu}^U)$. Suppose $a$ is such that $\mu(a) = a$ but $\tilde{\mu}^E(a) = u \in \mathcal{U}$. By construction of $Q$, all applicants in $\mu^E(u)$ rank $u$ first in their

in Ergin and Sönmez (2006).

submitted ranking for step E. Since $a$ ranks only acceptable universities, $uP_a\mu(a)$ so that we obtain a contradiction to (iv). If $a$ is such that $\mu^U(a) = u \in \mathcal{U}$, then the only possibility for $\mu^E(a) \neq \tilde{\mu}^E(a)$ is that $a$ receives a place at $u$ in step E of the ZVS procedure under $Q$. This is a contradiction to (iii). Hence, $\tilde{\mu}^E(a) = a$ if $\mu^E(a) = a$. Given the construction of $Q$, this implies that any applicant who receives a place at some university $u$ under $\mu^E$ must be assigned to $u$ in step E of the ZVS procedure for $Q$. Hence, $\mu^E = \tilde{\mu}^E$ and a completely analogous argument can be used to establish that $\mu^W = \tilde{\mu}^W$. To see that $\tilde{\mu}^U = \mu^U$, note that, by the above, any applicant $a$ with $\mu^U(a) \in \mathcal{U}$ will not be assigned in steps E or W of the ZVS procedure under Q. All of these applicants rank only their assigned university under $\mu^U$ for step U while all other unassigned applicants rank their six most preferred acceptable universities (w.r.t. $R$). If one of the unassigned applicants received a place in step U of the ZVS procedure under Q, (iv) would be violated.

Next, we show that no applicant has an incentive to deviate from the proposed strategy-profile. Let $\tilde{Q}_a$ be an alternative report for applicant $a$, $\tilde{Q} = (\tilde{Q}_a, Q_{-a})$, and $f^{ZVS}(\tilde{Q}) = (\tilde{\mu}^E, \tilde{\mu}^W, \tilde{\mu}^U)$. Note that the sets of top-grade and wait-time applicants are the same as under $Q$ since, for $i \in \{E, W\}$, $\tilde{Q}^i_a$ contains at least one university by the no empty lists assumption. Now suppose to the contrary that $\tilde{\mu}(a)\tilde{P}_a\mu(a)$. It cannot be the case that $\tilde{\mu}^E(a) = u = \mu^U(a)$ or $\tilde{\mu}^W(a) = u = \mu^U(a)$, since all applicants in $\mu^E(u)$ and $\mu^W(u)$ apply to $u$ in the first round of the Boston mechanism under $\tilde{Q}$ so that $\mu$ could not satisfy (iii) otherwise. We can show in a similar fashion that $a$ cannot obtain a university strictly preferred to $\mu(a)$ (w.r.t. $R_a$) in steps E or W. Hence, it remains to be shown that $a$ cannot prefer $\tilde{\mu}^U(a)$ over $\mu(a)$. Consider first an applicant $a$ such that $\mu^E(a) = \mu^W(a) = a$. Given (iii) and (iv), no alternative report $\tilde{Q}_a$ that leads to different assignments in steps E or W can be profitable for $a$. Suppose then that contrary to what we want to show, $\tilde{Q}_a$ is a profitable deviation for $a$. We can assume w.l.o.g. that $\tilde{Q}^E_a = Q^E_a$ and $\tilde{Q}^W_a = Q^W_a$ by the above. But then the set of applicants who are unmatched by the beginning of step U of the ZVS procedure under $\tilde{Q}$ contains in particular all applicants who are matched to some university under $\mu^U$. By assumption, $\tilde{\mu}^U(a)P_a\mu(a)$ so that there must be some university $u \in \mathcal{U}$ such that $\tilde{\mu}^U(a) = u$. But all applicants in $\mu^U(u)$ were "available" to $u$ in step U of the ZVS procedure under $\tilde{Q}$ since they all ranked only $u$. Hence, $u$ must prefer $a$ to at least one of the applicants assigned to it under $\mu^U$ or must have had some unfilled capacity in $\mu^U(u)$ even though $a$ is acceptable (w.r.t. $\succ^U_u$). This is a contradiction to (iv). Now consider an applicant $a$ such that for some $u \in \mathcal{U}$, $\mu^E(a) = u$. By the disjoint sets assumption, $a$ can never obtain a place at some university in step W. By (iv) and the construction of $Q$, there is no alternative report for $a$ such that she obtains a strictly preferred (w.r.t. $R_a$) university in step E. Thus the only way that $a$ could potentially improve upon her assignment under $\mu$ is that $\tilde{\mu}^E(a) = a(= \tilde{\mu}^W(a))$. But the only applicants who could take the leftover seat at $u$ in step E are those who are either unassigned under $\mu$ or who are matched to $u$ under $\mu^U$. In any case, for all universities $u' \neq u$, all applicants in $\mu^U(u')$ remain in the procedure by the beginning of step U under $\tilde{Q}$. If $\tilde{\mu}^U(a) P_a u$, we must thus obtain a contradiction to (iv). This completes the proof.

□

## An example showing that No Empty Lists is Restrictive

The following example shows that the No Empty Lists assumption is restrictive. There are seven applicants $a_1, a_2, a_3, a_4, a_5, a_6, a_7$ indexed in order of increasing average grades. For simplicity, we assume that there are only two universities $u$ and $u'$ who have one place to allocate in each of the three steps of the ZVS procedure. Preferences of applicants are as follows:

| $R_{\mathcal{A}}$ | $R_{a_1}$ | $R_{a_2}$ | $R_{a_3}$ | $R_{a_4}$ | $R_{a_5}$ | $R_{a_6}$ | $R_{a_7}$ |
|---|---|---|---|---|---|---|---|
| | $u$ | $u$ | $u'$ | $u'$ | $u'$ | $u$ | $u'$ |
| | | | $u$ | | | | |

Assuming that $a_5$ and $a_6$ are the applicants with the longest waiting time, $a_3 \succ_u^U a_2$ as well as $a_5 \succ_u^U a_4 \succ_u^U a_3$, and that all applicants submit their true ranking of universities for each step of the procedure, the outcome of the ZVS procedure is

$$f^{ZVSE}(R) = \begin{matrix} u & u' \\ a_1 & \emptyset \end{matrix}, \qquad f^{ZVSW}(R) = \begin{matrix} u & u' \\ a_6 & a_7 \end{matrix}, \qquad f^{ZVSU}(R, \succ^U) = \begin{matrix} u & u' \\ a_3 & \{a_4, a_5\} \end{matrix}.$$

Note that if we keep the profile of reports by everyone but $a_2$ fixed, $a_2$ cannot obtain a place at $u$ if she applies for a place in step E. Suppose then that she decides to apply only for step U and submits $Q^U_{a_2} = u$. If everyone else submits the same preferences as before, $a_3$ would be a top-grade applicant and could obtain a place at her most preferred university $u'$ in step E of the ZVS procedure. But then $a_2$ would receive a place at $u$ in step U of the procedure since no one else will apply to $u$ in that step. Thus, she benefits from not applying for a place in step E.

□

## Proof that $\mu_1$ and $\mu_2$ are the only stable matchings in Example 2

Note first that $a_1$ has to be assigned to $u_1$ in step E by (iii) and (iv). This implies that if $a_2$ is matched in step E, she must receive a place at $u_3$ by (iv) for $i = E$. Now note that if $a_2$ is matched in step E, $a_3$ must be matched to $u_2$ in step E. Otherwise, one place at $u_2$ would remain empty and $a_3$ would have to receive a place at $u_3$ in step U by (iii). But then $a_4$ and $a_5$ would take the two places at $u_2$ in step U. But this implies that $a_1$ must have strictly higher priority according $\succ_{u_1}^{U}$ than one of the applicants assigned a place at $u_1$ in step U (or $u_1$ does not fill its capacity reserved for $U$). Hence, $a_3$ must be matched to $u_2$ in step E if $a_3$ receives a place at $u_2$ in step E. It is easy to see that wait-time applicants cannot do better than being assigned a place in step W and that there is a unique stable matching for this subproblem. This is easily seen to imply that there is a unique stable matching where $a_2$ is matched in step E.

If on the other hand, $a_2$ is not matched in step E, she must receive a place at $u_1$ in step U by (iii). This implies directly that $a_4$ and $a_5$ must both obtain a place at $u_2$ in step U and that $a_3$ must be assigned to $a_3$ in step E. It is again easy to see that the wait-time applicants cannot do better than being assigned a place in step W.

□

## Omitted details of the ZVS Procedure

This appendix lists some of the more substantial simplifications made in the main body of the text. Readers interested in all details of the current ZVS procedure may still want to consult the *Vergabeverordnung ZVS [Stand: WS 2008/2009]* (available at www.zvs.de).

**Capacities:** The total number of places at each university is determined by the application of federal laws. For each state there is a so called *Kapazit"atsverodnung* (KaPVO) which prescribes a formula for calculating the number of applicants a university can admit on basis of the number of professors, available teaching facilities, and so on.[26]

**Special Quotas:** Up to approximately fifteen percent of total available places are allocated in advance among foreign applicants, applicants pursuing a second university degree, and so on. These applicants are not allowed to participate in the regular assignment procedure.

[26] There has been some discussion about the KaPVO in recent years, see e.g. *Die fiese Formel* in Die Zeit, Nr. 39(2007).

**Step E:** The education system in Germany is federalized and the general opinion is that average grades are not directly comparable across federal states. For this reason, there are actually sixteen separate assignment procedures in step E, one for each federal state. This is achieved by splitting the 20 percent of (remaining) capacity available in step E into sixteen parts. In the assignment procedure of a given federal state only those applicants are considered who have received their high school diploma in this state.

**Step U:**

- Once assignments are determined by the ZVS procedure described in section 2, successful applicants have to enroll at their assigned university. If some applicants fail to do so, their places are allocated according to the rules of step U. Here, only those applicants are considered who did not receive a place in previous rounds of the assignment procedure. Again, students have to enroll at their assigned university (if any) and if they fail to do so, another round of step U is used to allocate remaining places (again only students who were not previously assigned a place are considered). Any places that remain after all of this are allocated via lottery by universities.
- In order to prevent multiple rounds of the assignment procedure in step U, a university can demand the ZVS to overbook its capacities. Thus, a university with, say 100 places, may ask the ZVS to be assigned 150 applicants since it expects some students not to accept their assigned places.

**Lotteries:** If a university does not fill its capacity in the ZVS procedure, remaining places are allocated on basis of lottery. Each university conducts its own lottery and applicants have to apply to universities directly in order to participate.

## Evaluation Procedures in the current ZVS procedure

In this Appendix we provide some further details on the different evaluation procedures used by universities in step U. All of the below concerns the ZVS procedure for the winter term 2008/2009. The evaluation process takes place after assignments in steps E and W have been determined and only those applicants who did not receive a place in these steps are considered. In principle, the ZVS informs each university about all remaining applicants who have listed the university in their ranking for step U.

A university may, however, limit the set of applicants it will consider for step U in advance on basis of its rank in the preference lists submitted for step U, average grades, or a combination of the two criteria. For example, a university with, say, a hundred seats to be allocated in step

U may consider only the 300 applicants with the best average grades among those who ranked it first. This practice is called *pre-selection* and the ZVS informs each university only about those applicants who "survived" its pre-selection process.

In case an applicant is not rejected in the pre-selection process of a university, the ZVS provides the university with detailed information including its rank in the submitted preference list, average grade, waiting-time, and so on. Universities can then use average grades, interviews, statements of purpose, completion of on-the-job training in a relevant field, prizes in scientific competition, and so on, to evaluate remaining applicants. Furthermore, universities are allowed to split their capacity into several parts and to apply different admission criteria across these parts. For example, a university may decide to allocate 50 percent of places according to average grade and 50 percent on basis of performance in an interview. In this case, the university has to specify which place an applicant receives if she could be admitted in more than one of these quotas. The official information brochure of the ZVS states that *in the determination of an applicant's rank average grade has to be a decisive factor*,[27] although the exact requirement that needs to be fulfilled is not specified.

A university uses a *mechanical evaluation procedure* if the marginal cost of evaluating an additional candidate according to its criteria is negligible. Under this label we summarize all universities who do not use "subjective" criteria such as performance in interviews or the evaluation of statements of purpose. While such universities may still require to elicit additional information from applicants, computing the rank of an applicant is completely mechanical. In this category, we also include universities who use standardized tests in their evaluation procedure. For example, several universities offering medical subjects use the outcome of a standardized medical subject test in their evaluation. Applicants have to pay a fixed fee in order to take the test and tests are graded by an independent company. If, for example, a university uses some weighted average of average grade and performance in the standardized test to rank applicants, its cost of considering one additional applicant amounts to the (computational) cost of calculating the weighted average of two numbers.

In the first table, $\#\mathcal{U}$ lists the number of universities, $\#Pre_k$ lists the number of universities that consider only applicants who ranked them at least $k$th ($k = 1, \ldots, 4$), $\#M$ lists the number of universities who use a mechanical evaluation procedure, and $\#M + Pre$ lists the number of universities who have a mechanical evaluation procedure but only consider applicants who rank them sufficiently high.

In the second table, $\#INT$ is the number of universities that use interviews to allocate at

[27] *Merkblatt M09: Auswahlverfahren der Hochschulen*, available at www.zvs.de. Translation by the author.

| Subject | # $\mathcal{U}$ | #Pre 1 | #Pre 2 | #Pre 3 | #Pre 4 | # M | # M+Pre |
|---|---|---|---|---|---|---|---|
| Biology | 2 | 0 | 1 | 0 | 0 | 2 | 1 |
| Psychology | 11 | 2 | 1 | 1 | 0 | 11 | 4 |
| Vet. Med. | 5 | 3 | 0 | 0 | 0 | 2 | 0 |
| Pharmacy | 22 | 0 | 2 | 3 | 0 | 20 | 5 |
| Dentistry | 29 | 4 | 3 | 6 | 1 | 22 | 8 |
| Medicine | 34 | 6 | 4 | 9 | 1 | 24 | 12 |

Table 1: Preselection and Mechanical Evaluation

| Subject | $\#INT$ | $\#INT_{>0.5}$ | $\#INT_{\leq 0.5}$ |
|---|---|---|---|
| Biology | 0 | 0 | 0 |
| Psychology | 0 | 0 | 0 |
| Vet. Med. | 3 | 3 | 0 |
| Pharmacy | 2 | 2 | 0 |
| Dentistry | 7 | 2 | 5 |
| Medicine | 10 | 3 | 7 |

Table 2: Interviews in the Evaluation Process

least part of their capacity, $\#INT_{>0.5}$ is the number of universities who assign more than half of their seats on basis of interviews, and $\#INT_{\leq 0.5}$ is the number of universities who assign at most half (but at least one) of their seats on basis of interviews.

## A short history of the ZVS mechanism

The last major revision of the ZVS procedure took place in the winter term 2005/2006 where the three part (steps E, W and U) assignment procedure was introduced. There have been a number of changes from this procedure to the one currently employed procedure. The following is a list of the most important changes.

**Floating Quotas** Initially, places not taken in step E were allocated in step W (in addition to the 20 percent of places reserved for wait-time applicants). As Braun, Dwenger, and Kübler (2008) mention, this led to some universities allocating far more than 20 percent of their places to wait-time applicants since very few top-grade applicants demanded places in step E. Presumably, this one of the main reasons for changing the procedure and moving unused places directly from step E to step U.

**Preference Reversals in step U** Applicants were allowed to reverse their preferences in step U once the evaluation process of universities was complete. Initially, applicants who could be admitted at several universities in the first round of the assignment procedure of step U were allowed to take the offer of a university other than the one in highest reported preference. At some point between 2005/2006 and 2008/2009 applicants were even allowed to reorder their whole ranking of universities in step U after the evaluation process. We conjecture that one of the main reasons for abandoning this practice is that universities could not guarantee that a candidate they interviewed would accept an offer.[28]

**Algorithm in Step U** Previous versions of the description of the assignment procedure have been relatively vague concerning the exact form of the assignment procedure in step U. It was not exactly clear whether a version of the Boston mechanism in which universities made the offers or the college/university proposing deferred acceptance algorithm were used. The current version of the official description of the assignment procedure says that *at each point an applicant can be admitted at multiple universities, only the offer of the most preferred university remains valid.*[29] Thus, the ZVS algorithm looks at the lists submitted by universities to check whether an applicant can be admitted at more than one university. If this is the case, the applicant is deleted from the lists of all but her most preferred university among those that offered admission. If this procedure is iterated until no applicants are left, we arrive at exactly the description of the college proposing deferred acceptance algorithm.

## A.2 Appendix to Chapter 2

### Proof of Proposition 5

Suppose to the contrary that $\succ^1$ contains an ambiguous 1-tie but that $f$ is a strategy-proof and constrained efficient selection from the stable correspondence. W.l.o.g. we can assume that there are exactly four students $1, 2, 3, 4$ and two specialized schools $s_1, s_2$ such that

$$1 \succ_{s_1} 3 \succ_{s_1} 2 \text{ and } 2 \succ_{s_2} 4 \succ_{s_2} 1. \tag{1}$$

Let $s_3$ be one of the non-specialized schools. To derive a contradiction, we consider six preference profiles which define a cycle in the space of preference profiles. The following diagram summarizes the preference profiles used in our proof. Arrows indicate how we move between the profiles.

[28] Given that a couple of universities considered only applicants who ranked them first, it still guaranteed that an applicant could not have interviewed at some of the other universities.

[29] *Vergabeverordnung ZVS, Stand: WS 2008/2009* §10.(5), which can be found at www.zvs.de. Translation by the author.

$$
\begin{array}{c|cccc} R^2 & R_1^1 & R_2^1 & R_3^2 & R_4^1 \\ \hline & s_2 & s_1 & s_1 & s_3 \\ & s_3 & s_3 & s_3 & \\ & s_1 & s_2 & & \end{array}
\;\longleftarrow\;
\begin{array}{c|cccc} R^1 & R_1^1 & R_2^1 & R_3^1 & R_4^1 \\ \hline & s_2 & s_1 & s_3 & s_3 \\ & s_3 & s_3 & & \\ & s_1 & s_2 & & \end{array}
\;\longleftarrow\;
\begin{array}{c|cccc} R^6 & R_1^1 & R_2^1 & R_3^1 & R_4^2 \\ \hline & s_2 & s_1 & s_3 & s_3 \\ & s_3 & s_3 & & s_2 \\ & s_1 & s_2 & & \end{array}
$$

$$
\downarrow \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \uparrow
$$

$$
\begin{array}{c|cccc} R^3 & R_1^1 & R_2^3 & R_3^2 & R_4^1 \\ \hline & s_2 & s_1 & s_1 & s_3 \\ & s_3 & s_3 & s_3 & \\ & s_1 & & & \end{array}
\;\longrightarrow\;
\begin{array}{c|cccc} R^4 & R_1^1 & R_2^3 & R_3^2 & R_4^2 \\ \hline & s_2 & s_1 & s_1 & s_3 \\ & s_3 & s_3 & s_3 & s_2 \\ & s_1 & & & \end{array}
\;\longrightarrow\;
\begin{array}{c|cccc} R^5 & R_1^1 & R_2^1 & R_3^2 & R_4^2 \\ \hline & s_2 & s_1 & s_1 & s_3 \\ & s_3 & s_3 & s_3 & s_2 \\ & s_1 & s_2 & & \end{array}
$$

We start with the profile $R^1$. Let

$$
\mu = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_2 & s_1 & s_3 & 4 \end{pmatrix} \text{ and } \bar{\mu} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_2 & s_1 & 3 & s_3 \end{pmatrix}.
$$

It is straightforward that these are the only constrained efficient matchings for the profile $R^1$. Thus, we must have $f(R^1) = \mu$ or $f(R^1) = \bar{\mu}$. By the symmetries of the example, we can assume $f(R^1) = \mu$ without loss of generality.

Now let $R_3^2 : s_1, s_3$ and $R^2 = (R_3^2, R_{-3}^1)$. By strategy-proofness, $f_3(R^2) \neq 3$. Note that for $R^2$ there is no constrained efficient matching that assigns 3 to $s_3$. Hence, we must have $f_3(R^2) = s_1$. It is easy to see that this in conjunction with constrained efficiency implies

$$
f(R^2) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_2 & s_3 & s_1 & 4 \end{pmatrix}. \tag{2}
$$

Next, suppose 2 declares $s_2$ unacceptable, that is, consider $R_2^3 : s_1, s_3$ and the profile $R^3 = (R_1^1, R_2^3, R_3^2, R_4^1)$. By strategy-proofness we must have $f_2(R^3) = s_3$ so that constrained efficiency implies $f(R^3) = f(R^2)$.

Now consider the profile $R^4 = (R_1^1, R_2^3, R_3^2, R_4^2)$. By strategy-proofness, $f_4(R^4) \neq s_3$. Since $4 \succ_{s_2} 1$ and 1 and 4 are the only students who would like to be assigned to $s_2$, we have $f_1(R^4) \neq s_2 = f_4(R^4)$. If $f_1(R^4) = s_3$, 1 and 4 would form a SIC under $f(R^4)$ - a contradiction to constrained efficiency. Thus, $f_1(R^4) = s_1$. But then $f_3(R^4) \neq s_3$ since otherwise by $3 \succ_{s_1} 2$, 1 and 3 would form a SIC. Thus, $f_3(R^4) = 3$ so that

$$
f(R^4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_1 & s_3 & 3 & s_2 \end{pmatrix}. \tag{3}
$$

Now consider the profile $R^5 = (R_1^1, R_2^1, R_3^2, R_4^2)$. By strategy-proofness, $f_2(R^5) = s_3$ and similarly to above, $f(R^5) = f(R^4)$.

Finally, consider the profile $R^6 = (R_1^1, R_2^1, R_3^1, R_4^2)$. By strategy-proofness, $f_3(R^6) = 3$. Since the first choices of the other three agents are compatible, constrained efficiency implies $f(R^6) = f(R^5)$.

This yields the desired contradiction since $s_3 = f_4(R_4^2, R_{-4}^1) P_4^1 f_4(R^1) = 4$. Hence, there is no strategy-proof and constrained efficient mechanism. $\square$

## Proof of Proposition 6

Consider a $\succ^1$ that contains ambiguity at the top. Let $s_1, s_2, s_3$ be three specialized schools and $1, 2, 3, 4 \in I$ be four distinct students such that[30]

| $\succ_{s_1}$ | $\succ_{s_2}$ | $\succ_{s_3}$ |
|---|---|---|
| 1 | 2 | 1 |
| 3 | 3 | 2 |
| 2 | 1 | 4 |
| 4 | 4 | 3 |

Let $s_4$ be one of the non-specialized schools. Suppose to the contrary that there exists a strategy-proof and constrained efficient rule $f$.

The main part of the proof considers four preference profiles, which are summarized in the following diagram. As in the proof of Proposition 5, arrows indicate how we move between preference profiles.

$$
\begin{array}{c|cccc}
R^1 & R_1^1 & R_2^1 & R_3^1 & R_4^1 \\
\hline
 & s_4 & s_1 & s_4 & s_3 \\
 & & s_4 & s_3 & \\
 & & & s_1 &
\end{array}
\qquad\qquad
\begin{array}{c|cccc}
R^2 & R_1^2 & R_2^2 & R_3^2 & R_4^1 \\
\hline
 & s_2 & s_4 & s_4 & s_3 \\
 & s_4 & & s_3 & \\
 & & & s_2 &
\end{array}
$$

$$
\downarrow \qquad\qquad\qquad\qquad\qquad\qquad \downarrow
$$

$$
\begin{array}{c|cccc}
R^3 & R_1^1 & R_2^3 & R_3^1 & R_4^2 \\
\hline
 & s_4 & s_4 & s_4 & s_1 \\
 & & s_3 & s_3 & s_3 \\
 & & & s_1 & s_2
\end{array}
\quad \leftrightarrow \quad
\begin{array}{c|cccc}
R^4 & R_1^1 & R_2^3 & R_3^2 & R_4^2 \\
\hline
 & s_4 & s_4 & s_4 & s_1 \\
 & & s_3 & s_3 & s_3 \\
 & & & s_2 & s_2
\end{array}
$$

[30]Note that due to the symmetries of the definition it is without loss of generality to assume that $1 \succ_{s_3} 2$.

Let $R^1_1 : s_4; R^1_2 : s_1, s_4; R^1_3 : s_4, s_3, s_1; R^1_4 : s_3$, and $R^1$ be the resulting preference profile. It is easy to see that there are exactly two constrained efficient matchings at $R^1$,

$$\mu^1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_4 & 2 & s_1 & s_3 \end{pmatrix} \text{ and } \bar{\mu}^1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & s_1 & s_4 & s_3 \end{pmatrix}.$$

**Claim 1:** $f(R^1) = \mu^1$.

**Proof of Claim 1.** Suppose to the contrary that $f(R^1) = \bar{\mu}^1$. Starting from profile $R^1$ we will consider the following preference profiles for these students:

$$
\begin{array}{ccccc}
\begin{array}{c|cccc} R^1 & R^1_1 & R^1_2 & R^1_3 & R^1_4 \\ \hline & s_4 & s_1 & s_4 & s_3 \\ & & s_4 & s_3 & \\ & & & s_1 & \end{array}
& &
\begin{array}{c|cccc} R^{1,7} & R^{1,1}_1 & R^1_2 & R^{1,2}_3 & R^{1,3}_4 \\ \hline & s_4 & s_1 & s_3 & s_4 \\ & s_1 & s_4 & s_4 & \\ & & & & \end{array}
& \rightarrow &
\begin{array}{c|cccc} R^{1,8} & R^{1,1}_1 & R^1_2 & R^{1,2}_3 & R^{1,1}_4 \\ \hline & s_4 & s_1 & s_3 & s_4 \\ & s_1 & s_4 & s_4 & s_3 \\ & & & & \end{array}
\\
\downarrow & & \uparrow & & \downarrow \\
\begin{array}{c|cccc} R^{1,1} & R^1_1 & R^1_2 & R^{1,1}_3 & R^1_4 \\ \hline & s_4 & s_1 & s_4 & s_3 \\ & & s_4 & s_3 & \\ & & & & \end{array}
& &
\begin{array}{c|cccc} R^{1,6} & R^{1,1}_1 & R^{1,1}_2 & R^{1,2}_3 & R^{1,3}_4 \\ \hline & s_4 & s_1 & s_3 & s_4 \\ & s_1 & s_4 & s_4 & \\ & & s_3 & & \end{array}
& \leftarrow &
\begin{array}{c|cccc} R^{1,5} & R^{1,1}_1 & R^{1,1}_2 & R^{1,2}_3 & R^{1,1}_4 \\ \hline & s_4 & s_1 & s_3 & s_4 \\ & s_1 & s_4 & s_4 & s_3 \\ & & s_3 & & \end{array}
\\
\downarrow & & & & \uparrow \\
\begin{array}{c|cccc} R^{1,2} & R^1_{1,1} & R^1_2 & R^{1,1}_3 & R^1_4 \\ \hline & s_4 & s_1 & s_4 & s_3 \\ & s_1 & s_4 & s_3 & \\ & & & & \end{array}
& \rightarrow &
\begin{array}{c|cccc} R^{1,3} & R^{1,1}_1 & R^{1,1}_2 & R^{1,1}_3 & R^1_4 \\ \hline & s_4 & s_1 & s_4 & s_3 \\ & s_1 & s_4 & s_3 & \\ & & s_3 & & \end{array}
& \rightarrow &
\begin{array}{c|cccc} R^{1,4} & R^{1,1}_1 & R^{1,1}_2 & R^{1,2}_3 & R^1_4 \\ \hline & s_4 & s_1 & s_3 & s_3 \\ & s_1 & s_4 & s_4 & \\ & & s_3 & & \end{array}
\end{array}
$$

Let $R^{1,1}_3 : s_4, s_3$ and $R^{1,1} = (R^1_1, R^1_2, R^{1,1}_3, R^1_4)$. By strategy-proofness (for student 3) and constrained efficiency we must have that $f(R^2) = f(R^1)$. Let $R^{1,1}_1 : s_4, s_1$ and $R^{1,2} = (R^{1,1}_1, R^1_2, R^{1,1}_3, R^1_4)$. By strategy-proofness (for student 1) and $1 \succ_{s_1} 3 \succ_{s_1} 2 \succ_{s_1} 4$, we must have $f_1(R^{1,2}) = s_1$. Now if $f_2(R^{1,2}) = s_4$, then 1 and 2 form a SIC, a contradiction. Thus, by constrained efficiency, we must have that $f(R^{1,2}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_1 & 2 & s_4 & s_3 \end{pmatrix}$.

Let $R^{1,1}_2 : s_1, s_4, s_3$ and $R^{1,3} = (R^{1,1}_1, R^{1,1}_2, R^{1,1}_3, R^1_4)$. By strategy-proofness (for student 2) and constrained efficiency we must have that $f(R^{1,3}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_1 & s_3 & s_4 & 4 \end{pmatrix}$.

Let $R^{1,2}_3 : s_3, s_4$ and $R^{1,4} = (R^{1,1}_1, R^{1,1}_2, R^{1,2}_3, R^1_4)$. By strategy-proofness (for student 3) and constrained efficiency we must have that $f(R^{1,4}) = f(R^{1,3})$.

Let $R^{1,1}_4 : s_4, s_3$ and $R^{1,5} = (R^{1,1}_1, R^{1,1}_2, R^{1,2}_3, R^{1,1}_4)$. By strategy-proofness (for student 4) we must have $f_4(R^{1,5}) \in \{s_4, 4\}$. Suppose that $f_4(R^{1,5}) = s_4$. Let $\tilde{R}_4 : s_3, s_4$ and $\tilde{R} = (R^{1,1}_1, R^{1,1}_2, R^{1,2}_3, \tilde{R}_4)$. If $f_4(\tilde{R}) = s_3$, 4 could manipulate $f$ at the profile $R^{1,4}$ by submitting $\tilde{R}_4$. Given that $f_4(R^{1,5}) = s_4$, strategy-proofness thus implies $f_4(\tilde{R}) = s_4$ as well. But this contradicts the constrained efficiency of $f$ since 4 and 2 would then form a SIC at $f(\tilde{R})$ and $\tilde{R}$.

This contradiction shows that we must have $f_4(R^{1,5}) = 4$ and $f(R^{1,5}) = f(R^{1,3})$ as well.

Let $R_4^{1,3} : s_4$ and $R^{1,6} = (R_1^{1,1}, R_2^{1,1}, R_3^{1,2}, R_4^{1,3})$. By strategy-proofness we must have $f_4(R^{1,6}) = 4$. Since the top choices of the other students are compatible, constrained efficiency implies that $f(R^{1,6}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_4 & s_1 & s_3 & 4 \end{pmatrix}$.

Let $R^{1,7} = (R_1^{1,1}, R_2^{1}, R_3^{1,2}, R_4^{1,3})$. By strategy-proofness we must have that $f_2(R^{1,7}) = s_1$. Stability implies that $f_1(R^{1,7}) = s_4$ and hence $f(R^{1,7}) = f(R^{1,6})$.

Let $R^{1,8} = (R_1^{1,1}, R_2^{1}, R_3^{1,2}, R_4^{1,1})$. By strategy-proofness and constrained efficiency we must have $f_4(R^{1,8}) = s_3$ and $f_3(R^{1,8}) = 3$. Thus $f(R^{1,8}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_4 & s_1 & 3 & s_3 \end{pmatrix}$.

Since $f_2(R^{1,5}) = s_3$ and $f_2(R^{1,8}) = s_1$, this implies that 2 is strictly better off reporting $R_2^1$ rather than her true preference $R_2^{1,1}$ when the other students submit $R_1^{1,1}, R_3^{1,2}$, and $R_4^{1,1}$. This contradicts strategy-proofness and completes the proof of Claim 1. □

Now let $R_1^2 : s_2, s_4$, $R_2^2 : s_4$, $R_3^2 : s_4, s_3, s_2$, and $R^2 = (R_1^2, R_2^2, R_3^2, R_4^1)$. Similar to above, there are exactly two constrained efficient matchings at $R^2$,

$$\mu^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_2 & 2 & s_4 & s_3 \end{pmatrix} \text{ and } \bar{\mu}^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & s_4 & s_2 & s_3 \end{pmatrix}.$$

**Claim 2:** $f(R^2) = \bar{\mu}^2$

**Proof of Claim 2.** The proof is analogous to the proof of Claim 1. One just needs to switch the roles of 1 and 2 as well as $s_1$ and $s_2$ and note that at any profile in the proof of Claim 1, school $s_3$ is never acceptable for both 1 and 2 (and the proof only uses the fact $\{1, 2\} \succ_{s_3} 4 \succ_{s_3} 3$ and not how students 1 and 2 are ranked under $\succ_{s_3}$). □

Now let $R_2^3 : s_4, s_3$, $R_4^2 : s_1, s_3, s_2$, and $R^3 = (R_1^1, R_2^3, R_3^1, R_4^2)$. Let $\mu^3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_4 & s_3 & s_1 & s_2 \end{pmatrix}$.

**Claim 3:** $f(R^3) = \mu^3$.

**Proof of Claim 3.** By Claim 1 we have that $f(R^1) = \mu^1$. Now let $R_2^{3,1} : s_1, s_4, s_3$ and $R^{3,1} = (R_1^1, R_2^{3,1}, R_3^1, R_4^1)$. Since $f_2(R^1) = 2$, strategy-proofness and stability imply $f_2(R^{3,1}) = s_3$. Constrained efficiency then implies $f_3(R^{3,1}) = s_1$ so that $f(R^{3,1}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_4 & s_3 & s_1 & 4 \end{pmatrix}$.

Now let $R_4^{3,1} : s_3, s_1$ and $R^{3,2} = (R_1^1, R_2^{3,1}, R_3^1, R_4^{3,1})$. By strategy-proofness, $f_4(R^{3,2}) \neq s_3$ so that constrained efficiency implies $f(R^{3,2}) = f(R^{3,1})$.

Let $R_4^{3,2} : s_1, s_3, s_2$ and $R^{3,3} = (R_1^1, R_2^{3,1}, R_3^1, R_4^{3,2})$. By strategy-proofness and $f_4(R^{3,2}) = 4$,

$f_4(R^{3,3}) \notin \{s_1, s_3\}$. Thus, by constrained efficiency, $f(R^{3,3}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_4 & s_3 & s_1 & s_2 \end{pmatrix}$.

Let $R_1^{3,1} : s_4, s_3$ and $R^{3,4} = (R_1^{3,1}, R_2^{3,1}, R_3^1, R_4^{3,2})$. By strategy-proofness, $f_1(R^{3,4}) = s_4$ so that constrained efficiency implies $f(R^{3,4}) = f(R^{3,3})$.

Let $R_2^{3,2} : s_4, s_3$ and $R^{3,5} = (R_1^{3,1}, R_2^{3,2}, R_3^1, R_4^{3,2})$. By strategy-proofness we must have $f_2(R^{3,5}) = s_3$. Since $1 \succ_{s_3} 2$, stability implies $f_1(R^{3,5}) = s_4$ and thus $f(R^{3,5}) = f(R^{3,3})$.

Let $R_1^{3,2} : s_4$ and $R^{3,6} = (R_1^{3,2}, R_2^{3,2}, R_3^1, R_4^{3,2})$. By strategy-proofness, $f_1(R^{3,6}) = s_4$ and hence $f(R^{3,6}) = f(R^{3,3})$. Since $R^{3,6} = R^3$ this proves Claim 3. $\square$

Let $R^4 = (R_1^1, R_2^3, R_3^2, R_4^2)$.

**Claim 4:** $f_3(R^4) \in \{s_3, s_4\}$.

**Proof.** By Claim 2 we have that $f(R^2) = \bar{\mu}^2$. Starting from profile $R^2$ we will consider the following preference profiles for these students:

| $R^{4,1}$ | $R_1^2$ | $R_2^{4,1}$ | $R_3^2$ | $R_4^2$ |
|---|---|---|---|---|
| | $s_2$ | $s_4$ | $s_4$ | $s_3$ |
| | $s_4$ | $s_3$ | $s_3$ | |
| | | | $s_2$ | |

| $R^2$ | $R_1^2$ | $R_2^2$ | $R_3^2$ | $R_4^2$ |
|---|---|---|---|---|
| | $s_2$ | $s_4$ | $s_4$ | $s_3$ |
| | $s_4$ | | $s_3$ | |
| | | | $s_2$ | |

| $R^{4,10}$ | $R_1^{4,2}$ | $R_2^{4,1}$ | $R_3^2$ | $R_4^{4,3}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_1$ |
| | | $s_3$ | $s_3$ | $s_3$ |
| | | | $s_2$ | $s_2$ |

| $R^{4,2}$ | $R_1^2$ | $R_2^{4,1}$ | $R_3^2$ | $R_4^{4,1}$ |
|---|---|---|---|---|
| | $s_2$ | $s_4$ | $s_4$ | $s_3$ |
| | $s_4$ | $s_3$ | $s_3$ | $s_1$ |
| | | | $s_2$ | |

| $R^{4,9}$ | $R_1^{4,1}$ | $R_2^{4,1}$ | $R_3^2$ | $R_4^{4,3}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_3$ | $s_1$ |
| | $s_3$ | $s_3$ | $s_4$ | $s_3$ |
| | | | $s_2$ | $s_2$ |

| $R^{4,3}$ | $R_1^{4,1}$ | $R_2^{4,1}$ | $R_3^2$ | $R_4^{4,1}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_3$ |
| | $s_3$ | $s_3$ | $s_3$ | $s_1$ |
| | | | $s_2$ | |

| $R^{4,7}$ | $R_1^{4,1}$ | $R_2^{4,2}$ | $R_3^2$ | $R_4^{4,1}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_3$ |
| | $s_3$ | $s_1$ | $s_3$ | $s_1$ |
| | | | $s_2$ | |

| $R^{4,8}$ | $R_1^{4,1}$ | $R_2^{4,2}$ | $R_3^2$ | $R_4^{4,3}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_1$ |
| | $s_3$ | $s_1$ | $s_3$ | $s_3$ |
| | | | $s_2$ | $s_2$ |

| $R^{4,4}$ | $R_1^{4,1}$ | $R_2^{4,1}$ | $R_3^{4,1}$ | $R_4^{4,1}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_3$ |
| | $s_3$ | $s_3$ | $s_3$ | $s_1$ |
| | | | $s_1$ | |

| $R^{4,5}$ | $R_1^{4,1}$ | $R_2^{4,1}$ | $R_3^{4,1}$ | $R_4^{4,2}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_1$ |
| | $s_3$ | $s_3$ | $s_3$ | $s_3$ |
| | | | $s_1$ | $s_2$ |

| $R^{4,6}$ | $R_1^{4,2}$ | $R_2^{4,1}$ | $R_3^{4,1}$ | $R_4^{4,2}$ |
|---|---|---|---|---|
| | $s_4$ | $s_4$ | $s_4$ | $s_1$ |
| | | $s_3$ | $s_3$ | $s_3$ |
| | | | $s_1$ | $s_2$ |

Let $R_2^{4,1} : s_4, s_3$ and $R^{4,1} = (R_1^2, R_2^{4,1}, R_3^2, R_4^2)$. By strategy-proofness, $f_2(R^{4,1}) = s_4$ so that constrained efficiency implies $f(R^{4,1}) = \bar{\mu}^2$

Let $R_4^{4,1} : s_3, s_1$ and $R^{4,2} = (R_1^2, R_2^{4,1}, R_3^2, R_4^{4,1})$. By strategy-proofness, $f_4(R^{4,2}) = s_3$ so that, by stability, $f_2(R^{4,2}) = s_4$ and $f(R^{4,2}) = \bar{\mu}^2$.

Let $R_1^{4,1} : s_4, s_3$ and $R^{4,3} = (R_1^{4,1}, R_2^{4,1}, R_3^2, R_4^{4,1})$. By strategy-proofness, stability, and $1 \succ_{s_3} 2 \succ_{s_3} 4 \succ_{s_3} 3$, we must have $f_1(R^{4,3}) = s_3$. Constrained efficiency then implies that either $f_2(R^{4,3}) = s_4$ or $f_3(R^{4,3}) = s_4$. We show by contradiction that the second case is impossible. Suppose $f_3(R^{4,3}) = s_4$. Let $R_3^{4,1} : s_4, s_3, s_1$ and $R^{4,4} = (R_1^{4,1}, R_2^{4,1}, R_3^{4,1}, R_4^{4,1})$. By strategy-proofness $f_3(R^{4,4}) = s_4$ so that in particular $f_1(R^{4,4}) = s_3$ and $f_4(R^{4,4}) = s_1$. Let $R_4^{4,2} : s_1, s_3, s_2$

and $R^{4,5} = (R_1^{4,1}, R_2^{4,1}, R_3^{4,1}, R_4^{4,2})$. By strategy-proofness we must have $f_4(R^{4,5}) = s_1$. This is compatible with stability only if $f_3(R^{4,5}) = s_4$ and $f_1(R^{4,5}) = s_3$ so that $f(R^{4,5}) = f(R^{4,4})$. Let $R_1^{4,2} : s_4$ and $R^{4,6} = (R_1^{4,2}, R_2^{4,1}, R_3^{4,1}, R_4^{4,2})$. By strategy-proofness $f_1(R^{4,6}) = 1$. But note that $R^{4,6} = R^3$ and, by Claim 3 above, $f_1(R^3) = s_4$. This is a contradiction and hence we must have that $f_2(R^{4,3}) = s_4$. Constrained efficiency then implies $f(R^{4,3}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ s_3 & s_4 & s_2 & s_1 \end{pmatrix}$.

Let $R_2^{4,2} : s_4, s_1$ and $R^{4,7} = (R_1^{4,1}, R_2^{4,2}, R_3^2, R_4^{4,1})$. By strategy-proofness, $f_2(R^{4,7}) = s_4$ so that by stability $f(R^{4,7}) = f(R^{4,3})$.

Let $R_4^{4,3} : s_4, s_3, s_2$ and $R^{4,8} = (R_1^{4,1}, R_2^{4,2}, R_3^2, R_4^{4,3})$. By strategy-proofness, $f_4(R^{4,8}) = s_1$. Stability then implies $f_2(R^{4,8}) = s_4$ and $f(R^{4,8}) = f(R^{4,3})$.

Let $R^{4,9} = (R_1^{4,1}, R_2^{4,1}, R_3^2, R_4^{4,3})$. Strategy-proofness implies $f_2(R^{4,9}) = s_4$ so that, by stability, $f(R^{4,9}) = f(R^{4,3})$.

Let $R^{4,10} = (R_1^{4,2}, R_2^{4,1}, R_3^2, R_4^{4,3})$ and note that $R^{4,10} = R^4$. By strategy-proofness $f_1(R^4) = 1$. But then constrained efficiency implies $f_3(R^4) \in \{s_4, s_3\}$ since either $f_2(R^4) = s_4$ or $f_3(R^4) = s_4$, and if $f_2(R^4) = s_4$ then $f_3(R^4) = s_3$. □

Combining Claim 3 and Claim 4 we see that student 3 has an incentive to submit $R_3^2$ when other students submit their preferences from the profile $R^3$ since $f_3(R^4) P_3^1 f_3(R^3) = s_1$ given that $f_3(R^3) = s_1$ by Claim 3 and $f_3(R^4) \in \{s_3, s_4\}$ by Claim 4. This contradicts strategy-proofness of $f$ and completes the proof. □

## Proof of Proposition 7

We show first that no ambiguos 1-ties and no ambiguity at the top imply that **O1** and **O2** are satisfied. Fix some $k \leq K$ and $i \in L_k$. Suppose to the contrary that there exists a specialized school $s_1 \in S^1$ such that $r_{k+l}(\succ_{s_1}) = i$ for some $l \geq 3$. On the other hand since $i \in L_k$ there exists a specialized school $s_2 \in S^1 \setminus \{s_1\}$ such that $r_k(\succ_{s_2}) = i$. For any $l' \in \{1, \ldots, l\}$ let $i_{l'} = r_{k+l'}(\succ_{s_2})$. Now if there is a student $j$ such that $i_l \succ_{s_1} j \succ_{s_1} i$ we have found an ambiguous 1-tie since $l \geq 3$ so that at least two distinct students rank between $i$ and $i_l$ with respect to $\succ_{s_2}$. If $i_l = r_{k+l-1}(\succ_{s_1})$ and $i_{l-1} \succ_{s_1} i_l$ we obtain a contradiction since $i \succ_{s_2} i_{l-2} \succ_{s_2} i_{l-1}$. If $i_l \succ_{s_1} i \succ_{s_1} i_{l-1}$, there has to be a student $j$ such that $i_l \succ_{s_2} j$ and $j \succ_{s_1} i_l$ yielding another contradiction. Hence, we must have $i \succ_{s_1} i_l$. But then there has to exist a student $j$ such that $i_l \succ_{s_2} j$ and $j \succ_{s_1} i$. No matter whether $j \succ_{s_1} i_{l-1}$ or $i_{l-1} \succ_{s_1} j$ we obtain an ambiguous 1-tie. This shows that **O1** has to be satisfied.

To see that **O2** must be satisfied note that if **O1** is satisfied we must have $|L_1| = 2$ if $|I| > 3$: If $|L_1| = 1$ we must have $|I| = 1$ and if $|L_1| = 3$ we must have $I = L_1$ since otherwise one of the students in $L_1$ would have to rank fourth at some specialized school. By similar arguments we must have $|L_2| \in \{1, 2\}$ and $K = 2$ if $|L_2| = 2$. Let $L_1 = \{1, 2\}$. If $L_2 = \{3, 4\}$ but there exist two specialized schools $s_1, s_2 \in S^1$ such that $r_3(\succ_{s_1}) = 1$ and $r_3(\succ_{s_2}) = 2$ we obtain an ambiguous 1-tie. If $L_2 = \{3\}$, let 4 be one of the students in $L_3$. Suppose there exist two specialized schools $s_1, s_2 \in S^1$ such that $r_3(\succ_{s_1}) = 1$ and $r_3(\succ_{s_1}) = 2$. By **O1** we must have $2 \succ_{s_1} 3 \succ_{s_1} 1 \succ_{s_1} 4$ and $1 \succ_{s_2} 3 \succ_{s_2} 2 \succ_{s_2} 4$. Now since $4 \in L_3$ and $|L_1 \cup L_2| = 3$, at least one agent in $\{1, 2, 3\}$ must have lower priority than 4 for some specialized school. By **O1** this agent cannot be 1 or 2. This implies that there exists a third specialized school $s_3$ such that $\{1, 2\} \succ_{s_3} 4 \succ_{s_3} 3$. Hence, $\succ^1$ contains ambiguity at the top (with respect to schools $s_1, s_2, s_3$).

Now suppose that $\succ^1$ satisfies **O1** and **O2**. Note that if $|I| > 3$, we must have $|L_1| = 2$, $|L_k| = 1$, for all $k \in \{1, \ldots, K-1\}$, and $|L_K| \in \{1, 2\}$. Now by **O2** there cannot be an ambiguous 1-tie involving the two students in $L_1$ since at most one of them can rank third. By **O1** there cannot be an ambiguous 1-tie between an agent $i \in L_k$ and an agent $j \in L_{k'}$ for $k < k' \leq K$ since $i$ always has at least $(k+2)$nd highest priority at specialized schools. Lastly, there cannot be an ambiguous 1-tie between two students in $L_K$ (if $|L_K| = 2$) since only one student in $L_{K-1}$ can rank in between these two agents by **O1** and **O2**.

Now suppose that there exist four distinct students $i_1, i_2, i_3, i_4$ and three specialized schools $s_1, s_2, s_3$ such that $i_1 \succ_{s_1} i_3 \succ_{s_1} i_2 \succ_{s_1} i_4$, $i_2 \succ_{s_2} i_3 \succ_{s_2} i_1 \succ_{s_2} i_4$, and $\{i_1, i_2\} \succ_{s_3} i_4 \succ_{s_3} i_3$. Clearly, we cannot have $i_1, i_2 \in L_K$ since this would imply $i_4 \in L_1 \cup \ldots \cup L_{K-1}$ while $i_4 \notin \{r_1(\succ_{s_1}), \ldots, r_{K+2}(\succ_{s_1})\}$, contradicting **O1**. By **O1**, we can also not have that $i_1, i_2 \in L_1$ and it is easy to see that $i_1$ and $i_2$ cannot belong to different $L_k$s. This completes the proof.

□

## Proof of Theorem 2

(i) Fix an arbitrary school choice problem $R$ and let $\mu := f^{ETB}(R)$ be the matching produced by the SDA-ETB algorithm. Let $(\mu^t)_{t \geq 1}$ be the sequence of temporary assignments in the SDA-ETB. We show that there are no stable improvement cycles (SICs) at $\mu$ and $R$ by contradiction.

Suppose that $i_1, \ldots, i_m$ is a SIC at $\mu$ and $R$, and let $s_l := \mu(i_l)$ for all $l \leq m$. We assume in the following that the cycle is minimal in the sense that no strict subset of students $i_1, \ldots, i_m$ forms a SIC. Note that since $s_{l+1} P_{i_l} s_l$, $i_l$ must have applied to $s_{l+1}$ before applying to $s_l$ in the SDA-ETB. We start with a few preliminary observations about SICs that are summarized in the following Lemma.

**Lemma 1.** *(i) $s_l \neq s_{l'}$ for all $l \neq l'$.*

*(ii) $s_l \in S^0$ for at least one $l \leq m$.*

*(iii) $s_l \in S^1$ for at least one $l \leq m$.*

**Proof.**

(i) It is clear that no specialized school can appear more than once on a SIC since $D_s(\mu)$ contains at most one student if $s \in S^1$. If $s_l = s_{l'} = s \in S^0$, $i_1, \ldots, i_{l-1}, i_{l'}, \ldots, i_m$ is a SIC of smaller size which contradicts the assumed minimality of the cycle.

(ii) Suppose to the contrary that $\{s_1, \ldots, s_m\} \subseteq S^1$ and let $t_1$ be the first round of the SDA-ETB (under $R$)[31] in which a student $i_l$ is rejected by $s_{l+1}$. But then, there must be a student $j \in \mu^{t_1}(s_{l+1}) \setminus \mu(s_{l+1})$ such that $j \succ_{s_{l+1}} i_l$. Since $s_{l+1} \in S^1$, this implies that $i_l \notin D_{s_{l+1}}(\mu)$; contradiction.

(iii) Suppose to the contrary that $\{s_1, \ldots, s_m\} \subseteq S^0$.

Consider first the case $|I| \leq p$ and note that $m \geq 2$. Again, let $t_1$ be the first round of the SDA-ETB in which a student $i_l$ is rejected by $s_{l+1}$. If there was a specialized school $s \in S^1$ such that $|\mu^{t_1}(s)| = q^1$, there could not have been a round $t' > t_1$ in which a student $i_{l'} \neq i_l$ was rejected by $s_{l'+1}$ given $|I| \leq p$; contradiction. Now let $t'' > t_1$ be some round of the SDA-ETB in which a student $i_{l'} \neq i_l$ is rejected by $s_{l'+1}$. Since $|\mu^{t''}(s_{l+1})| \geq q_{s_{l+1}}$ and $|\mu^{t''}(s_{l'+1})| > q_{s_{l'}+1}$, there could not have been a specialized school $s \in S^1$ such that $|\mu^{t''}(s)| = q_s$ given that $|I| \leq p$. This implies that all rejections by non-specialized schools on the SIC were based on the labeling of students so that $i_1 > i_2 > \ldots > i_m > i_1$[32]; contradiction.

If $|I| > p$ note that the construction of $\succeq^0$ and the just completed argument imply that no student on the SIC belongs to the upper segment $L_1 \cup \ldots \cup L_{p-2}$. Furthermore, the cycle cannot consist exclusively of students who are strictly ordered according to $\succeq^0$. This implies that the only remaining possibility for a SIC consisting only of non-specialized schools is $m = 2$ and $i_1, i_2 \in L_K$. The proof is completed by noting that, by construction of $\succeq^0$, no student in $L_K$ can cause a student in $L_k$ to be rejected by a non-specialized school for all $k < K$.

[31] This qualifying statement will henceforth be omitted and we will speak of the SDA-ETB. There is no ambiguity involved here since the problem $R$ is fixed throughout.

[32] Remember that we identify students with their labels.

□

Now consider the case of $|I| \leq p$. By Lemma 1 we can assume w.l.o.g. that $s_1 \in S^0$ and $s_2 \in S^1$. Since $|I| \leq p$, we must have $m \leq 3$. Suppose first that $m = 3$ and that $s_3 \in S^1$ so that $s_1$ is the only non-specialized school on the SIC. Note that $i_3$ must have been rejected by $s_1$ before $i_2$ was rejected by $s_3$: Otherwise $i_2 \notin D_{s_3}(\mu)$ since at least one higher priority student must have been rejected by $s_3$ in the course of SDA-ETB. Similarly, $i_2$ must have been rejected by $s_3$ before $i_1$ was rejected by $s_2$. But this implies that at the point where $i_1$ was supposedly rejected by $s_2$, at least $q^0_{(1)}$ students were temporarily matched to $s_1$ and $q^1$ students were temporarily matched to $s_3$. Since $|I| \leq p$, $i_1$ could not have been rejected by $s_2$; contradiction. Now suppose that $s_3 \in S^0$. As in the previous case, $i_2$ must have been rejected by $s_3$ before $i_1$ was rejected by $s_2$. This implies that there is a round $t$ of SDA-ETB such that $i_2 \in \mu^t(s_3)$, $\mu^t(i_1)R_{i_1}s_2$, and $s_3P_{i_2}\mu^{t+1}(i_2)$. Now in some round $t' > t$, $i_1$ must have been rejected by $s_2$. If $s_1P_{i_3}\mu^{t'}(i_3)$, we obtain an immediate contradiction since at the point were $i_1$ was supposedly rejected by $s_2$, at least $q^0_{(1)} + q^0_{(2)}$ students must have been matched to $s_1$ and $s_3$ so that there could not have been $q^1$ students apart from $i_1$ applying to $s_2$ in round $t'$. If $\mu^{t'}(i_3)R_{i_3}s_1$, we similarly obtain a contradiction to the assumption that $i_3$ was rejected by $s_1$ in some later round of SDA-ETB. Hence, we must have $m = 2$. As in the previous cases, $i_2$ must have been rejected by $s_1$ before $i_1$ was rejected by $s_2$. This implies that there is a round $t$ of the SDA-ETB such that $i_2 \in \mu^t(s_1)$, $\mu^t(i_1)R_{i_1}s_2$, and $s_1P_{i_2}\mu^{t+1}(i_2)$. If $\mu^t(i_1)P_{i_1}s_2$, $i_1$ could not have been rejected by $\mu^t(i_1)$ and $s_2$ in subsequent rounds of SDA-ETB given that $|I| \leq p$. Hence, we must have $\mu^t(i_1) = s_2$ by strict preferences. If $i_1 \notin \mu^{t+1}(s_2)$, it has to be the case that $\mu^t(s_1) \succ_{s_2} i_1$ since $i_2$ would not have been rejected in $TB(\mu^t)$ otherwise. But then $i_1$ could not have subsequently obtained a place at $s_1$; contradiction. If $i_1 \in \mu^{t+1}(s_2)$, there must be a student $j \in \mu(s_2)$ such that $\mu^t(j)P_js_2$. If $\mu^t(j) \neq s_1$, let $t'$ be the round where $j$ was rejected by $\mu^t(j)$. Given $|I| \leq p$, there cannot be a round $t'' > t'$ in which $i_1$ was rejected by $s_2$. But then $i_2$ must have been rejected by $s_2$ before round $t'$ so that in particular $i_2 \notin D_{s_2}(\mu)$; contradiction. Hence, we must have $\mu^t(j) = s_1$ for any $j \in \mu(s_2)$ such that $\mu^t(j)P_js_2$. Iterating this argument it is easy to see that there must be a round $t' > t$ of SDA-ETB such that $|\mu^{t'}(s_2)| = q^1$, $i_1 \in \mu^{t'}(s_2)$, $|\mu^{t'}(s_1)| \geq q_{s_1} + 1$, and $\mu^{t'}(s_1) \succ_{s_2} i_1$. But then $i_1$ could not have obtained a place at $s_1$ subsequently to being rejected by $s_2$; contradiction.

Now we consider the case that $|I| > p$. Note that we can assume w.l.o.g. that $i_1$ is the student with the lowest label among all students on the SIC. We distinguish two subcases.

**Case 1:** $i_1 \geq p$.

We assume for now that $|L_K| = 2$. It will become clear from our arguments that there cannot be SIC in case $|L_K| = 1$ either.

If $s_2 \in S^0$, we must have $i_1, i_2 \in L_K$ by the construction of $\succeq^0$ and the assumption that $p \leq i_1 < i_2$. This implies in particular that $L = 2$, $i_1 = K+1$, $i_2 = K+2$, and $s_1 \in S^1$. By exogenous tie-breaking in the SDA-ETB at most $q_{s_2} - 1$ students with lower labels than $K+1$ could have applied to $s_2$. If more than $q^1 - 1$ students with lower labels than $K+1$ applied to $s_1$, stability of $\mu$ and limited p-variability imply that $K$ must have been rejected by $s_1$ in the SDA-ETB and $K+1 \succ_{s_1} K \succ_{s_1} K+2$. But then we cannot have that $K+2 \in D_{s_1}(\mu)$. Hence, at most $q^1 - 1$ students with lower labels than $K+1$ could have applied to $s_1$. But then SDA-ETB would have assigned $K+1$ to $s_2$ and $K+2$ to $s_1$; contradiction. Hence, we must have $s_2 \in S^1$.

Now suppose that $s_2 \in S^1$ and $i_2 = i_1 + 2$. Since $s_2 \in S^1$, the stability of $\mu$ w.r.t. $\succ^1$ and limited p-variability imply that $i_1 = K$, $i_2 = K+2$, and $K+2 \succ_{s_2} K$. If $m = 2$ we must have $s_1 \in S^0$ by Lemma 1. Given exogenous tie-breaking in the SDA-ETB and $\mu(K) = s_1$, at most $q_{s_1} - 1$ students with labels lower than $K$ could have applied to $s_1$. As above, $K+2$ must have been rejected by $s_1$ before $K$ was rejected by $s_2$. It is easy to see that we obtain a contradiction unless $K+1$ applied to $s_1$ in the SDA-ETB. Since ties in the lower segment are broken last, this implies that there is a round $t$ of the SDA-ETB procedure such that $\{K+1, K+2\} \subset \mu^t(s_1)$ and $i_1 \in \mu^t(s_2)$. Since $K+2 \succ_{s_2} K$ by the stability of $\mu$, we must have $K \succ_{s_2} K+1$ by limited p-variability. But then $K+1$, and not $K+2$, would have been rejected by $s_1$. Since there are no further rejections after tie-breaking in the lower segment this is a contradiction. If $m = 3$, $i_3 = K+1$ (this is the only possibility given the definition of $i_1$), and $s_3 \in S^0$, $i_2$ must have been rejected by $s_3$ before $i_1$ was rejected by $s_2$. Suppose first that $s_1 \in S^0$. By exogenous tie-breaking in the SDA-ETB and limited p-variability at most $q_{s_1} - 1$ students with lower labels than $K$ could have applied to $s_1$. Similarly, at most $q_{s_3} - 1$ lower labeled students could have applied to $s_3$. Furthermore, it cannot be the case that $s_1 P_{K+2} s_3$ since there would be a SIC of size 2 otherwise. But then neither $i_3$ nor $i_2$ would have been rejected by $s_1$ and $s_3$, respectively; contradiction. If $s_1 \in S^1$, it is easy to see that there must have been a round $t$ of the SDA-ETB such that $\{i_2, i_3\} \subset \mu^t(s_3)$ and $i_1 \in \mu^t(s_2)$. But then, $i_2$ would not have been rejected by $s_3$ since $i_2 \succ_{s_2} i_1 \succ_{s_2} i_3$ given the stability of $\mu$ and limited p-variability. An analogous argument can be used to show that $m = 3$, $i_3 = K+1$, and $s_3 \in S^1$ is also impossible. Hence, we must have $i_2 = i_1 + 1$.

Now suppose we have shown that, for some $l \leq m$, $i_{l'} = i_{l'-1} + 1$ and $s_{l'} \in S^1$, for all $l' \in \{2, \ldots, l\}$. We now establish that $m > l$, $i_{l+1} = i_l + 1$, and $s_{l+1} \in S^1$. This inductive argument completes the proof since it contradicts the finiteness of the set of students.

Suppose first that $m = l$. Since there has to be at least one non-specialized school on the SIC by Lemma 1, we must have $s_1 \in S^0$. Note that it has to be the case that $i_m \in \{K+1, K+2\}$. Otherwise we could use exactly the same argument used to establish that a SIC cannot contain only specialized schools to derive a contradiction since we assumed $i_1 \geq p-1$. Suppose first that $i_m = K+1$ so that $i_{m-1} = K$. By minimality of the cycle, we must have $s_{l+1} P_{i_l} s_1$ for all $l < m$. Furthermore, at most $q_{s_1} - 1$ students indexed lower than $i_1$ could have applied to $s_1$ by exogenous tie-breaking in the SDA-ETB. Since all schools except $s_1$ are specialized, $i_m$ must have been rejected by $s_1$ before $i_l$ was rejected by $s_{l+1}$ *for all* $l \leq m-1$. Similar to above this implies that there must have been a round $t$ of SDA-ETB such that $\{K+1, K+2\} \subset \mu^t(s_1)$ and $K \in \mu^t(s_m)$. By stability of $\mu$ and limited p-variability, we must have $K+1 \succ_{s_m} K \succ_{s_m} K+2$. This implies again that $K+2$, and not $K+1$, would have been rejected by $s_1$. Since there are no rejections after tie-breaking in the lower segment this is a contradiction. If $i_m = K+2$, we obtain an immediate contradiction since it is easy to see that the minimality of the cycle implies that at most $q_{s_1} - 1$ other students could have applied to $s_1$ in the SDA-ETB. Hence, $i_m$ could not have been rejected by $s_1$ in the SDA-ETB; contradiction.

The proofs that $s_{l+1} \in S^1$ and $i_{l+1} = i_l + 1$ are virtually identical to the proofs that $s_2 \in S^1$ and $i_2 = i_1 + 1$. The details are omitted.

**Case 2:** $i_1 < p$.

Note first that there has to exist an $l \leq m-1$ such that $i_l = p-1$ and $i_{l+1} \in L_{p-1} \cup \ldots \cup L_K$. Otherwise, the SIC would consist entirely of students in $L_1 \cup \ldots \cup L_{p-2}$. The proof that this is impossible is completely analogous to the proof for the case of $|I| \leq p$.

Hence, there has to exist an index $l$ with the above mentioned properties. Note that in particular $s_{l+1} \in S^1$ since $i_l$ cannot envy a student in $L_{p-1}$ for a non-specialized school. Furthermore, by limited p-variability $i_{l+1}$ is the only student in $L_{p-1} \cup \ldots \cup L_K$ who can rank higher at $s_{l+1}$ than $i_l$. This implies $|\mu(s_{l+1}) \cap (L_1 \cup \ldots \cup L_{p-2})| = q^1 - 1$. For all $l' \in \{2, \ldots, l\}$ we must have $\mu(s_l) \subset L_1 \cup \ldots \cup L_{p-1}$ and $|\mu(s_l)| = q_{s_l}$. Now note that $|I| > p$ implies $|L_1 \cup \ldots \cup L_{p-2}| = p - 1$ so that we must have $l \leq 2$.

Now let $l \leq 2$, $i_l = p - 1$ and $i_{l+1} \in L_{p-1}$. Note that there cannot be an $l < l' \leq m$ such that $i_{l'} \in L_1 \cup \ldots \cup L_{p-2}$: Otherwise, we would have that $\{i_{l'}, \ldots, i_m\} \subset L_1 \cup \ldots \cup L_{p-2}$. If $l = 1$, this yields a contradiction to the assumption that $i_1$ was the lowest labeled agent on the SIC. If $l = 2$, we would have that $\mu(s_{l'+1}) \cup \ldots \cup \mu(s_1) \subset L_1 \cup \ldots \cup L_{p-2}$. Since $|\mu(s_{l+1}) \cap (L_1 \cup \ldots \cup L_{p-2})| = q^1 - 1$, this contradicts $|L_1 \cup \ldots \cup L_{p-2}| = p - 1$. By construction of $\succeq^0$ and limited p-variability, we must thus have $i_{l'+1} = i_{l'} + 1$ for all $l' \in \{l, \ldots, m-3\}$, and $\{s_{l+1}, \ldots, s_{m-1}\} \subset S^1$. If $s_m \in S^0$, it has to be the case that $\{i_{m-1}, i_m\} = L_K$. As above, $i_{m-1}$ must have been rejected by $s_m$ before $i_{m-2} = K$ was rejected by $s_{m-1} \in S^1$. Since we break ties in the lower segment last and $i_m \in \mu(s_m)$, there must have been a round $t$ of the SDA-ETB such that $\{i_{m-1}, i_m\} \subset \mu^t(s_m)$, $|\mu^t(s_m)| = q_{s_m} + 1$, and $i_{m-2} \in \mu^t(s_{m-1})$. But then $i_{m-1}$ could not have been rejected by $s_m$; contradiction. Hence, we must have $\{s_{l+1}, \ldots, s_m\} \subset S^1$.

Now suppose that $l = 2$, $s_1 \in S^1$, and $s_2 \in S^0$. It has to be the case that $i_1$ was rejected by $s_2$ before $i_l$ was rejected by $s_{l+1}$ for all $l \in \{2, \ldots, L\}$. Let $t$ be the round of SDA-ETB in which $i_1$ was rejected by $s_2$ and let $t' > t$ be the round of SDA-ETB in which $i_2$ applied to $s_2$. By limited p-variability we must have $|\mu^{t'}(s_3) \cap (L_1 \cup \ldots \cup L_{p-2})| \geq q^1 - 1$ and since $t' > t$ it has to be the case that $\mu^{t'}(s_2) \setminus \{i_2\} \subset L_1 \cup \ldots \cup L_{p-2}$, $|\mu^{t'}(s_2)| \geq q^0_{(1)} + 1$, and $i_1 \notin \mu^{t'}(s_3) \cup \mu^{t'}(s_2)$. But since $|L_1 \cup \ldots \cup L_{p-2}| = p - 1$, there cannot be a specialized school $s \in S^1 \setminus \{s_3\}$ such that $\mu^{t'}(s) \subset L_1 \cup \ldots \cup L_{p-2}$ and $|\mu^{t'}(s)| = q^1$. Since $i_2$ is the highest indexed student in the upper segment, she could not have obtained a place at $s_2$; contradiction. Next, consider the case $s_1 \in S^0$, $s_2 \in S^1$. As above, $i_m$ must have been rejected by $s_1$ before $i_l$ was rejected by $s_{l+1}$ for all $l \leq m - 1$. Let $t$ be the round of the SDA-ETB in which $i_m$ was rejected by $s_1$. If $\mu^t(s_1) \subset L_1 \cup \ldots \cup L_{p-2}$, it is easy to see that $i_1$ could not have been rejected by $s_2$ given that $|L_1 \cup \ldots \cup L_{p-2}| = p - 1$ since $i_2$ must have been rejected by $s_3$ in some earlier round. By minimality of the cycle, we must have $s_l P_{i_l} s_1$ for all $l \neq m$. This implies that there is an agent $j \in I \setminus (L_1 \cup \ldots \cup L_{p-2} \cup \{i_1, \ldots, i_m\})$ such that

$\mu^{t+1}(j) = \mu^t(j) = s_1$. Given the above the only possibility is that $\{i_m, j\} = L_K$. Since ties in the lower segment are broken last, we must have $\mu^t(i_{m-1}) = s_m$. As usual, stability and limited p-variability imply that $i_m$ would not have been rejected by $s_1$; contradiction. The only remaining case to consider for $l = 2$ is $s_1, s_2 \in S^0$. Here, $i_m$ must have been rejected by $s_1$ in particular before $i_2$ was rejected by $s_3$. If $i_1$ was rejected by $s_2$ before $i_2$ was rejected by $s_3$, $i_2$ could not have been rejected by $s_3$ in a subsequent round. If $i_2$ was rejected by $s_3$ before $i_1$ was rejected by $s_2$, $i_1$ could not have been rejected by $s_2$ in a subsequent round. The details are similar to the previous cases and omitted.

Thus, the only remaining case is $l = 1$ and $s_1 \in S^0$ (by Lemma 1). As usual, $i_m$ must have been rejected by $s_1$ before $i_l$ was rejected by $s_{l+1}$ for all $l \leq m-1$. Similar to the proof that $l = 2, s_1 \in S^0, s_2 \in S^1$, we can show that if $t$ is the round in which $i_m$ was rejected by $s_1$ we must have $\mu^t(s_1) \setminus \{i_m\} \subset L_1 \cup \ldots \cup L_{p-2}$. If $t' > t$ is the round in which $i_1$ is rejected by $s_2$, we must have $|\mu^{t'}(s_2) \cap (L_1 \cup \ldots \cup L_{p-2})| = q^1 - 1$. Since $|I| \leq p-1$, this implies that if $t'' > t'$ denotes the round of SDA-ETB in which $i_2$ applies to $s_1$, we must have $\mu^{t''}(s_1) \subset L_1 \cup \ldots \cup L_{p-2}$ and there could not have been a specialized school $s \neq s_2$ such that $\mu^{t''}(s_2) \subset L_1 \cup \ldots \cup L_{p-2}$ and $|\mu^{t''}(s_2)| = q^1$. But then $i_2$ could not have obtained a place at $s_2$; contradiction.

(ii) We first consider the case $|I| \leq p$. Let $i \in I$ be a student and $R_i$ be an arbitrary strict preference relation for this student. Let $top_j(R_i)$ be the $j$th most preferred school according to $R_i$ if $R_i$ contains at least $j$ acceptable schools, and $top_j(R_i) = i$ if $R_i$ contains less than $j$ acceptable schools.

Note that if $f_i^{ETB}(R) \notin \{top(R_i), top_2(R_i), top_3(R_i)\}$ for some $i \in I$, student $i$ cannot manipulate at the profile $R$:[33] Let $s_1 = top(R_i), s_2 = top_2(R_i)$, and $s_3 = top_3(R_i)$. Given that $|I| \leq p$, we must have $s_1, s_2, s_3 \in S^0$ and any specialized school $s$ must have at least $q_{s_1} + q_{s_2} + q_{s_3}$ free places in the matching $f^{ETB}(R)$. But then, no matter which preference relation $i$ submits, no specialized school can ever fill its capacity. By the rules of the tie-breaking subroutine, tie-breaking decisions will thus always be based on the (fixed) labels of students. By strategy-proofness of the SDA for fixed tie-breaking rules, $i$ cannot manipulate the SDA-ETB.

Now consider a profile $R$ and a student $i$ such that $f_i^{ETB}(R) \in \{top_2(R_i), top_3(R_i)\}$.

[33] It is obvious that the SDA-ETB never assigns a student to an unacceptable school. This implies in particular that $\{top(R_i), top_2(R_i), top_3(R_i)\} \subset S$ in the above situation.

Suppose there was an alternative report $R_i'$ for $i$ such that $f_i^{ETB}(R_i', R_{-i}) = top(R_i) =: s_1$. Let $s_2 := top(R_i')$ and note that we must have $s_1 \neq s_2$. Denote by $(\mu^t)_{t\geq 1}$ and $(\tilde{\mu}^t)_{t\geq 1}$ the sequences of temporary assignments of the SDA-ETB under $R$ and $R'$, respectively. Let $t_1$ be the round of the SDA-ETB under $R$ in which $i$ is rejected by $s_1$ and let $t_2$ be the round of the SDA-ETB under $R'$ in which $i$ is rejected by $s_2$. Consider first the case $s_1 \in S^0$ and $s_2 \in S^1$. We must have $|\mu^{t_1}(s_1)| = q_{s_1} + l$ for some $l \geq 1$. If there was a specialized school $s$ such that $|\mu^{t_1}(s)| = q_s$, it would have to be the case that $|\mu^t(s_2)| \leq q_{s_2} - l$. But this would imply that in the SDA-ETB under $R$ no student is ever rejected by $s_2$, so that $f_i^{ETB}(R') = s_2$. Continuing this line of reasoning it is easy to see that all tie-breaking decisions in the SDA-ETB under $R$ must have been made conditional on the fixed labeling of students and no specialized school could have rejected any student in the course of this algorithm. But the same statements must hold for the SDA-ETB under $R'$ since $i$ was rejected by $s_1$ in the SDA-ETB under $R$. Now there must be at least $q_{s_1}$ students with lower labels than $i$ who applied to $s_1$ in the SDA-ETB under $R$. But all of these students will apply to $s_1$ in the SDA-ETB under $R'$ given the above so that $i$ cannot obtain a place at $s_1$. Next, consider the case $s_1 \in S^0$ and $s_2 \in S^1$. Note that if $t_2 = 1$, $i$ cannot end up matched to $s_1$ in the SDA-ETB under $R'$ if $top_2(R_i') = s_1$ since there must be a round of this procedure in which the temporary assignment is exactly the same as in round $t_1$ of the SDA-ETB under $R$. If $top_2(R_i') \neq s_1$, we must have $f_i^{ETB}(R') = top_2(R_i')$ given $|I| \leq p$. Thus, $t_2 > 1$ and there has to be a school $s \neq s_2$ that has to reject a student in some round $t < t_2$ of the SDA-ETB under $R'$. If $s \neq s_1$, $s_1$ could not have rejected any student in the SDA-ETB under $R$ and $R'$ given that $|I| \leq p$ which contradicts $f_i^{ETB}(R) \neq s_1$. Hence, $s = s_1$ and we must have $|\tilde{\mu}^{t_2}(s_1)| > q_{s_1}$. But then subroutine $\mathbf{TB}(\tilde{\mu}^{t_2})$ ensures that all students in $\tilde{\mu}^{t_2}(s_1)$ have higher priority for $s_2$ than $i$ and that $i$ could not have obtained a place at $s_1$ in one of the subsequent rounds. Now suppose $s_1 \in S^1$ and note that $i$ could not obtain $s_1$ by any misrepresentation if $\mu^1(i) \neq s_1$. Hence, there must be school $s \neq s_1$ that had to reject at least one student prior to $t_1$. Since $|I| \leq p$, $i$ is matched to $top(R_i')$ if $top(R_i') \in S \setminus \{s_1, s\}$. Hence, we must have $s = s_2$ and $s_1, s_2$ are the only schools who had to reject a student in the SDA-ETB under $R'$. If $s_2 \in S^1$ and $f_i^{ETB}(R') = s_1$, $i$ could also manipulate if $s_1$ and $s_2$ were the only schools available. This contradicts strategy-proofness of the SDA when $S^0 = \emptyset$. So suppose that $s_2 \in S^0$. Since $i$ was rejected subsequently to a rejection at $s_2$ in the SDA-ETB under $R$, all students in $\mu^{t_1}(s_2)$ must have had higher priority for $s_1$ than $i$. This implies that in the SDA-ETB under $R'$ ultimately the same set of students will be rejected by $s_2$ as in the SDA-ETB under $R$. Since all students who have applied to $s_1$ prior to $t_1$ in the SDA-ETB under $R$ also apply to $s_1$ in some round of the SDA-ETB under $R'$, we cannot have $f_i^{ETB}(R') = s_1$.

It remains to be shown that $i$ cannot obtain $top_2(R_i)$ if $f_i^{ETB}(R) = top_3(R_i)$. Given $|I| \leq p$, it is easy to see that the only potentially profitable manipulation is for $i$ to rank $top_2(R_i)$ first. The proof can be completed using a similar case distinction as above and the details are omitted.

Next, we prove the statement for the case of $|I| \geq p + 1$. Note first that students outside the upper segment $L_1 \cup \ldots \cup L_{p-2}$ cannot influence tie breaking in the upper segment. To see this note that the only possible effect such a student could have on this tie-breaking decision is to initiate a rejection chain leading to the rejection of the student indexed $p-1$ by some specialized school $s$. Since there cannot be more than one student from $L_{p-1}$ who has higher priority for $s$ than $p-1$, this implies that there are $q^1 - 1$ students from the upper segment temporarily matched to $s$. Now remember that we only use temporary assignments for tie-breaking in the upper segment if some specialized school has filled all of its places with students from the upper segment. But if $p - 1$ is temporarily matched to some specialized school $s' \in S^1 \setminus \{s\}$ together with $q^1 - 1$ other students from the upper segment, no student from the upper segment is rejected by a non-specialized school given that $|I| \leq p$.[34] The proof that no student in the upper segment can manipulate tie-breaking in the upper segment to her benefit is completely analogous to the proof for the case of $|I| \leq p$ and the details are omitted.

It remains to be shown that no student can profitable manipulate the tie-breaking decision in the lower segment. In particular, the proof is complete unless $|L_K| = 2$. Now note that the endogenous tie-breaking of the SDA-ETB in the lower segment ensures that there are no additional rejections after the tie-breaking stage since (i) only student $K$ can ever rank lower than one of $K+1$ and $K+2$, and (ii) $K$ cannot rank below both of these students at some specialized school. Hence, a student can profitably manipulate tie-breaking in the lower segment only if she obtains a better school prior to tie-breaking. Now suppose that contrary to what we want to show some student $i$ can profitably manipulate the tie-breaking procedure when the profile of (true) preferences is $R$ by submitting $R'_i$. Note that it has to be the case that the tie between $K+1$ and $K+2$ needs to be broken endogenously under $R$ and $R' = (R'_i, R_{-i})$: Otherwise we could use the same strict priority structure

[34]It is precisely at this point where we need that the upper segment contains $p-1$ students and not $p$ students if $|I| \geq p + 1$.

under $R$ and $R'$ so that we obtain a contradiction to the strategy-proofness of SDA for a fixed strict priority structure. This already implies that neither $K+1$ nor $K+2$ can manipulate the SDA-ETB procedure to their benefit. Let $s$ and $s'$ be the schools to which $K+1$ and $K+2$ are temporarily matched before tie-breaking in the lower segment under $R$ and $R'$, respectively. By the exogenous tie-breaking of the SDA-ETB exactly $q_s - 1$ students in $I \setminus \{K+1, K+2\}$ apply to $s$ in the course of SDA-ETB under $R$ and exactly $q_{s'} - 1$ students in $I \setminus \{K+1, K+2\}$ apply to $s'$ in the course of SDA-ETB under $R'$. Note that $s = s'$ unless $i$ applies to $s'$ (prior to tie-breaking) under $R'_i$ but not under $R_i$. Since there are no rejections after tie-breaking in the lower segment, $s \neq s'$ would imply that $s' = f_i^{ETB}(R')$ and thus $s = f_i^{ETB}(R) R_i f_i^{ETB}(R') = s'$. So we may assume that $s = s'$. But then we would obtain the same final matching for students outside the lower segment (under both $R$ and $R'$), if we assumed (contrary to fact) that $s$ could admit $q_s + 1$ students and has a strict priority ranking of students (by arbitrarily breaking the remaining ties in $\succeq^0$). This is again a contradiction to strategy-proofness of the SDA for a fixed strict priority structure and completes the proof.

□

## A.3 Appendix to Chapter 3

### Proof of Theorem 4

$(i) \Rightarrow (ii)$ For agent $v \in V$ and a contract $x \in X(v)$, let $S_{x,v}$ denote the direction of the contract relative to $v$, that is, $S_{x,v} = U_v$ if and only if $x$ is an upstream contract for $v$. Let $\overline{S}_{x,v}$ denote the complementary direction, that is, $\overline{S}_{x,v} = D_v$ if and only if $S_{x,v} = U_v$.

The proof is by contradiction: Suppose that $(G_X, c, q)$ is weakly acyclic but for some preference profile $R \in \mathcal{R}_{(c,q)}$ there exists a chain stable matching $\mu$ which is not group stable. By the definition of group stability there must then be a coalition $A$ that blocks $\mu$ via network $\mu'$. We show that $\mu' \setminus \mu$ contains a blocking chain of $\mu$. The following procedure, which we call the *chain algorithm*, will be key to the proof.

**Step 1:** Let $v_1 \in A$ and $x \in Ch_{v_1}(\mu(v_1) \cup \mu'(v_1)) \setminus \mu(v_1)$ be arbitrary.

If $x \in Ch_{v_1}(\mu(v_1) \cup \{x\})$, set $x_1 := x$ and $B_1 := \emptyset$.

Else let $x_1 \in Ch_{v_1}\left[\mu(v_1) \cup \{x\} \cup \overline{S}_{x,v_1}(\mu'(v_1))\right] \setminus (\mu(v_1) \cup \{x\})$ be arbitrary.

If $x_1 \in Ch_{v_1}(\mu(v_1) \cup \{x_1\})$, set $B_1 := \emptyset$.

Else set $B_1 := \{v_1\}$.

$\vdots$

**Step** $k+1$**:** Let $v_{k+1} \neq v_k$ be the other node involved with $x_k$.

If $x_k \in Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \{x_k\})$, stop.

If $x_k \in Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \mu'(v_{k+1}))$ but $x_k \notin Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \{x_k\})$, let $x_{k+1} \in Ch_{v_{k+1}}[\mu(v_{k+1}) \cup \{x_k\} \cup \overline{S}_{x_k,v_{k+1}}(\mu'(v_{k+1}))] \setminus (\mu(v_{k+1}) \cup \{x_k\})$ be arbitrary and set $B_{k+1} := B_k$.

If $x_k \notin Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \mu'(v_{k+1}))$, let $x \in Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \mu'(v_{k+1})) \setminus \mu(v_{k+1})$ be arbitrary.

If $x \in Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \{x\})$ set $x_{k+1} := x$ and $B_{k+1} := B_k$.

Else let $x_{k+1} \in Ch_{v_{k+1}}[\mu(v_{k+1}) \cup \{x\} \cup \overline{S}_{x,v_{k+1}}(\mu'(v_{k+1}))] \setminus (\mu(v_{k+1}) \cup \{x\})$ be arbitrary.

If $x_{k+1} \in Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \{x_{k+1}\})$, set $B_{k+1} := B_k$.

Else set $B_{k+1} := B_k \cup \{v_{k+1}\}$.

$\vdots$

The sequence $\{B_k\}$ produced by the algorithm is a stack of agents who are marked for later processing. In order to show that this algorithm is well defined and terminates after a finite number of rounds we need the following lemma.

**Lemma 2.** *Let $v \in A$ be arbitrary.*

*(i) There exists a contract $x \in Ch_v(\mu(v) \cup \mu'(v)) \setminus \mu(v)$.*

*(ii) If $x \in Ch_v(\mu(v) \cup \mu'(v)) \setminus \mu(v)$ and $x \notin Ch_v(\mu(v) \cup \{x\})$ then $Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))] \setminus (\mu(v) \cup \{x\}) \neq \emptyset$.*

*(iii) If $x \in Ch_v(\mu(v) \cup \mu'(v)) \setminus \mu(v)$, $x \notin Ch_v(\mu(v) \cup \{x\})$, $y \in Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))] \setminus (\mu(v) \cup \{x\})$, and $y \notin Ch_v(\mu(v) \cup \{y\})$ then $\{x, y\} \subseteq Ch_v(\mu(v) \cup \{x, y\})$.*

**Proof of Lemma 2:**

(i) Let $v \in A$ be arbitrary. Since $\mu \in \mathcal{CS}(R) \subseteq \mathcal{IS}(R)$ it cannot be the case that $Ch_v(\mu(v) \cup \mu'(v)) \subseteq \mu(v)$. Otherwise revealed preference would imply that $Ch_v(\mu(v) \cup \mu'(v)) = \mu(v)$ so that $\mu(v) R_v \mu'(v)$. This contradicts the assumption that $\mu$ is blocked by coalition $A$ via $\mu'$.

(ii) If $x \in Ch_v(\mu(v) \cup \mu'(v)) \setminus \mu(v)$, SSS implies that $x \in Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))]$. If $x \notin Ch_v(\mu(v) \cup \{x\})$ but $Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))] \setminus (\mu(v) \cup \{x\}) = \emptyset$ we would have $Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))] \subseteq \mu(v) \cup \{x\}$. Hence, by revealed preference, $Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))] = Ch_v(\mu(v) \cup \{x\})$. Since $x \notin Ch_v(\mu(v) \cup \{x\})$ this is a contradiction.

(iii) If $y \in Ch_v[\mu(v) \cup \{x\} \cup \overline{S}_{x,v}(\mu'(v))] \setminus (\mu(v) \cup \{x\})$, SSS implies that $y \in Ch_v(\mu(v) \cup \{x, y\})$. If $y \notin Ch_v(\mu(v) \cup \{y\})$ and $x \notin Ch_v(\mu(v) \cup \{x, y\})$, we would have $Ch_v(\mu(v) \cup \{x, y\}) \subseteq \mu(v) \cup \{y\}$. Hence, by revealed preferences $Ch_v(\mu(v) \cup \{x, y\}) = Ch_v(\mu(v) \cup \{y\})$. Since $y \notin Ch_v(\mu(v) \cup \{y\})$ this is a contradiction.

□

Lemma 2 implies that the case distinction made by the algorithm is exhaustive: For the first step this is immediate. So let $l \geq 1$ be some non-terminal step of the chain algorithm and let $\{v_k\}_{k=1}^{l}$ and $\{x_k\}_{k=1}^{l}$ denote the sequences of agents and contracts considered by the chain algorithm up to step $l$. Let $v_{l+1}$ be the agent considered by the algorithm in step $l+1$. We must have $v_{l+1} \in A$ since the algorithm considers only contracts in $\mu' \setminus \mu$. By Lemma 2, one of the following cases has to apply.

Case 1: $x_l \in Ch_{v_{l+1}}(\mu(v_{l+1}) \cup \{x_l\})$,

Case 2: $\{x_l, x_{l+1}\} \subseteq Ch_{v_{l+1}}(\mu(v_{l+1}) \cup \{x_l, x_{l+1}\})$,

Case 3: $x_{l+1} \in Ch_{v_{l+1}}(\mu(v_{l+1}) \cup \{x_{l+1}\})$, or

Case 4: $\{x, x_{l+1}\} \subseteq Ch_{v_{l+1}}(\mu(v_{l+1}) \cup \{x, x_{l+1}\})$ for some $x \in \overline{S}_{x_{l+1}, v_{l+1}}[\mu'(v_{l+1}) \setminus (\mu(v_{l+1}) \cup \{x_l\})]$.

Note that the cases are not in general mutually exclusive. The algorithm first checks whether Case 1 applies and then proceeds to check whether Case 2 or 3 applies. We now show that the algorithm must terminate after a finite number of rounds. Suppose to the contrary that this is not the case. Then, since the set of agents is finite there must exist indices $k$ and $l$ such that $k < l$ and $v = v_k = v_l$. We can assume w.l.o.g. that all agents between $v_k$ and $v_l$ are different. Since all agents considered by the chain algorithm are part of the blocking coalition $A$ we must have $\mu'(v_j) P_{v_j} \mu(v_j) R_{v_j} \emptyset$ for all $j \in \{k, \ldots, l-1\}$. If $l-k \geq 2$, $v_k, \ldots, v_{l-1}$ is a cycle in $G_X$. Since $\{x_{k+1}, \ldots, x_{l-1}\} \subseteq \mu'$ no agent on the cycle is capacity constrained and we have found a 2 cycle in the market structure. If $l - k = 1$, note that since $x_k \notin Ch_{v_{k+1}}(\mu(v_{k+1}) \cup \{x_k\})$ (as the algorithm does not terminate in step

$k+1$) we must have $x_k \neq x_{k+1}$. But then $|\mu \cap X(v_k, v_{k+1})| \geq 2$ and since preferences conform to capacities, $c(v_k, v_{k+1}) \geq 2$. In both cases we thus obtain a contradiction to weak acyclicity.

In the following we use the chain algorithm as a subroutine for another algorithm that finds a chain block of $\mu$ given weak acyclicity. This completes the proof since $\mu$ was assumed to be chain stable. The algorithm uses a special version of the chain algorithm that starts from a given agent-contract pair: Taking $v \in V$ and $x \in \mu'(v) \setminus \mu(v)$ as given, the algorithm proceeds as if the first step of the chain algorithm yielded $v_1 := v$, $x_1 := x$ and $B_1 = \emptyset$ (so that $v_2$ is the other agent involved with $x$). The *chain block* algorithm works as follows.

**Step 1:** Run the chain algorithm and let the resulting sequence of agents, contracts, and stacks be denoted by $\{v_k^1, x_k^1, B_k^1\}_{k\geq 1}$. Let $K_1$ be the last step of the algorithm. If $B_{K_1}^1 = \emptyset$ stop. Otherwise let $j_1$ be the highest index $j$ such that $v_j^1 \in B_{K_1}^1$. Set $v_1^2 := v_{j_1}^1$ and proceed with step 2.

$\vdots$

**Step $t+1$:** Let $x_1^{t+1} \in \overline{S}_{x_{j_t}^t, v_1^{t+1}}[\mu'(v_1^{t+1}) \setminus (\mu(v_1^{t+1}) \cup \{x_{j_t-1}^t\})]$ be such that $\{x_{j_t}^t, x_1^{t+1}\} \subseteq Ch_{v_1^{t+1}}(\mu(v_1^{t+1}) \cup \{x_{j_t}^t, x_1^{t+1}\})$. Run the chain algorithm starting at $(v_1^{t+1}, x_1^{t+1})$ and let the resulting sequence of agents, contracts and stacks be denoted by $\{v_k^{t+1}, x_k^{t+1}, B_k^{t+1}\}_{k\geq 1}$. Let $K_{t+1}$ be the last step of the algorithm. If $B_{K_{t+1}}^{t+1} = \emptyset$ stop. Otherwise let $j_{t+1}$ be the highest index $j$ such that $v_j^t \in B_{K_{t+1}}^{t+1}$. Set $v_1^{t+2} := v_{j_{t+1}}^{t+1}$ and proceed with step $t+2$.

$\vdots$

Note that the algorithm is well defined since an agent is put on the stack if and only if only Case 4 applies. The proof that the algorithm terminates in finite time is analogous to the corresponding proof for the chain algorithm and the details are omitted.

Let $T$ be the last iteration of the chain algorithm. We now show how to find a blocking chain of $\mu$. Since $K_T$ is the last step of this iteration we have $x_{K_T-1}^T \in Ch_{v_{K_T}^T}(\mu(v_{K_T}^T) \cup \{x_{K_T-1}^T\})$. Since $B_{K_T}^T = \emptyset$ one of the Cases 1-3 holds for each step of the last ($T$th) iteration of the chain algorithm. If there is a step of the $T$th iteration for which Case 3 holds, let $j$ be the last step with this property. Then $x_j^T, \ldots, x_{K_T-1}^T$ is a blocking chain if $v_j^T = s_{x_j^T}$. To see this note that by the definition of $j$, Case 2 must hold for all steps $k \in \{j+1, \ldots, K_T-1\}$ and Case 1 holds for step $K_T$. Furthermore, $v_k^T = b_{x_{k-1}^T} = s_{x_k^T}$ for all $k \in \{j+1, \ldots, K_T-1\}$. If $v_j^T = b_{x_j^T}$, the sequence in reverse order is a blocking chain.

If there is no step of the $T$th iteration such that Case 3 holds consider the sequence of agents and contracts in the $T-1$st iteration of the chain algorithm starting at $v^{T-1}_{j_{T-1}}$. By definition of $j_{T-1}$ there is no step $i > j_{T-1}$ such that only Case 4 applies. If there is a step $i > j_{T-1}$ of the $T-1$st iteration such that Case 3 applies, let $l$ be the largest index with this property. Then $x^{T-1}_l, \ldots, x^{T-1}_{K_{T-1}-1}$ is a blocking chain. If there is no step for which Case 3 applies, $x^{T-1}_{K_{T-1}-1}, \ldots, x^{T-1}_{j_{T-1}}, x^T_1, \ldots, x^T_{K_T-1}$ must be a blocking chain. In both cases it might be necessary to reverse the order of the sequence to obtain a blocking chain as above.

Thus, the above algorithm finds a chain block of $\mu$ within $\mu' \setminus \mu$. This contradicts the assumption that $\mu$ is chain stable and completes the proof that $(i) \Rightarrow (ii)$.

$(ii) \Rightarrow (iii)$ This follows immediately from the observation that $\mathcal{GS}(R) \subseteq C(R)$, for all $R \in \mathcal{R}_{(c,q)}$.

$(iii) \Rightarrow (iv)$ Suppose that (iii) holds but that there exists a preference profile $R \in \mathcal{R}_{(c,q)}$ and a network $\mu \in \mathcal{CS}(R) \setminus \mathcal{E}(R)$. By definition, this means that there exists a network $\mu'$ that all agents weakly and some strictly prefer over $\mu$. But then $\mu'$ weakly dominates $\mu$ via $V$, which contradicts (iii).

$(iv) \Rightarrow (i)$ This will follow from the proof that $(iii) \Rightarrow (i)$ in Theorem 5 since chain stable networks exist for all profiles in $\mathcal{R}_{(c,q)}$.

□

## Proof of Theorem 5

$(i) \Rightarrow (ii)$ This follows from the implication $(i) \Rightarrow (ii)$ of Theorem 1 that was proven above and the existence of chain stable networks on the domain $\mathcal{R}_{(c,q)}$.

$(ii) \Rightarrow (iii)$ Since a group stable network is efficient and individually stable, this implication is again immediate.

$(iii) \Rightarrow (i)$ Suppose first that there is a 2 cycle $v_1, \ldots, v_n$ in $(G_X, c, q)$. We show that there exists a profile $R \in \mathcal{R}_q$ such that $\mathcal{E}(R) \cap \mathcal{IS}(R) = \emptyset$. Since we can always construct the preference profile in such a way that no agent on the cycle wants to sign a contract with the outside world and vice versa, we can assume that $V = \{v_1, \ldots, v_n\}$. Let $x_i \in X$ denote some contract between $v_i$ and $v_{i+1}$ for $i \in \{1, \ldots, n\}$. The preferences of agents on the cycle are defined as follows: For all $i \in \{1, \ldots, n\}$, Agent $v_i$ strictly prefers signing only contract $x_i$ over signing the set of contracts $\{x_{i-1}, x_i\}$, over signing no contracts at all. These are the only acceptable sets of contracts. It is easy to check that the resulting profile $R \in \mathcal{R}_{(c,q)}$.

We show that for this profile any individually stable network must assign the empty set of contracts to all agents. By construction of the preference profile, individual stability demands that $v_1$ is assigned either the empty set of contracts or only $x_1$. So suppose that there is an individually stable network $\mu$ that includes $x_1$. But $v_2$ must be assigned either the empty set of contracts or only contract $x_2$ in an individually stable network. Hence, any individually stable network must assign the empty set of contracts to agent $v_1$. A simple repetition of this argument establishes that the unique individually stable network is the empty network. But this network is (strictly) Pareto dominated by the complete network.

Now consider the case that there is some pair of agents $v, w$ such that $(v, w) \in G_X$ and $c(v, w) \geq 2$. Take two contracts $x, y \in X(v, w)$. Let $\{x\}P_v\{x, y\}P_v\emptyset P_v\{y\}$ and $\{y\}P_w\{x, y\}P_w\emptyset P_w\{x\}$, and assume that no other sets of contracts are acceptable to either $v$ or $w$. As above only the empty network is individually stable but Pareto dominated by the network $\{x, y\}$. This completes the proof that $(iii) \Rightarrow (i)$.

$\square$

## Proof of Theorem 6

$(i) \Rightarrow (ii)$ Suppose to the contrary that $(G_X, c, q)$ is weakly acyclic but that for some preference profile $R \in \mathcal{R}_{(c,q)}$ there exists a core stable network $\mu$ which is not individually stable. Let $v_0 \in V$ be an agent who would like to drop some of the contracts in $\mu(v_0)$, that is, $Ch_{v_0}(\mu(v_0)) \neq \mu(v_0)$. This implies in particular that $\mu(v_0) \neq \emptyset$. But then it has to be the case that $Ch_{v_0}(\mu(v_0)) \neq \emptyset$ since $\mu$ is at least individually rational. We denote by $\mu'$ the network that results from $\mu$ when contracts in $\mu(v_0) \setminus Ch_{v_0}(\mu(v_0))$ are deleted. For the following let $V_0 := \{v_0\}$.

Let $V_1 \subset V \setminus V_0$ be the (nonempty) set of agents who are involved with some contract in $Ch_{v_0}(\mu(v_0))$ and let $W$ be the set of agents who are involved with one of the contracts $v_0$ wants to drop, that is, with one of the contracts in $\mu(v_0) \setminus Ch_{v_0}(\mu(v_0))$. We must have $V_1 \cap W = \emptyset$ since $c(v, w) \leq 1$ for all $v, w \in V$ and $\mu$ is individually rational. Furthermore, $\mu$

cannot contain a contract between a pair of agents in $V_1 \times (W \cup V_1)$ if the market structure is weakly acyclic. Otherwise, there would be two agents $w_1 \in V_1$ and $w_2 \in W \cup V_1$ such that $\mu$ contains a contract between $w_1$ and $w_2$. By definition of $V_1$ and $W$, $\mu$ also contains contracts between $v_0$ and both, $w_1$ and $w_2$. Since $\mu$ is individually rational, none of the three agents can be capacity constrained. Hence, we have found a 2 cycle in $(G_X, c, q)$, a contradiction. On the other hand, $\mu$ has to contain at least one contract between an agent in $V_1$ and an agent in $V \setminus (W \cup V_0 \cup V_1)$. Otherwise $\mu'$ weakly dominates $\mu$ via the coalition $V_0 \cup V_1$ since (i) all agents in $V_1$ would be indifferent between these two networks, (ii) $v_0$ strictly prefers $\mu'$ over $\mu$, and (iii) no agent in $V_0 \cup V_1$ signs a contract with an agent in $V \setminus (V_0 \cup V_1)$. Thus, there has to be a nonempty set of agents $V_2 \subset V \setminus (W \cup V_0 \cup V_1)$ who sign a contract with some agent in $V_1$ under $\mu$.

Now suppose that for some $k \geq 2$ we have shown that there is a sequence of sets of agents $V_1, \ldots, V_k$ such that, for all $l \in \{2, \ldots, k\}$, $V_l \subset V \setminus (W \cup V_0 \cup \ldots \cup V_{l-1})$ and the set of all agents who sign a contract with agents in $V_{l-1}$ under $\mu$ is $V_{l-2} \cup V_l$. If the market structure is weakly acyclic, $\mu$ cannot contain a contract between a pair of agents in $V_k \times (W \cup V_0 \ldots \cup V_k)$. The argument is similar to above. On the other hand, $\mu$ has to contain at least one contract between an agent in $V_k$ and an agent in $V \setminus (W \cup V_0 \ldots \cup V_k)$. Otherwise $\mu'$ weakly dominates $\mu$ via the coalition $V_0 \cup \ldots \cup V_k$ since (i) all agents in $V_1 \cup \ldots \cup V_k$ would be indifferent between these two networks, (ii) $v_0$ strictly prefers $\mu'$ over $\mu$, and (iii) no agent in $V_0 \cup \ldots \cup V_k$ signs a contract with an agent in $V \setminus (V_0 \cup \ldots \cup V_{k-1})$. Thus, there has to be a nonempty set of agents $V_{k+1} \subset V \setminus (W \cup V_0 \cup \ldots \cup V_k)$ who sign a contract with some agent in $V_k$ under $\mu$.

The above argument is valid for any $k$ and the procedure would thus run forever, contradicting the finiteness of $V$. This completes the proof that $(i) \Rightarrow (ii)$.

$(ii) \Rightarrow (i)$ Note that $\mathcal{C}(R) \subseteq \mathcal{IS}(R)$ for all $R \in \mathcal{R}_{(c,q)}$ implies that an efficient and individually stable network always exists. Hence, the statement follows from $(iii) \Rightarrow (i)$ in Theorem 5.

□

## Proof of Theorem 7

$(i) \Rightarrow (ii)$ Before going to the details of the proof, note that any network $\mu$ with $|\mu \cap X(v, w)| \leq 1$ for all $v, w \in V$ defines a (unique) subgraph $G_\mu$ of $G_X$ that includes an edge from $v$ to

$w$ if and only if $\mu$ contains a $x$ with $s_x = v$ and $b_x = w$.

By Theorem 4 we know that, for all $R \in \mathcal{R}_{(c,q)}$, $\mathcal{CS}(R) \subseteq \mathcal{C}(R)$ if the market structure is weakly acyclic. Hence, we only need to show that $\mathcal{C}(R) \subseteq \mathcal{CS}(R)$ if the market structure is strongly acyclic. The proof will be by contradiction: Assume that for some $R \in \mathcal{R}_{(c,q)}$ there exists a network $\mu \in \mathcal{C}(R) \backslash \mathcal{CS}(R)$. Since strong implies weak acyclicity we have that $\mathcal{C}(R) \subseteq \mathcal{IS}(R)$ by Theorem 6. Hence, there must be a chain $x_1, \ldots, x_n \notin \mu$ that blocks $\mu$. Consider the network $\mu'$ that results from $\mu$ when we add the contracts $x_1, \ldots, x_n$ and delete contracts in $\mu(s_{x_1}) \setminus Ch_{s_{x_1}}(\mu(s_{x_1}) \cup \{x_1\})$, $\mu(b_{x_n}) \setminus Ch_{b_{x_n}}(\mu(b_{x_n}) \cup \{x_n\})$, and $\mu(b_{x_i}) \setminus Ch_{b_{x_i}}(\mu(b_{x_i}) \cup \{x_i, x_{i+1}\})$ for all $i < n$. Note that $\mu'$ and $\mu$ can both contain at most one contract between each pair of agents since the market structure is strongly acyclic and (i) $\mu'(v) P_v \mu(v) R_v \emptyset$, for all $v \in \{s_{x_1}, b_{x_1}, \ldots, b_{x_n}\}$, (ii) $\mu'(v) \subseteq \mu(v)$, for all $v \in V \setminus \{s_{x_1}, b_{x_1}, \ldots, b_{x_n}\}$, as well as (iii) $\mu(v) R_v \emptyset$, for all $v \in V$. Now let $A$ be the set of agents who are in the same connected component of $G^{\mu'}$ as $s_{x_1}$.[35] We claim that $\mu'$ weakly dominates $\mu$ via $A$. As noted above, all agents involved with some contract in the blocking chain strictly prefer $\mu'$ over $\mu$. Now suppose there is an agent $\hat{v} \in A \setminus \{s_{x_1}, b_{x_1}, \ldots, b_{x_n}\}$ such that $\mu(\hat{v}) P_{\hat{v}} \mu'(\hat{v})$. Since we have only deleted some contracts involving agents on the blocking chain this means that there is a contract $x \in \mu \setminus \mu'$ which involves $\hat{v}$ and an agent $\tilde{v} \in \{s_{x_1}, b_{x_1} \ldots, b_{x_n}\}$. Since $\mu'(\hat{v}) = \mu(\hat{v}) \setminus \{x\}$ and $\mu(\hat{v}) R_{\hat{v}} \emptyset$, $\mu'$ cannot contain a contract between $\hat{v}$ and $\tilde{v}$. Given that $\hat{v}$ is in the same connected component of $G^{\mu'}$ as $s_{x_1}$, $G^{\mu' \cup \{x\}}$ contains a cycle $v_1, \ldots, v_n$ with $\{\hat{v}, \tilde{v}\} \subset \{v_1, \ldots, v_n\}$. We now show that this must be a restricted 2 cycle. This contradiction shows that $\mu'$ must dominate $\mu$ via $A$ so that $\mu$ could not have been in the core.

Note that the following two cases cannot occur: (i) $\tilde{v} = s_{x_1}$ and $x$ is an upstream contract for $\tilde{v}$, and (ii) $\tilde{v} = b_{x_n}$ and $x$ is a downstream contract for $\tilde{v}$. In case (i) we would have that $x \in Ch_{\tilde{v}}(\mu(\tilde{v}))$, due to the individual stability of $\mu$, and $x \notin Ch_{\tilde{v}}(\mu(\tilde{v}) \cup \{x_1\})$. But then CSC would be violated since $x_1$ is a downstream contract for $\tilde{v}$ and $x$ is an upstream contract for $\tilde{v}$. Case (ii) can be handled similarly. This shows that if $\tilde{v}$ is a passing node of the cycle $v_1, \ldots, v_n$, she must be one of the intermediate agents in the blocking chain. But this means that $\tilde{v}$ signs an upstream <u>and</u> a downstream contract in $\mu'$. Since $\mu'(\tilde{v}) P_{\tilde{v}} \mu(\tilde{v}) R_{\tilde{v}} \emptyset$, $\tilde{v}$ cannot be capacity constrained on the cycle if she is a passing node. But $\tilde{v}$ is the only potentially capacity constrained agent on the cycle. To see this, note that all agents in $\{v_1, \ldots, v_n\} \setminus \{\hat{v}, \tilde{v}\}$ sign contracts with both of their neighbors under

[35] Two nodes $v$ and $w$ are in the same connected component of a directed graph $G$ if there is a sequence of edges connecting $v$ and $w$.

$\mu'$ and that $\mu(\hat{v})$ contains contracts between $\hat{v}$ and $\tilde{v}$ as well as her other neighbor on the cycle. For each agent $v \in (\{s_{x_1}, b_{x_1}, \dots, b_{x_n}\} \cap \{v_1, \dots, v_n\}) \setminus \{\tilde{v}\}$, $\mu'(v) P_v \mu(v) R_v \emptyset$ so that $\mu'(v)$ is an acceptable set of contracts and $v$ is not capacity constrained. For each agent $v \in \{v_1, \dots, v_n\} \setminus \{s_{x_1}, b_{x_1}, \dots, b_{x_n}\}$ we have $\mu'(v) \subseteq \mu(v)$ and, since $\mu$ is a core allocation, $\mu(v) P_v \emptyset$ so that $v$ is not capacity constrained. In both cases $v$ cannot be capacity constrained on the cycle.

$(ii) \Rightarrow (i)$ If weak acyclicity is not satisfied, consider the counterexamples used to prove that $(iii) \Rightarrow (i)$ in Theorem 5. In both types of examples it is easy to check that the core consists of the complete network while the empty network is the unique chain stable network. Now suppose that weak acyclicity is satisfied but that there is a restricted 2 cycle $v_1, \dots, v_n$ for which only the source $v_1$ is capacity constrained (the case where $v_1$ is a sink can be handled similarly). Let $x_1, \dots, x_n$ be an accompanying sequence of contracts, that is, $x_k$ is a contract between agents $v_k$ and $v_{k+1}$ (where $n+1 := 1$). As in the proof that $(iii) \Rightarrow (i)$ in Theorem 5 we can assume that $V = \{v_1, \dots, v_n\}$. We now define a preference profile for the agents starting with $R_{v_1} = \{x_1\}, \{x_n\}$. Let $k \in \{2, \dots, n\}$ be arbitrary and set $R_{v_k} = \{x_{k-1}, x_k\}$ if $v_k$ is a passing node, and $R_{v_k} = \{x_{k-1}, x_k\}, \{x_k\}, \{x_{k-1}\}$ in any other case. Let the resulting profile be denoted by $R$ and note that $R \in \mathcal{R}_{(c,q)}$ since $v_1$ is the only capacity constrained agent among $v_1, \dots, v_n$.

Now let $j$ be the smallest index in $\{2, \dots, n\}$ such that $v_j$ is a sink. Note that such an index must exist since $G_X$ contains no directed cycles and all agents between $v_1$ and $v_j$ are passing nodes. Consider the network $\mu = \{x_j, \dots, x_n\}$. Clearly, $\mu \notin \mathcal{CS}(R)$ since it is blocked by the chain $x_1, \dots, x_{j-1}$. We now show that $\mu \in \mathcal{C}(R)$.

Suppose to the contrary that $\mu$ is weakly dominated by some network $\mu'$ via a coalition $A$. Since agents $v_{j+1}, \dots, v_n$ get their most preferred set of contracts under $\mu$, $A \cap \{v_1, \dots, v_j\} \neq \emptyset$. Since agents $v_2, \dots, v_{j-1}$ are all passing nodes by the definition of $j$, the construction of $R$ implies $\{x_1, \dots, x_{j-1}\} \subseteq \mu'$. Hence, we must have $\{v_1, \dots, v_j\} \subseteq A$. Since $\{x_j\} P_{v_j} \{x_{j-1}\}$ this implies $x_j \in \mu'$ and thus $v_{j+1} \in A$. Continuing with this line of reasoning it is easy to see that we must have $\mu' = \{x_1, \dots, x_n\}$ and $A = \{v_1, \dots, v_n\}$. But then $v_1$ is worse off compared to $\mu$ since $\emptyset P_{v_1} \{x_1, x_n\}$. Hence, it has to be the case that $\mu \in \mathcal{C}(R)$.

$\square$

## Discussion of the main results

For this Appendix we assume that there are no capacity constraints and let $\mathcal{R}$ denote the set of all preference profiles satisfying strict preferences, no externalities, SSS, and CSC. The following statements are easily seen to be true for any given $R \in \mathcal{R}$.

$\mathcal{CS}(R) = \mathcal{GS}(R) \Rightarrow \mathcal{CS}(R) \subseteq \mathcal{C}(R), \mathcal{CS}(R) \subseteq \mathcal{E}(R), \mathcal{GS}(R) \neq \emptyset, \mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$

$\mathcal{CS}(R) \subseteq \mathcal{C}(R) \Rightarrow \mathcal{CS}(R) \subseteq \mathcal{E}(R), \mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$

$\mathcal{CS}(R) \subseteq \mathcal{E}(R) \Rightarrow \mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$

$\mathcal{GS}(R) \neq \emptyset \Rightarrow \mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$

In this Appendix we show that all other implications of Corollary 3 are not necessarily true for any given preference profile. Chain stable networks in the examples can be calculated using the *T-algorithm* of Ostrovsky (2008). All counterexamples except example 3 use a supply chain model with five agents and the following graph of potential interactions:

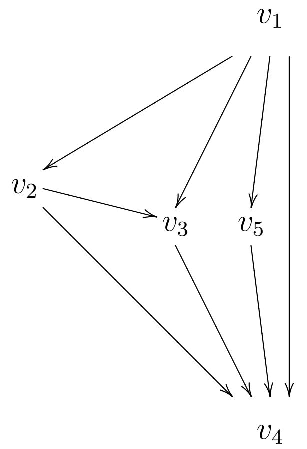


**Figure 3:** Graph $G_3$ of potential interactions.

Throughout the Appendix we use the following notation: $x_i^j$ denotes some contract in which agent $i$ sells something to agent $j$. Agent $v_5$ will only be needed for the last example.

1. There exist profiles $R \in \mathcal{R}$ such that $\mathcal{CS}(R) \subseteq \mathcal{C}(R)$, $\mathcal{CS}(R) \subseteq \mathcal{E}(R)$, and $\mathcal{E}(R) \cap \mathcal{IS}(R) \neq \emptyset$, but $\mathcal{CS}(R) \neq \mathcal{GS}(R)$ and $\mathcal{GS}(R) = \emptyset$.

   Suppose preferences of the agents are[36]

| $R^1$ | $R^1_{v_1}$ | $R^1_{v_2}$ | $R^1_{v_3}$ | $R^1_{v_4}$ |
|---|---|---|---|---|
| | $\{x_1^2\}$ | $\{x_1^2, x_2^3\}$ | $\{x_1^3, x_3^4\}$ | $\{x_1^4, x_3^4\}$ |
| | $\{x_1^2, x_1^3\}$ | | $\{x_1^3, x_2^3, x_3^4\}$ | $\{x_1^4\}$ |
| | $\{x_1^4\}$ | | $\{x_3^4\}$ | $\{x_3^4\}$ |

[36] As in all examples that follow it is easy to check that SSS and CSC are indeed satisfied.

Using the T-Algorithm it is easy to show that the unique chain stable network is given by $\mu = \{x_1^4, x_3^4\}$.

To see that $\mu \in \mathcal{C}(R^1)$ note that $v_4$ cannot be made better off and the only network which makes $v_1$ and $v_3$ better off without hurting $v_2$ is $\{x_1^2, x_1^3, x_2^3, x_3^4\}$. This network would make $v_4$ worse off and thus does not weakly dominate $\mu$ in the sense of the core. Hence, $\mu$ is in the core and thus in particular efficient.

On the other hand, $\mu$ is blocked by $\{v_1, v_2, v_3\}$ via $\{x_1^2, x_1^3, x_2^3, x_3^4\}$ so that $\mu$ is not group stable. The only other nonempty individually stable network is $\{x_3^4\}$ which is not even chain stable. Since a group stable matching has to be individually stable, this shows that $\mathcal{GS}(P) = \emptyset$.

2. There exist preference profiles $R \in \mathcal{R}$ such that $\mathcal{IS}(R) \cap \mathcal{E}(R) \neq \emptyset$ , but $\mathcal{CS}(R) \cap \mathcal{C}(R) = \emptyset$ and $\mathcal{CS}(R) \cap \mathcal{E}(R) = \emptyset$.

   Consider the following preference profile

| $R^2$ | $R^2_{v_1}$ | $R^2_{v_2}$ | $R^2_{v_3}$ | $R^2_{v_4}$ |
|---|---|---|---|---|
| | $\{x_1^4, x_1^3\}$ | $\{x_1^2, x_2^3, x_2^4\}$ | $\{x_1^3, x_2^3, x_3^4\}$ | $\{x_1^4, x_2^4\}$ |
| | $\{x_1^2, x_1^3\}$ | $\{x_2^3\}$ | $\{x_2^3\}$ | $\{x_2^4, x_3^4\}$ |
| | $\{x_1^4\}$ | $\{x_2^4\}$ | $\{x_1^3\}$ | $\{x_1^4\}$ |
| | $\{x_1^2\}$ | | | $\{x_3^4\}$ |
| | $\{x_1^3\}$ | | | $\{x_2^4\}$ |

   The unique chain stable network is given by $\mu = \{x_1^4, x_2^3\}$. But the network $\mu' = \{x_1^2, x_1^3, x_2^3, x_2^4, x_3^4\}$ is individually stable as well as efficient, and makes all agents better off (note that this network is blocked by the chain $x_1^4$).

3. There exist preference profiles $R \in \mathcal{R}$ such that $\mathcal{CS}(R) \subseteq \mathcal{E}(R)$ but $\mathcal{CS}(R) \cap \mathcal{C}(R) = \emptyset$.

   In the example of section 3.2 the unique chain stable network is the efficient network $\{x(M, D2)\}$. The unique core network is given by $\{x(M, S), x(S, D1), x(M, D1)\}$.

4. There exist preference profiles $R \in \mathcal{R}$ such that $\mathcal{GS}(R) \neq \emptyset$ but $\mathcal{CS}(R) \setminus \mathcal{E}(R) \neq \emptyset$, $\mathcal{CS}(R) \setminus \mathcal{GS}(R) \neq \emptyset$, and $\mathcal{CS}(R) \setminus \mathcal{C}(R) \neq \emptyset$.

   Preferences are given by:

| $R^2$ | $R^3_{v_1}$ | $R^3_{v_2}$ | $R^3_{v_3}$ | $R^3_{v_4}$ | $R^3_{v_5}$ |
|---|---|---|---|---|---|
| | $\{x_1^4\}$ | $\{x_1^2, x_2^4\}$ | $\{x_1^3, x_3^4\}$ | $\{x_5^4\}$ | $\{x_1^5, x_5^4\}$ |
| | $\{x_1^2\}$ | | | $\{x_2^4, x_3^4, x_5^4\}$ | |
| | $\{x_1^2, x_1^3, x_1^5\}$ | | | $\{x_3^4\}$ | |
| | $\{x_1^3\}$ | | | $\{x_1^4\}$ | |

For this profile there are two chain stable networks: $\{x_1^4\}$ and $\{x_1^3, x_3^4\}$.

The first network is also group stable, but the second is not even efficient as the network $\{x_1^2, x_1^3, x_1^5, x_2^4, x_3^4, x_5^4\}$ makes all agents (weakly) better off.

## Feasibility Restrictions on Networks

Instead of requiring preferences to conform to an exogenously given capacity vector, one could also restrict the set of feasible networks. Given a capacity vector $(c, q)$, the set of feasible networks $\mathcal{M}_{(c,q)}$ can be defined as follows: $\mu \in \mathcal{M}_{(c,q)}$ if and only if $\mu$ does not violate the capacity constraints, that is, for all $v \in V$, (i) $|\{w \in V \setminus \{v\} : s_x = w \text{ for some } x \in \mu(v)\}| \leq q_v^U$, (ii) $|\{w \in V \setminus \{v\} : b_x = w \text{ for some } x \in \mu(v)\}| \leq q_v^D$, and (iii) $|\mu(v) \cap X(v, w)| \leq c(v, w)$, for all $w \in V \setminus \{v\}$. Given some network $\mu \in \mathcal{M}_{(c,q)}$ let $G^\mu$ be the directed graph which contains one edge from $v$ to $w$ for each contract $x \in \mu$ such that $s_x = v$ and $b_x = w$. Note that in contrast to $G_X$ and $G_\mu$ used in the main text, this graph may contain multiple edges between a given pair of agents. The following shows how the acyclicity condition developed in Chapter 3 can be expressed in this framework.

**Proposition 8.** *The market structure $(G_X, c, q)$ is weakly acyclic if and only if $G^\mu$ is a forest for all $\mu \in \mathcal{M}_{(c,q)}$.*[37]

The proof is straightforward and omitted here. It is not clear, how strong acyclicity could have been formulated in this framework. This is the main reason for requiring preferences to conform to capacities instead.

## Beyond the Supply Chain Model

Some of the main results of chapter continue to hold without the CSC assumption. Let $(G_X, c, q)$ be a market structure and let $\hat{\mathcal{R}}_{(c,q)}$ be the set of all preference profiles that satisfy strict preferences, no externalities, and SSS, and conform to capacities. The following theorem summarizes

[37] A forest is a directed graph containing no directed or undirected cycles.

the results that carry over to this more general setting (in which the existence of chain stable networks cannot be guaranteed).

**Theorem 8.** *1. The following are equivalent:*

*(i)* $(G_X, c, q)$ *is weakly acyclic.*

*(ii)* $\mathcal{CS}(R) = \mathcal{GS}(R)$ *for all* $R \in \hat{\mathcal{R}}_{(c,q)}$.

*(iii)* $\mathcal{CS}(R) \subseteq \mathcal{C}(R)$ *for all* $R \in \hat{\mathcal{R}}_{(c,q)}$.

*(iv)* $\mathcal{CS}(R) \subseteq \mathcal{E}(R)$ *for all* $R \in \hat{\mathcal{R}}_{(c,q)}$.

*2. If* $\mathcal{GS}(R) \neq \emptyset$ *for all* $R \in \hat{\mathcal{R}}_{(c,q)}$ *then* $(G_X, c, q)$ *is weakly acyclic.*

*3. If* $\mathcal{IS}(R) \cap \mathcal{E}(R) \neq \emptyset$ *for all* $R \in \hat{\mathcal{R}}_{(c,q)}$ *then* $(G_X, c, q)$ *is weakly acyclic.*

*4.* $(G_X, c, q)$ *is weakly acyclic if and only if* $\mathcal{C}(R) \subseteq \mathcal{IS}(R)$ *for all* $R \in \hat{\mathcal{R}}_{(c,q)}$.

The proof of this Theorem follows directly from the proofs of Theorems 4,5, and 6: The proof of Theorem 4 does not use the CSC assumption. The counterexample used in the proof that *(iii)* $\Rightarrow$ *(i)* in Theorem 5 belongs to the larger domain $\hat{\mathcal{R}}_{(c,q)}$. The proof of Theorem 6 does not use CSC to prove sufficiency of weak acyclicity. The proof of necessity uses the same counterexample as Theorem 5. Weak acyclicity may not be sufficient for the existence of a group stable or efficient and individually stable network since the existence of a chain stable network is not guaranteed by Ostrovsky (2008)'s existence result.

One application of Theorem 8 is the *roommate problem* introduced in Gale and Shapley (1962). In this problem $2n$ agents have to be assigned among $n$ rooms that each have place for 2 agents. Each agent can share a room with any other agent, has a strict preference relation over potential roommates, and does not care about which room she is assigned to (only the roommate matters). If we want to allow agents to have any (rational) preference relation over potential roommates, this model does not belong to the class of supply chain models with same side substitutable and cross side complementary preferences: In order to write this problem as a supply chain model, we would have to define a directed graph of potential trading relationships. In order to allow all potential roommate combinations we would need to introduce an arbitrarily directed edge between all pairs of agents. It is easy to see that if $n \geq 2$, at least one agent has to be an intermediary if we require the market structure to be free of directed cycles. The preferences of such an agent would be severely restricted by the assumption of CSC: The intermediary would be required to either declare all upstream or all downstream agents as unacceptable roommates. Given that the direction of the edges introduced is arbitrary, this is not a satisfactory embedding of the roommate problem. If we dispense with the assumption that preferences satisfy the CSC condition this problem does not occur since it is easy to see that SSS does not restrict the set of allowed preference relations. Hence, any roommate problem can be formulated as a supply chain model in which agents' preferences satisfy SSS (but are allowed to violate CSC). Note that since each agent is looking for at most one partner chain stability reduces to pairwise stability and (any) market structure is weakly acyclic. The above theorem then implies that in the roommate problem any pairwise stable matching lies in the core. Since it is a trivial fact that any core matching is pairwise stable in the roommate problem we obtain the following corollary to Theorem 8.

**Corollary 4.** *For roommate problems the set of pairwise stable matchings coincides with the core.*