5-2015

The Right Fit: Improving the Algorithm that Matches Students to Colleges Based on Test Scores

Carolyn A. Shira
College of William and Mary

THE RIGHT FIT:
Improving the Algorithm that Matches Students to Colleges Based on Test Scores

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Arts in Economics from
The College of William and Mary

By
Carolyn Ann Shira

Accepted for Honors
(Honors)

Donald E. Campbell (Economics), Director

Robert B. Archibald (Economics)

Rex K. Kincaid (Mathematics)

Williamsburg, VA
May 6, 2015

THE RIGHT FIT:
Improving the Algorithm that Matches Students to Colleges Based on Test Scores

Carly Shira
Advisor: Dr. Donald. E Campbell
College of William and Mary, Department of Economics

Abstract: Introduced by Gale and Shapley in 1962, the deferred acceptance algorithm has been applied to an array of markets, including that of college admissions (a two-sided matching game assigning students to colleges). In the case where colleges' preferences are completely determined by students' test scores, this research proves that the Student Optimal Deferred Acceptance (SODA) algorithm never rewards a student for doing less well on an exam, i.e. satisfies "*positive responsiveness*." The SODA algorithm is unique in that it yields an outcome that is *stable*, *fair*, and elicits *positive responsiveness* in the scenario of two students, two colleges, and two tests. Further, we compare this algorithm relative to other matching mechanisms, extracting key features that set SODA apart. There is no two-sided stable matching algorithm (one-to-one matching or many-to-one matching) that is strategy proof; however, when the preferences of the college are completely determined by the test scores of the students, regardless of the number of tests, we show that SODA not only satisfies *positive responsiveness*, but that it is also *strategy proof* in the case of any number of students, any number of colleges, and any number of tests. The implications of this are pertinent to the future of the college admissions process.

1- Introduction to Matching and Marriage

1.1 A Brief Introduction to Matching and Current Applications

The moment where a matching mechanism determines one's fate is the moment when the gravity of game theory, mathematical economic modeling, and reality converge. Day to day, "matching" determines who gets what job in the labor market, who will receive a kidney transplant, which child is accepted into a particular charter school, and more theoretically, who marries whom. Matching mechanisms tackle the challenge of the efficient allocation of *discrete units*: men, women, students, schools, jobs, residencies, etc. Certain markets exist in which only one side of the equation has preferences, such as matching donor recipients to kidneys (i.e. kidneys do not have preferences): these games are referred to as *one-sided matching* mechanisms. Other markets exist in which both sides have preferences, such as when matching men and women to be married: such models are called *two-sided matching* games. This research focuses on two-sided matching.

Tracing back to a paper written by David Gale and Lloyd E. Shapley in 1962 that compared the process of students applying to colleges to the a hypothetical situation in which men propose to women, matching has gained both popularity and prominence in modern markets. As mentioned above, these markets are vast and diverse. Matching algorithms have been applied to the *school choice problem*, the *college admissions problem*, kidney transplants, and entry-level medical residency programs; these mechanisms have also been applied to purely theoretical situations ranging from marriage (Gale & Shapley, 1962) to sorority rush (Mongell & Roth, 1991).

The original work of Gale and Shapley, *College Admissions and the Stability of Marriage*, concludes with the following statement:

> *The reader who has followed us this far has doubtless noticed a trend in our discussion. In making special assumptions needed in order to analyze our problem mathematically, we necessarily moved further away from the original college admission question, and eventually in discussing the marriage problem, we abandoned reality altogether and entered the world of mathematical make believe.*
>
> (Gale & Shapley, 1962)

This reflection has been transformed into a call to action. While the matching algorithm originally explored by Gale and Shapley was only applicable in the land of mathematical make-believe, matching has transcended "make-believe" into reality – into markets with significant implications for and effects on society's welfare.

However, even before Gale and Shapley published this paper opening the floodgates for academic literature on matching, residency programs matching medical students to hospitals were utilizing an algorithm called NIMP (now called NRMP), National Intern Matching Program (or presently, National Residency Matching Program). Used for the first time in 1951 and remaining much the same today, the algorithm is a voluntary process in which interns apply to a central clearing house, rating the hospitals in order of their preferences. The hospitals also rate the students in order of their preferences, subsequently submitting them. The central clearinghouse then takes these preferences into account to create a match – or a resident to hospital pairing. Each resident can only be paired to one hospital, and each hospital can be paired with a number of students not exceeding the hospital's capacity. Despite the fact that this is a voluntary process, the majority of potential residents seeking hospitals decide to apply. Within the first few years of its implementation, there was a participation rate of over 95% in the procedure (Roth and Sotomayor, 1990). While this rate has since dropped a few percentage points, the sweeping participation not only demonstrates the widespread effect that matching algorithms have had on society, but also exemplifies that potential participants are willing to put their futures in the hands of a central clearing house. This systematic approach must, however, respond to the demands of the demographic it serves. In recent years, NRMP has been modified to account for married couples and has instituted more elaborate tie-breaking measures (*Couples in the Match*, 2015).

NRMP does not stand alone in its practical use. School Choice is another prominent example of the application of matching mechanisms. School choice emerged out of America's racial historical context and organization of public schools. The neighborhood in which a student lives largely determines where that student will attend public school. The quality of that public school, however, is also largely dependent upon the neighborhood in which it is located. Thus emerges the problem of systematic socio-economic segregation in the public school system and the correlated issue of de facto racial segregation. Public education faces the challenge of the opportunity gap: the disparity in the resources allotted and achievement of different demographic groups. The Great Schools Partnership defines the opportunity gap as "the ways in which race, ethnicity, socioeconomic status, English proficiency, community wealth, familial situations, or other factors contribute to or perpetuate lower educational aspirations, achievement, and attainment for certain groups of students" (*Opportunity Gap*, 2013). The opportunity gap often refers to inputs: "the unequal or inequitable distribution of resources and opportunities" (*Opportunity Gap*, 2014). One way to counteract the detrimental effects of disparities in resource endowments is to cross boundary lines, and parents noticed. In the storm of education challenges, parents demanded a greater voice in choosing where their child attends public school. School choice is one result among numerous political prescriptions attempting to solve educational inequality.

In the Boston and New York Public School systems, much like other school choice programs around the country, the original system for assigning students to schools via the school choice program schools was highly inefficient, often manipulated by students and parents (unsuccessfully), and failed to fulfill the objective of equal education opportunity (Ehlers & Klaus, 2012). As a result, in 2003, the New York City Department of Education contacted Alvin Roth, Harvard professor and economic specialist in matching games, requesting that he design a mechanism similar to NIMP to better the student assignment process of school choice. Mirroring this, Atila Abdulkadiroglu, a professor then at Columbia University and now at Duke, and Tayfun Sönmez, a professor at Harvard University at the time and now at Boston College, wrote an article published in the Boston Globe that explored the shortcomings of Boston's student assignment mechanism. This article initiated the partnership between the school system and the two professors, as the Boston public school system ultimately consulted Abdulkadiroglu and Sömnez to design a new matching mechanism for the school choice program. (Abdulkadiroglu, 2013).

School choice matching mechanisms are under constant evaluation in the search for improvement and adaptation. Abdulkadiroglu, Pathack, Roth, and Sömnez (2006) and Pathak and Sömnez (2008) demonstrated that the improved Boston mechanism, designed by Abdulkadiroglu and Sömnez, was subject to manipulation by parents, and it often benefited "sophisticated agents" (Ehlers & Klaus, 2012). In practice, the mechanism may favor a certain student demographic over another because "naively truth-telling students (or parents) tend to be the worst off students under the Boston mechanism, which is *not* strategy proof" (a concept that will be discussed in depth at the end of Chapter 1) (Ehlers & Klaus, 2012). With the gravity of student futures resting in the hands of an algorithm, school choice continues to be a controversial topic.

Similar to the school choice problem, the *college admissions problem* is an abstract representation of a real world decision. College admissions are a two-sided matching game that takes into consideration preferences of both the students and the colleges. *Preferences*, as the name suggests, are the order in which each ranks his/her/the institution's possible matches—specifically, the students that the school wishes to enroll and the schools that the student wishes to attend. In the modern college admissions process, schools prefer some students to others because of residency, GPA, standardize test scores, starting a non-profit at age eight, or other admirable characteristics deemed necessary for a dynamic incoming class. For the purposes of this research, however, we assume that the preferences of the schools are *completely determined by the test scores of the students.* Students, on the other hand, prefer some schools to others because of specialty, prestige, location, legacy, cost, social opportunities, and/or the host of other attractive college features. This research takes into account the flexibility of those student preferences by using students' stated preferences – rather than arbitrary ranking of colleges, for instance. Similar to the National Residency Matching Program, students can be paired with only one college, and colleges can be paired with many students. Thus, the college admissions problem is a many-to-one, two-sided matching game. Before delving into the specifics of this problem, it will be useful to introduce the basics of the mathematical approach to matching using Gale and Shapley's theoretical framework via the marriage problem.

1.2 Exploring The Marriage Problem

1.2.1 Meet your Match – An Introduction to Marriage

Imagine 20 women and 20 men in small, isolated village. (Apologies for the hetero-normativity of this example.) Each man needs to find a suitable wife, and each individual has to be happy enough such that he or she is not going to leave his or her partner for another. Is there a way to match each man to a woman so that the arrangements are stable? In other words, is there an algorithm that could be used to "successfully" match men and women in this thought experiment? Referred to as the marriage problem, this is an example of a one-to-one, two-sided matching market, explored originally in a paper by Gale and Shapley (1962). Two sets of agents are at play: men and women. Each man and each woman have preferences over the potential partners, and each person can ultimately be matched with only one partner of the opposite gender or remain unmatched if they prefer.

The theorems introduced in the original paper have opened the door to the concept's application in a range of markets and theoretical applications. As Roth and Sotomayor state, it is "helpful to remember that much of our interest in this problem is motivated by labor markets, rather than by marriage in its full human complexity" (Roth & Sotomayor, 1999). In this case, the examination of marriage, devoid of its "full human complexity", is motivated by the desire to investigate the *college admissions* problem.

We will introduce the basics of a matching, the properties of stability and optimality, and the deferred acceptance algorithm under the conceptual framework of the marriage model before exploring specific applications of these concepts through the use of the case of two students, two colleges, and two test scores in Chapter 2 focusing on two algorithms: Student Optimal Deferred Acceptance (SODA) and College Optimal Deferred Acceptance (CODA). In addition, we will summarize select theorems fundamental to the positive responsiveness proof presented in Chapter 3. Chapter 4 serves as a comparative study of other existing algorithms, delving into why certain properties are desirable, the limitations to certain assumptions, and ultimately how other existing and hypothetical algorithms violate one or more of the desirable properties that SODA satisfies.

1.2.2 The Nitty Gritty of the Marriage Problem

Gale and Shapley (1962) define the marriage problem as consisting of two finite and disjoint sets of men and women, $M=\{m_1, m_2, m_3, \ldots, m_n\}$ and $W=\{w_1, w_2, w_3, \ldots, w_n\}$, respectively. Just as on our isolated island, each man has preferences over the women, and each woman has preferences over the men. Certain assumptions need to be made about individual preferences. First preferences are *strict*: a woman is never indifferent between two alternative men, and a man is never indifferent between two alternative women. (This assumption can be relaxed, but it is maintained throughout this thesis for expositional purposes.) There will always exist a distinguishing feature that a female prefers about one man to another, and vice versa. Each man (or woman) thus has a distinct rank in the woman's (or man's) preference profile. This assumption is fairly realistic considering that preferences are a "knife-edge" phenomenon, where the slightest disparity in desirable characteristics may determine an individual's distinct rank (Roth & Sotomayor, 1999). In addition, preferences are *transitive*. For example, if a woman prefers m_1 to m_2 and m_2 to m_3, then she must prefer m_1 to m_3. In other words, the lower down a man is ranked on a woman's preference profile, the less preferable he is, and if he is ranked third, then the woman cannot prefer him to the man ranked first. Again, this applies to both men and women. Finally, preferences are *complete*. Upon having to make the decision between two men, the woman is always able to determine whom she prefers; she is never unable to make a choice. In other words, given two alternatives, excluding the situation in which in which the woman prefers to remain unmatched, she is able to rank any given man (and vice versa for the preferences of men). These lead to the assumption that preferences are *rational*, in the sense that they are "*acyclic*" (Roth & Sotomayor, 1999). Strength of preferences is not taken into account –for the purposes of this research we simply account for the order in which each agent lists preferences. In addition, each agent only cares about his or her own matching; individual utility is detached from the outcome of other players in the game. What I will call *comparative wellbeing* is not a factor in individual preferences.

A *matching* α, or the outcome of the game, consists of the pairs of men and women, such that $\alpha = \{(m_1,w_1), (m_2,w_2), \ldots\}$, and each pair is a *match*. In other words, a matching is the group of bilateral pairings, or matches, between a man and a woman. Each man can be paired to only one woman, and each woman can only be paired to one man. However, a man or woman can decide that he or she would rather remain unmatched, or *single*. If a woman would prefer to remain single than to be matched with a particular man (…ouch) then that man is *unacceptable* to that particular woman. Conversely, a man is *acceptable* to a woman if she prefers being with him than to being unmatched. The same is true for the preferences of men over women. (Campbell, 2006)

If there exists a man and woman who mutually prefer each other to the woman and man, respectively, with whom they are currently matched, then the pair can upset the matching by *blocking*. If no pair wants to *block*, then that matching is *stable*. In other words, a *stable matching* is one in which there exists no woman/man pair that mutually prefer each other to his or her current match. More formally, there is no match such that w_a would rather be paired with m_a than m_b with whom she is currently matched, and m_a would rather be paired with w_a than w_b, his current match. In this example, m_a and w_a would block, and the matching would not be stable. Given stability is a desirable property, it is important to prove that stable matches are guaranteed for any group of men and women. In order for a mechanism to satisfy stability, the outcome must be stable for any specification of agent preferences. (Campbell, 2006)

Theorem 1: (Gale and Shapley) *There always exists a stable set of marriages.*

In order to prove Theorem 1 for any marriage market, Gale and Shapley put forth a procedure that always elicits a stable matching. The rules of this matching algorithm, coined the *deferred acceptance algorithm*, are as follows:

Each woman in the set W proposes to her most preferred man from the set M, the man who ranks first in her preference ordering (for this example, we will be in a progressive society in which women propose to men.) Each man subsequently rejects or accepts the proposal. If a man receives more than one proposal, he accepts the proposal from the most preferred woman according to his preference ordering. If a man finds his only proposer unacceptable, he can decide to remain unmatched for the first round. The pairs that remain at the end of stage one are "engaged". In other words, each woman whose proposal was not denied during the first round is *provisionally matched* with her most preferred man. (Please note that I will be using the terms "stage" and "round" interchangeably.)

The women who were rejected in the first round then propose to the man second on their preference orderings, the man next preferred to the one whom they proposed to in round 1. Each man, once again, accepts the proposal from the woman he most prefers (choosing from a group consisting of new proposers and the woman with whom he is provisionally matched), and he is then provisionally matched with her for this round. A man remains unmatched if he has rejected all proposals in round 1 and round 2.

The algorithm follows in this fashion, with women proposing to their next choice in the subsequent round if they are rejected or remaining provisionally matched if they are the most preferred of the group of proposers. In every round, each man rejects the proposals of unacceptable women and rejects all proposals other than that of his most preferred woman of the set of proposers and his provisional match from the previous round. A male and female pair are engaged if they are provisionally matched at the end of a round. A man remains unmatched if he has rejected all proposals in all previous rounds.

The algorithm terminates when no woman is rejected. At this point each woman is in one of two situations: matched with a man (she is engaged) or rejected by all men that she finds acceptable (she is single). Each man has accepted a proposal in some round of the algorithm (he is engaged) or has rejected all proposals (he is single) or has received no proposals. The engagements, or provisional matches, are then consummated—resulting in a stable set of marriages.

Figure 1: *Female Optimal Deferred Acceptance Algorithm: A Woman's Path to Marriage*

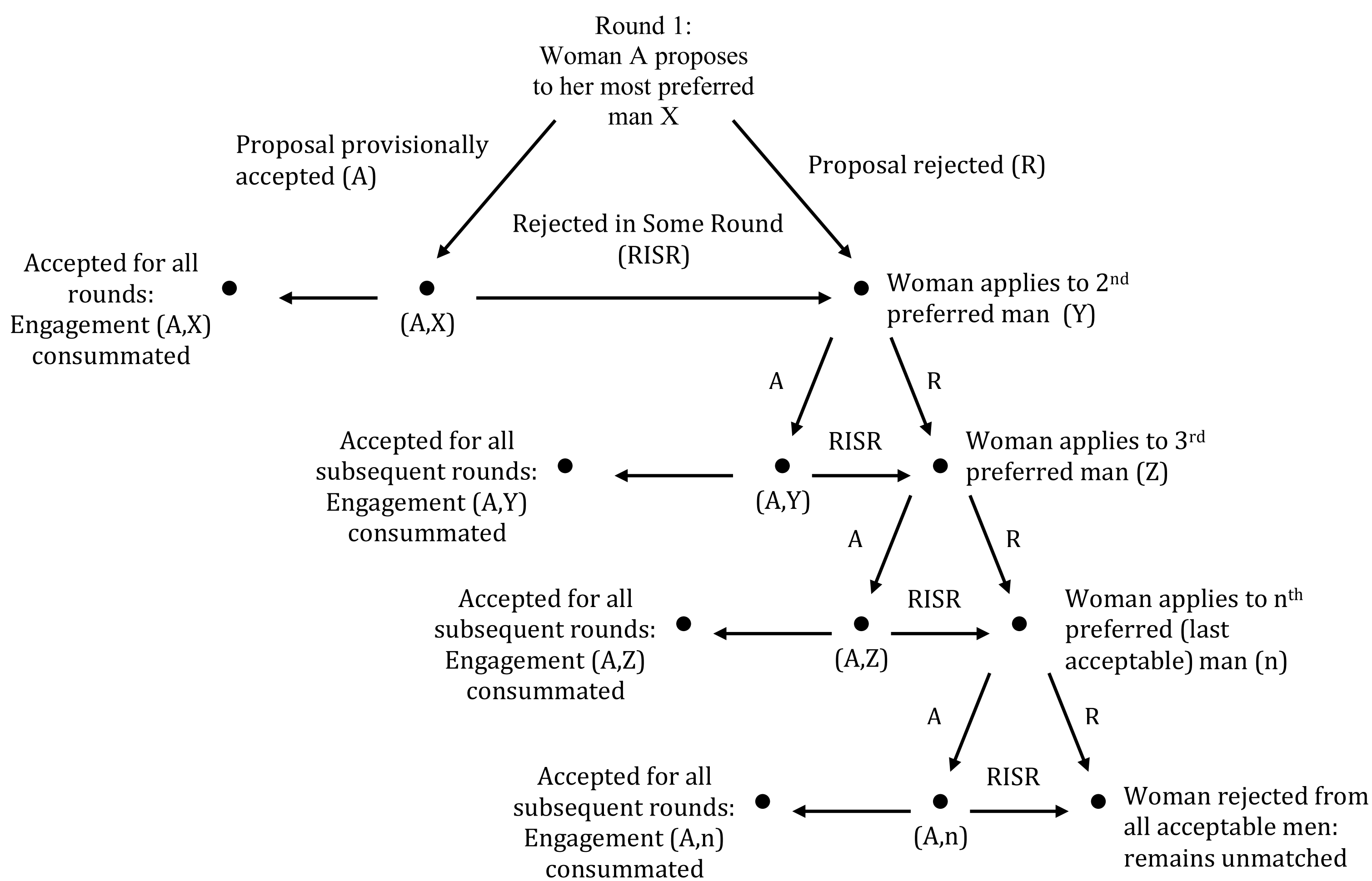


Figure 1 is a visual representation of how the deferred acceptance algorithm proceeds in practice. At the start of the procedure, the woman proposes to her most preferred man. She then faces two possibilities: being accepted or rejected. If she is rejected, she proposes to her next preferred man. If she is accepted, she either remains accepted in all additional rounds (denoted by the horizontal arrow pointing to the left) or is rejected in some subsequent round. This rejection is represented by the right-facing, horizontal arrow leading to the node in the second layer of the tree. From here the woman faces the same options of acceptance or rejection, then continued acceptance or rejection in some subsequent stage of the algorithm. Once again, if she is rejected at a later stage of the algorithm, she follows the horizontal arrow pointing to her next action: proposing to the man next on her preference list.

A striking phenomenon results from this structure: the algorithm favors the proposing agents over the recipients of proposals (Gale and Shapley, 1962). If women are proposing, then no woman strictly prefers any other stable matching to the one that results from the algorithm. If men are proposing, then no man strictly prefers any other stable match. This is the concept of optimality (*School Choice*, 2014).

Definition (Roth and Sotomayor): *For any given marriage market, a stable matching is male* ***optimal*** *if every man likes it at least as well as any other matching. A stable matching is female* ***optimal****, if every female likes it at least as well as any other matching.*

Depending on the set of women and men and their preferences, a number of stable matchings are possible. If women and men have strict preferences, then each member of the group that proposes is matched with his or her "most preferred *achievable*" partner, where an agent of the opposite type is *achievable* if he or she is a partner in some stable outcome (Roth & Sotomayor, 1999). This leads to the understanding that there can only be one optimal matching for each side; only one male-optimal stable matching and only one female-optimal stable matching exist.

Theorem 2 (Gale and Shapley): *When all men and women have strict preferences, there always exists an M-optimal stable matching, and a W-optimal stable matching. The W-optimal stable matching is the matching α_w produced by the algorithm when the women propose. The M-optimal stable matching is the matching α_M produced by the algorithm when the men propose.* (Roth and Sotomayor, 1999)

It is convenient to conceptualize optimality as stable matchings on opposite ends of the "favorable" spectrum. Women, as a group, enjoy the set of stable matches produced by the algorithm when they propose, and vice versa for men. In addition, the stable matching that women most prefer just so happens to be the stable matching that the men least prefer. Roth and Sotomayor provide a corollary to this theorem, originally proposed by Knuth, proving that "when all agents have strict preferences, the M-optimal stable matching is the worst stable matching for the women; that is, it matches each woman with her least preferred achievable mate. Similarly, the W-optimal stable matching matches each man with his least preferred achievable mate" (Roth & Sotomayor, 1999). Conceptually, if stable matchings were represented by utility functions, there would be an inverse relationship between the utility of men and women according to what faction proposed. The deferred acceptance algorithm acts as a short cut to find the stable matching that is liked at least as much as any other stable matching by the proposing set of agents. (*School Choice*, 2014).

The above assertion results in an incentive for the recipient to manipulate the game in order to obtain a more favorable matching. That is, the deferred acceptance algorithm guarantees a stable matching when both parties act according to their true preferences, but it is not *strategy proof*. This captured by Roth's (1985) impossibility theorem:

Theorem (Roth): *"No stable matching mechanism exists for which stating the true preferences is a dominant strategy for every agent"* (Roth and Sotomayor, 1999).

If men and women were to submit their preferences to a match-maker, the proposers would have no incentive to manipulate the system because the deferred acceptance algorithm guarantees that each member of this group gets their most preferred achievable partner. However, if the women were proposing, a man may benefit (by getting a more preferable matching) by deviating from his true preferences. An algorithm is *strategy proof* if neither party has an incentive to deviate from their true preferences, or more formally, "a matching mechanism will be called *strategy proof* if it is a dominant strategy for each player to state his or her true preferences in the strategic game it induces" (Roth & Sotomayor, 1999). In Chapter 3, this paper will present an existing theorem that no proposer can deviate from truthful revelation under the requirement of satisfying stability. This will be crucial to the proof that the student optimal deferred acceptance algorithm never rewards a student for doing worse on an exam. In order for a game to be entirely "strategy proof," however, neither the proposer nor the recipient can have the incentive (or, in certain cases, the capability) to misrepresent preferences.

1.3 Where We Go From Here

The theorems and assumptions about the deferred acceptance algorithm addressed in the marriage problem, a one-to-one matching game, are in many ways directly transferable to the college admission problem, a many-to-one matching game. In Chapter 2, we will briefly disclose certain assumptions and theorems necessary to continue for the college admissions problem before delving into two approaches to deferred acceptance.

SODA, student optimal deferred acceptance, and CODA, college optimal deferred acceptance, are the focus of Chapter 2. Using CODA, each college "proposes" to the institution's most preferred student in the first round. If the college is "*acceptable*" to the student then the student accepts the proposal, and a match is *provisionally* made. If the student gets more than one proposal, then the student accepts the preferred college, and the college extends a proposal to its next preferred student in the next round. The algorithm terminates when each student is either matched with a college or has rejected all proposals. Correspondingly, using SODA, the students do the proposing. All students "propose" to their most preferred college at the beginning of round 1. A college can accept students up to its capacity and will accept those with the highest test scores (specific to that college's test). If a student is denied from a college, she proposes to her next ranked school in the following round. In Chapter 2, we will use the special case in which there are two students, two schools, and two tests to introduce a basic example of how CODA can reward a student for doing worse on an exam. That observation was the motivation for the present honors thesis.

Chapter 2 – The 2x2x2 Case

2.1 Defining the College Admissions Problem

As stated in the introduction, the college admissions problem is a two-sided, many-to-one matching game that takes into account the preferences of both colleges and students. There are two finite and disjoint sets of students and colleges, $S=\{s_1, s_2, \ldots, s_m)$ and $C=\{c_1, c_2, \ldots, c_n)$, respectively. A *matching* π is the outcome of a game such that each college is matched with no greater than its capacity of q_C students, and each student is matched with one college. A *match* is bilateral pairing of a student to a college. (If $|q_C| > 1$, colleges will have multiple matches in the outcome, where students will have only one match.) Colleges have preferences over students, as students have preferences over colleges. Each college's preferences are based solely on students' test scores on the specific test used by that college, with the student who receives the highest test score ranking first, the student who receives the second highest test score ranking second, and so on, down to the student who receives the lowest acceptable test score ranking last. T_C denotes a test used by a certain college C.

We need to impose certain assumptions about the preferences of each party. Just as in the marriage problem, the preferences of both colleges and students are *strict*, *transitive*, and *complete* (Roth & Sotomayor, 1999). A student is never indifferent between two alternative schools, and school is never indifferent between two students. (The indifference assumption can be relaxed, but we maintain it here to keep the analysis relatively simple.) Because test scores determine college preferences, a corresponding assumption is necessary: no two students receive exactly that same test score, thus allowing for distinct ranks in each college's preference ordering. Students' preference orderings of colleges are determined on a subjective basis. In addition, if a student ranks college X first, college Y second, and college Z third, then it is not possible for the student to prefer college Z to college X. Similarly, if a college ranks student A first, B second, and C third, the college cannot prefer C to A. Complete preferences take on a slightly nuanced definition; among the colleges that the student finds acceptable, the student is always able to make a decision about which college she prefers given two alternative schools – there is no college left behind, or unevaluated in this situation. (This can also be relaxed to include the case of a student who would rather not attend college than be enrolled in college C.) Mirroring this, colleges are able to rank any given student who takes the entrance exam used by that college. The combination of the assumptions that preferences are *strict*, *transitive*, and *complete* comprises the assumption that preferences are *rational* (Roth & Sotomayor, 1999). Once again, we assume that *comparative wellbeing* holds no weight in determining the preferences of either the schools or students. In other words, the student only cares about his assignment – not his assignment relative to another student's assignment, and a school only cares about its incoming class – not its incoming class in comparison to that of a competitive institution.

In addition to the classic economic assumptions about each party's preferences, we make certain implicit assumptions. For instance, do expressed student preferences in fact promote student welfare? Creating an informed personal preference ordering requires accurate self-evaluation. Depending on the perceived vs. actual skill-set of a student, what if the student would be more successful if he were assigned a school other than his top choice (or other than a college which he finds acceptable)? We ignore this possibility; we assume that satisfying the expressed preferences of students promotes student welfare.

Unlike the marriage problem, colleges have an incoming *class*, not an incoming *student*. The many-to-one aspect of the college admissions problem raises an important question as to whether preferences over individual students (which are equivalent to preferences over test scores) translate to preferences over a desirable incoming class (a conglomerate of students) (Roth & Sotomayor, 1999). For the purposes of this paper, we use colleges' preferences over individuals as representative of their preferences over groups of individuals. Specifically, we assume the *replacement property*. If the set G′ of students can be obtained from G by removing some student g from G and replacing her with g′, then a college will prefer G′ to G if and only if it prefers g′ to g. (Note that a student's preferences are complete if for any two acceptable colleges C and D, either C is preferred to D or D is preferred to C. The same is true for college preferences. Further, a college's preferences over groups of students need not be complete, even if it has complete preferences over individuals.) (Roth & Sotomayor, 1999).

A student is *acceptable* to the college if the college prefers that student to leaving an empty seat. A college is *acceptable* to a student if that student prefers the college to remaining unmatched. If an agent of the opposite type is unacceptable, then a student does not include that college on her preference ordering, and a college does not include that student on its preference ordering (Campbell, 2006). From the perspective of a college, a test threshold determines whether or not a student is acceptable. An acceptable student must receive a test score of at least the test threshold of $T_C(0)$ for a given college C. Throughout this paper, we assume that the threshold is a test score of 0 for all colleges; if the student takes the test used by a certain college, he or she is automatically deemed acceptable. In other words, if a student can write her name on the test, she is good to go (setting an impressively high standard).

Given a certain outcome, if there exists a student-college pair that mutually prefers each other to their current matches, they can upset the outcome by blocking (Campbell, 2006). If no such pair exists, the outcome is stable.

Definition: *An outcome is **stable** if there is no match such that some student A would rather be paired with college X than college Y, with which she is currently matched, and college X would rather be paired with student A than some student B, who is currently assigned to X; A and X would block. If there is not such blocking pair, regardless of the specification of student and college preferences, then we say the mechanism is **stable**.*

Dealing with the case of a single student-college pair, the above definition of stability is specifically "pair-wise" stability; however, a coalition of students and colleges may be able to benefit by acting outside of the mechanism. If no coalition is able to benefit by blocking, the outcome of a particular algorithm is considered *group stable*. For the purposes of this research, however, we focus on pair-wise stability because satisfying group stability is contingent upon satisfying pair-wise stability (Roth and Sotomayor, 1999).

Gale and Shapley, the first to draw an economic comparison between marriage and college admissions, expanded the deferred acceptance algorithm from the marriage model to the college admissions problem. Not only proving that the deferred acceptance algorithm always yields a stable outcome, but also proving that deferred acceptance yields "an optimal assignment of applicants," we refer to this as student optimal deferred acceptance or SODA (Gale & Shapley, 1962).

Theorem (Gale and Shapley): *"Every applicant is at least as well off under the assignment given by the deferred acceptance procedure as he would be under any other stable assignment."* (Gale & Shapley, 1962).

Given a set of students and colleges, a number of stable outcomes may exist. Similar to the marriage problem, the deferred acceptance algorithm acts as a short cut to find either the student optimal or the college optimal stable matching, depending upon who "proposes." If a number of stable matchings exist, they exist on a continuum favoring either colleges or students. The student optimal and college optimal matchings exist on opposite ends of this continuum, where the best matching for students is systematically the worst matching for colleges (*School Choice*, 2014). From this we can deduce that if the college optimal and student optimal stable outcomes are the same, there must be only one stable matching.

Definition: *An outcome π is considered student* ***optimal*** *if there is no other stable matching that some student prefers to π.* Similarly, an outcome *π* is college* ***optimal*** *if there exists no other stable matching that some college prefers to π*.* (Campbell, 2006)

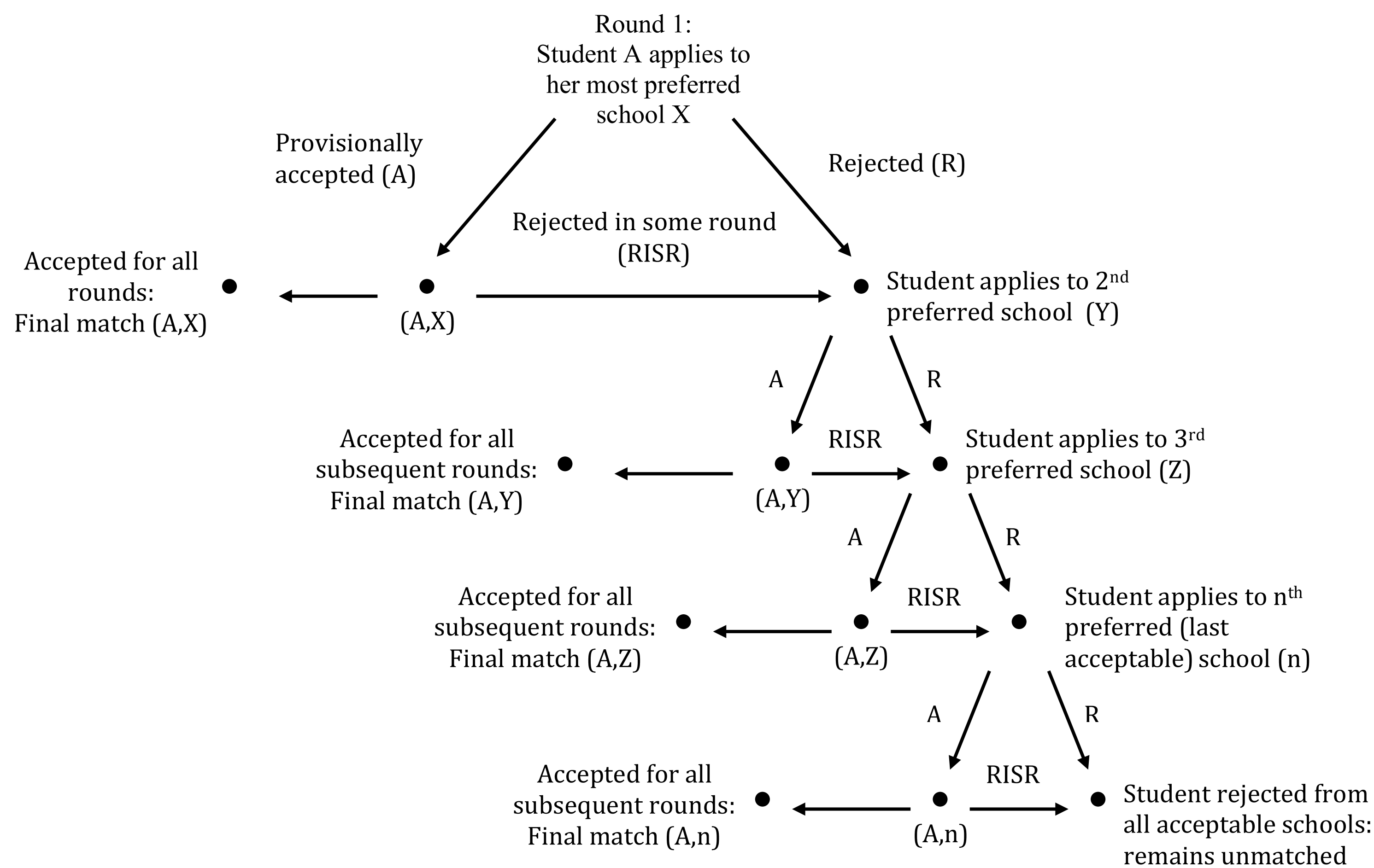
In section 2, we map the path of a student on the way to college acceptance (Figure 2.1) mirroring the path of a woman on the way marriage (Figure 1). Section 3 follows this explanation with a simple example of both CODA and SODA with certain desirable properties such as *fairness* and *responsiveness* embedded within the example. Section 4 provides a proof that in the 2x2 case, any mechanism that yields an outcome that is both responsive and fair must be equivalent to SODA.

2.2 SODA and CODA in Action

SODA proceeds as follows. In round 1, each student in the set S proposes to his or her most preferred school from the set C. Each school accepts the proposals of the students with the highest test scores up to its capacity of q_C students. If the school accepts the proposal, then the school and the student are provisionally matched. If the school rejects the proposal (because it has reached its capacity and the proposing student has a lower test score than all other students that have been provisionally accepted by that college or because the college finds the student unacceptable), the student then proposes to her second preferred school in the following round. Once again, the school either accepts her proposal (if she has one of the q_C highest test scores of the group of proposing students comprised of the school's previous round of provisionally matched students and new applicants) or rejects the proposal because she does not have one of the q_C highest test scores. A given college can have empty seats if it receives fewer proposals than its capacity in rounds 1 and 2 or if it receives fewer than its capacity of acceptable proposals in rounds 1 and 2. (We assume that all students who apply are acceptable because they have a test scores greater than the test threshold of 0.)

The algorithm proceeds in this fashion, with students proposing to their next choice in the subsequent round if they are rejected or remaining provisionally matched if they are among the students with the q_C highest tests scores of group of proposers. In every round, schools reject the applications of unacceptable students and reject all proposals of students beyond those with the q_C highest test scores on the exams specific to each college. A student whose proposal has been accepted and the school that accepted are provisionally matched at the end of a round.

Figure 2.1: *Student Optimal Deferred Acceptance Algorithm: The Path of a Student*



The algorithm terminates when no student is rejected. A student will either be provisionally matched with a college at this point or has been rejected by all colleges that she finds acceptable. The provisional matches become final student-to-college assignments. Figure 2.1 follows the path of a student through an application of SODA to the college admissions problem.

CODA is the converse; colleges offer admission. In the first round, each college proposes to its most preferred students up its capacity. Each student accepts or rejects the schools' proposals. If a student receives more than one proposal, she chooses her most preferred school and rejects all other proposals. If a student finds her proposer(s) unacceptable, she can remain unmatched for the first round. A student who accepted a proposal and the college that proposed are provisionally matched at the end of round 1. After the first round, if a college has open seats, it offers admission to the next highest ranked students up to its capacity. A student once again accepts the proposal of its most preferred school and rejects all others. And so on. The algorithm terminates when no college is either rejected or has exhausted its list of acceptable students.

2.3 A 2x2 Example

In the following example, there are two students and two colleges: S={A,B} and C={X,Y}. Each college has a capacity of one student, $|q_X| = 1$ and $|q_Y| = 1$. College X admits students based on their quantitative test score, while college Y uses the verbal score. Student A prefers college Y to X, and student B prefers X to Y. The table below lists the test scores of A and B on T_X and T_Y (note that T_X and T_Y can be different components (e.g. math and verbal) of the same test weighted differently into the admissions decisions of various colleges):

Figure 2.2: Scenario 1

Student	Quantitative Score (T_X)	Verbal Score (T_Y)
A	80	60
B	70	90

Given the test scores and preferences of the students, the following preference orderings emerge:

A	B
Y	X
X	Y

X	Y
A	B
B	A

Applying CODA, in round 1, college X offers admission to A, and college Y offers admission to B. Both students accept the proposals because each finds both schools acceptable, resulting in provisional matches (A,X) and (B,Y). The algorithm terminates as no student rejects any college's proposal. The matches (A,X) and (B,Y) are finalized: π_{CODA} = { (A,X) , (B,Y) }.

Applying SODA, in round 1, student A proposes to Y, and student B proposes to X. Both colleges accept the proposals resulting in the provisional matches (A,Y) and (B,X). The algorithm terminates as no college rejects any proposal: π_{SODA} = { (A,Y) , (B,X) }. Note that both students prefer π_{SODA} to π_{CODA}, and both colleges prefer π_{CODA} to π_{SODA}.

Figure 2.3 lists the test scores of A and B on T_X and T_Y where the only difference is that student B receives a lower verbal score.

Figure 2.3: Scenario 2

Student	Quantitative Score (T_X)	Verbal Score (T_Y)
A	80	60
B	70	50*

The change in B's verbal score is drastic enough such that both colleges now prefer student A to student B. The preferences of the students and colleges are as follows:

A	B
Y	X
X	Y

X	Y
A	A
B	B

Applying CODA first, colleges X and Y propose to A in round 1. A accepts her most preferred college, Y, and rejects X. The provisional match (A,Y) is made. In round 2, college X proposes to its next ranked student B who accepts the proposal, resulting in the provision match (B,X). The algorithm terminates as no college is rejected. The provisional matches are finalized, and with student A assigned to Y and student B assigned to X: π^*_{CODA} = { (A,Y) , (B,X) }.

Using SODA, the process is exactly the same as when student B receives a higher test score. Each student applies to her most preferred college who accept the proposals, and the algorithm terminates as π^*_{SODA} = { (A,Y) , (B,X) }.

Scenario 2 presents an issue. By receiving a lower test score on T_Y, student B is rewarded by getting assigned to a more preferred college X. This violates responsiveness:

Definition: *A mechanism that assigns students to colleges is* ***responsive*** *if it never assigns a student to a college that she prefers to the one that she would have been assigned if she had earned a higher score on any single exam, assuming no change in any student's preferences and no change in any other test score.*

The following section proves that SODA always yields an outcome that is both responsive and fair. Some responsive mechanisms exist where there is never a case in which a student can get assigned a more preferred college if she gets a higher test score (such as the mechanism that ignores test scores entirely). SODA, however, sometimes assigns a student to a better college when she gets a high score: this is the distinction between responsiveness and positive responsiveness.

2. 4 SODA: Responsive and Fair for the 2x2 Case

Definition: *When test scores generate college preferences, an assignment is* ***fair*** *if we cannot find a student S and a college C such that S prefers C to the college assigned to S, and S has a higher score on the test used by C than some student assigned to C.*

Note that if we treat each college's ranking of students by the students' test scores on the test used by that college as preferences, then fairness and stability are identical: an outcome is stable if there exists no student-school pair that prefer each other to their respective current matches. Formally, no student-school pair exists such that student B prefers college X to her current school assignment, and college X prefers student B to a student who is currently assigned to X. Student B and college X would block, upsetting the algorithm, thus violating stability. Stability and fairness are interchangeable when test scores completely determine college preferences.

Gale and Shapley (1962) prove that the deferred acceptance algorithm always results in a stable outcome. Because fairness and stability are interchangeable when student test scores determine college preferences, this theorem can be expanded to include fairness. The deferred acceptance algorithm, thus, always yields a fair outcome.

Definition: *Two mechanisms are considered* ***equivalent*** *if they always deliver the same outcome for any given pattern of individual test scores and college preferences.*

The following proves that in the case where there are two students, two colleges, and two tests, a mechanism that yields a stable, fair, and responsive outcome is equivalent to SODA. This proof sets the conceptual stage for the generalizable proof in Chapter 3 that addresses SODA's responsiveness with any number of students, colleges, and tests.

Lemma: In the 2x2 case, there is only one fair matching if and only if there is a student-college pair such that each ranks the other first.

Proof:

(i) Suppose that student A ranks Y first, and college Y ranks A first because A earned a higher T_Y score than B. Then if A is not assigned to Y the outcome is not fair. That leaves only one possible assignment: (A,Y), (B,X). This is fair because, even if B prefers Y to X, she cannot claim to have a higher T_Y score than A.

(ii) Suppose there is no student-college pair such that each ranks the other first. Suppose that A ranks Y first. Then Y ranks B first. Then B must rank X first. It follows that X ranks A first. The following pattern of preferences emerges:

A	B		X	Y
Y	X		A	B
X	Y		B	A

Note that both (A,X), (B,Y) and (A,Y), (B,X) are fair. □

Theorem: For the 2x2 case, any mechanism that yields an outcome that is **responsive** and **fair** is **equivalent** to SODA.

Proof:

If there is only one fair outcome, then every fair mechanism delivers that outcome. Because SODA is fair (stable), then when there is only one fair assignment, all fair mechanisms yield an outcome equivalent to that of SODA.

Suppose there is more than one fair assignment. In the 2X2 case, this means that there are exactly two fair assignments, (A,X), (B,Y) and (A,Y), (B,X). Then we cannot have both students preferring the same college, and we cannot have both colleges ranking the same student first. In other words, in order to have more than one fair assignment, no student-school pair can mutually prefer each other.

Suppose without a loss of generality, A ranks Y first and B ranks X first. If X ranks B first and Y ranks A first, then there is only one fair assignment: (A,Y), (B,X). Therefore, as the lemma above proves, if there are two fair assignments, then we must have this pattern:

A	B
Y	X
X	Y

X	Y
A	B
B	A

If the mechanism yields (A,Y), (B,X), then it agrees with SODA. Suppose that a fair mechanism yields (A,X), (B,Y) in this case.

Now, suppose B had earned a lower score on T_Y such that B's position in Y's ranking changes. The colleges' new preference orderings are:

X	Y
A	A
B	B

If the mechanism is responsive it cannot reward B for getting a lower score, and hence the assignment remains (A,X), (B,Y). We have:

A	B		X	Y		
Y	X		A	A	$\longrightarrow$	(A,X), (B,Y)
X	Y		B	B		

But this is not fair: A prefers Y to X and gets a better Y-score than B. Therefore, in the 2X2 case a fair and responsive mechanism is always equivalent to SODA. □

It is important to note that the above is an "equivalence" theorem. We are *not* claiming that SODA is the only algorithm that yields a stable, fair, and responsive outcome; we instead argue that if another mechanism yields a stable, fair, and responsive outcome, it will be equivalent to the outcome of SODA. Dubins and Freedman (1981) in fact provided an alternative algorithm characterized by the same properties. The following chapter delves into a generalized proof that SODA is responsive for any number of students, colleges, and tests.

3 – SODA is Responsive

When the preferences of colleges are completely determined by the test scores of the students, regardless of the number of tests, we prove that SODA does not incentivize students to perform below capacity and does not reward students for doing less well on an exam. We further propose that SODA is strategy proof in that students cannot benefit by misrepresenting their preferences nor by under-performing on a test in the case of any number of students, any number of colleges, and any number of tests. This chapter explores the above assertions. Sections 3.2 and 3.3 contain the proofs leading to the conclusion that SODA does not reward a student for doing less well on an exam while section 3.4 explores the possibility of strategy proofness.

Why should we care? As we will explore in chapter 4, responsiveness encompasses aspects of both allocative efficiency and fairness, and its presence helps to create a desirable incentive structure. Responsiveness ensures that "better" students go to "better" schools while stripping away the possibility for a more preferred assignment as a result of getting a lower test score. Students, thus, have an incentive to perform to the best of their capability on *every* test.

3.1 Setting the Stage

Definition: *A mechanism that assigns students to colleges is* ***responsive*** *if it never assigns a student to a college that she prefers to the one that she would have been assigned if she had earned a higher score on any single exam, assuming no change in any student's preferences and no change in any other test score.*

There are n students and m colleges. The preferences of the college are completely determined by the test scores of the students—with the student who receives the highest test score on the exam used by that college ranking first, the student who receives the second highest test score ranking second, down to the student with the lowest test score for that college. Let T_C denote the test used by college C, which has room for q_C students. Let α denote a particular application of the SODA algorithm. Then α(1) is the state of SODA in round 1, α(2) is the state of SODA in round 2, and so on. Every college C individually defines $T_C(0)$, where $T_C(0)$ is the lowest test score the college is willing to admit – the threshold – below which the college C prefers to remain unmatched. If the college C prefers to remain unmatched to accepting the student, the student is *unacceptable* to college C. (All students in each round must have test scores greater than $T_C(0)$ in order to be admitted.) If there are less than q_C test scores among the students who applied to C in round 1 that exceed the threshold, $T_C(0)$ then carries over to the next round. This pattern continues as $T_C(0)$ trumps each subsequent test threshold if the college is not yet at full capacity. This may happen as a result of the following two scenarios: less than q_C students apply to college C for every round, or less than q_C students who apply have test scores greater than the initial threshold of $T_C(0)$. Assume $T_C(0) = 0$ for every college C, and assume all students receive a score greater than zero.

Let $P_C(1)$ denote the set of students who apply to college C in the first round. (Then $P_C(1)$ is the set of students who rank college C first.) Let $A_C(1)$ denote the set of students provisionally accepted at the end of round 1. Set $A_C(1) = P_C(1)$ if $P_C(1)$ has no more than q_C members; otherwise $A_C(1)$ is the set of students in $P_C(1)$ with the q_C highest scores on T_C. Then $R_C(1) = P_C(1) \backslash A_C(1)$, the students who are rejected by college C at the end of round 1. Let $T_C(1)$ denote the lowest test score provisionally accepted in round 1. (If there are greater than q_C applicants, $T_C(1)$ acts as the functional threshold for provisional acceptance in the next round.) It follows that $\alpha(1)$ specifies $P_C(1)$, $A_C(1)$, $R_C(1)$, and $T_C(1)$ for each college C. Once a college reaches its full capacity of q_C students in any round, that college remains at full capacity until the algorithm terminates because there are at least q_C students who applied with test scores that exceed the threshold. (If at any stage t $|P_C| \geq q_C$, college C will accept, or be matched with, its capacity of q_C students.)

Given $\alpha(1)$, let $P_C(2)$ be the union of $A_C(1)$ and $R_{C-1}(1)$, where $R_{C-1}(1)$ is the set of students who were rejected in round 1 from the college ranked directly above college C on their preference profiles. $P_C(2)$ is the set of students who were either provisionally accepted at the end of round 1 or who applied to C at the beginning of round 2. Set $A_C(2) = P_C(2)$ if $P_C(2)$ has no more than q_C members; otherwise $A_C(2)$ is the set of students in $P_C(2)$ with the q_C highest scores on T_C. ($A_C(2)$ is the set of students who are accepted by college C at the end of round 2.) In order for a student in set $R_{C-1}(1)$ to gain provisional acceptance into college C, she must have a test score that at least exceeds $T_C(1)$ and that is ultimately equal to or greater than $T_C(2)$. (The lowest test score provisionally accepted in round 2 is $T_C(2)$.) Set $R_C(2) = P_C(2) \setminus A_C(2)$, the students who are rejected by college C at the end of round 2. Then $\alpha(2)$ specifies $P_C(2)$, $A_C(2)$, $R_C(2)$, and $T_C(2)$ for each college C.

Given $P_C(t)$, $A_C(t)$, $R_C(t)$, and $T_C(t)$ we define $P_C(t+1)$, $A_C(t+1)$, $R_C(t+1)$, and $T_C(t+1)$ as follows: $P_C(t+1)$ is the union of $A_C(t)$ and $R_{C-1}(t)$, where $R_{C-1}(t)$ is the set of students who were rejected from the college ranking just above college C in round t on their preference lists. Set $A_C(t+1) = P_C(t+1)$ if $P_C(t+1)$ has no more than q_C members; otherwise $A_C(t+1)$ is the set of students in $P_C(t+1)$ with the q_C highest scores on T_C. In order for a student in $R_{C-1}(t)$ to gain provisional acceptance into college C she must have a test score that at least exceeds $T_C(t)$ and that is ultimately greater than or equal to $T_C(t+1)$. ($T_C(t+1)$ is the lowest test score provisionally accepted.) And $R_C(t+1) = P_C(t+1) \backslash A_C(t+1)$. Similarly, $\alpha(t+1)$ specifies $P_C(t+1)$, $A_C(t+1)$, $R_C(t+1)$, and $T_C(t+1)$ for each college C.

Given the set of students, the preferences of the students, the test scores of the students (determining the preferences of the colleges), the set of colleges, and the capacity of each college, let α_H be the application of SODA. Let α_L be the application of SODA in the identical situation *except* that student B has a lower score on T_C for some college C. This includes the case C=X, where X is the college to which B is assigned by α_L. We prove that if α_H assigns B to college Y, and B prefers X to Y then α_L cannot assign B to X. Contextually, student B cannot benefit from getting a lower test score. We begin by establishing this for the specific case in which X is B's most-preferred college. In order to prove that SODA is responsive, however, we must first prove that truthful revelation is the dominant strategy for students. Section 3.2 explores this existing proof.

3.2 The Dominant Strategy for Students

Theorem 1: *Truthful revelation is the dominant strategy for students using SODA.* (Dubins & Freedman, 1981).

No student will be assigned a preferred college (according to her true preferences) when she misrepresents her true preferences. The above theorem is deduced from the marriage problem. Dubins and Freedman (1981) in conjunction with Roth (1985) found that "the mechanism that yields M-optimal stable matching (in terms of the stated preferences) makes it a dominant strategy for each man to state his true preferences" (Roth & Sotomayor, 1999). The same goes for a W-optimal stable outcome. As colleges cannot misrepresent preferences, we focus on students. Theorem 1 can be generalized: Any mechanism that always yields an S-optimal stable matching always elicits truthful revelation of student preferences. This is crucial for the proof that SODA is responsive.

3.3 Proof that SODA is responsive

The above theorem is integral to the second step in the proof that SODA is responsive. However, we first explore the special case in which X is student B's top ranked institution. With Theorem 1 and Step 1 as platforms, Step 2 proves that a student can never benefit by receiving a lower score on any test using the SODA algorithm.

Theorem 2: *SODA is **responsive**.*

Step 1: If X is B's most-preferred college and α_L assigns B to college X then α_H assigns B to X.

Proof: If α_L assigns B to college X, then X does not reject B in any round of $\alpha_{L.}$ In fact B belongs to $A_X(t)$ in every round t of α_L because B will apply to her most-preferred college in round 1.

The following is a proof by induction. Let Γ denote the set of colleges and Σ the set of students. Every student applies to her most-preferred college at the beginning of round 1 of both α_H and α_L. Therefore, $P_C(1)$ is the same for $\alpha_H(1)$ as for $\alpha_L(1)$ for every college C in Γ. Given that B is accepted into X in $\alpha_L(1)$, B must be accepted into X in

$\alpha_H(1)$ because the only difference in the scores between α_L and α_H is that B's T_X score is higher for α_H, and the colleges preferences are completely determined by test scores. $A_X(1)$ is thus the same set for $\alpha_H(1)$ as it is for $\alpha_L(1)$. If $A_X(1)$ and $P_X(1)$ are identical between $\alpha_H(1)$ and $\alpha_L(1)$, it follows that $R_X(1)$ must be the same set in $\alpha_H(1)$ as $\alpha_L(1)$. If, and only if, student B's test score was the lowest provisionally accepted by X for $\alpha_L(1)$, the functional threshold for college X in round 1 is determined by this score. The functional threshold for college X would then be higher in $\alpha_H(1)$ than $\alpha_L(1)$ (Assuming student B at least surpasses the test score of the student ranked directly above her from α_L to α_H, the ranking of students in $A_X(1)$ will change while the set remains the same). If student B has a test score of anything other than $T_X(1)$ for α_L, the functional threshold remains the same for both $\alpha_H(1)$ than $\alpha_L(1)$. Again, while student ranking in the X's preference ordering changes assuming student B at least surpasses the test score of the student ranked directly above her from α_L to α_H, the set $A_X(1)$ remains the same. For the sake of simplicity, assume student B's test score is never the lowest provisionally accepted. (Student B never determines the functional threshold $T_X(t)$.) For every college C in $\Gamma \setminus\{X\}$ the ranking of the students in $\Sigma\setminus\{B\}$ determined by T_C is the same for α_H as for α_L. Therefore, $T_C(1)$ is the same for α_H as for α_L for every college C in $\Gamma \setminus\{X\}$, and $A_C(1)$ is the same set for $\alpha_H(1)$ as it is for $\alpha_L(1)$ for every college C in $\Gamma \setminus\{X\}$. If $A_C(1)$ and $P_C(1)$ are the same in $\alpha_H(1)$ and $\alpha_L(1)$ for every college C in $\Gamma \setminus\{X\}$, it follows that $R_C(1)$ is the same in $\alpha_H(1)$ and $\alpha_L(1)$ for the same set of colleges. Therefore, we must have $\alpha_H(1) = \alpha_L(1)$.

Suppose that $\alpha_H(\tau) = \alpha_L(\tau)$ for all $\tau \leq t$. We show that $\alpha_H(t+1) = \alpha_L(t+1)$: For every college C, the sets $P_C(t)$, $A_C(t)$, and $R_C(t)$ and the threshold $T_C(t)$ are the same for

$\alpha_H(t)$ as for $\alpha_L(t)$. By hypothesis, B belongs to $A_X(t)$ and to $A_X(t+1)$ for α_L in the case where X is student B's most preferred college. Therefore, B belongs to $A_X(t)$ and to $A_X(t+1)$ for α_H because B's T_X score is higher for α_H (note the colleges' preferences are completely determined by test scores), and for every college C in $\Gamma \setminus \{X\}$ the sets $P_C(t)$, $A_C(t)$, and $R_C(t)$ and the threshold $T_C(1)$ are the same for $\alpha_H(t)$ as for $\alpha_L(t)$. It follows that for every college C the set $P_C(t+1)$ is the same for $\alpha_H(t+1)$ as for $\alpha_L(t+1)$. And for every college C the ranking of the students in $\Sigma\setminus\{B\}$ according T_C is the same for α_H as for α_L. Therefore, for every college C the set $A_C(t+1)$ is the same for $\alpha_H(t+1)$ as for $\alpha_L(t+1)$. We must have $\alpha_H(t+1) = \alpha_L(t+1)$.

Further, suppose α_L assigns B to X, and X ranks first in B's preference ordering. Suppose that α_H is an application of SODA for which everything is identical to α_L except that B gets a higher test score on T_Y from some $Y \neq X$. Then α_H assigns B to X because B never applies to any school other than X and all the T_X scores (including that of B) are the same for α_H and α_L. □

Thus, in the case where student B is accepted by her most preferred college X for α_L and α_H in all rounds, she is neither rewarded nor punished for receiving a lower test score. The question emerges, however, as to what would happen if X were not B's most-preferred college. Could B's higher score displace some student D, leading to a chain reaction that steers B to getting assigned a less preferable college than the college to which she is assigned with a lower test score? The following proof deals with the case in which X is not B's most preferred college. It is best to approach the subsequent step in the proof with the following in mind:

(i) If B ranks college X in position τ or above in her preference ordering, and B is assigned to X when B receives the low score, then B will be assigned to X or better according to her own preferences when B receives a high score. This is a supposition.

(ii) B can never benefit by misrepresenting her preferences (Theorem 1).

Step 2: *If α_H assigns B to a college, that college ranks at least as in B's preference ordering as the college assigned to B by α_L.*

Proof:

Let X be the college to which B is assigned by α_L. Let π_B be the rank of X in p(B), which is B's preference ordering of colleges. (π_B=1 if X is most preferred by B, π_B=2 if X is preferred to every college except the one that ranks first, and so on.) Note that π_B specifically refers to college X's position and no other college's ranking.

We know that if π_B=1 then α_H assigns B to a college that ranks at least as high as X in p(B). Suppose that we have established that if $\pi_B \leq \tau$ then α_H assigns B to a college that ranks at least as high as X in p(B). We prove that if $\pi_B = \tau+1$ then α_H assigns B to a college that ranks at least as high as X in p(B). In other words, if α_L assigns B to college X, B cannot benefit by receiving a lower test score because with a higher score, B is assigned to a college she prefers at least as much as college X (assuming no change in any student's preferences and no other changes in test scores).

Denote p(B) as follows:

$$\begin{array}{c} \underline{p(B)} \\ C_1 \\ C_2 \\ \vdots \\ C_{\tau\text{-}1} \\ C_\tau \\ X \\ \vdots \end{array} \qquad \xrightarrow{\alpha_L} \quad BX$$

That is, C_1 is B's most-preferred college, C_2 is B's second-preferred college, down to C_τ, the college that is in the τth place for B, with X ranking next in the (τ+1)th position. The preferences of all students in $\Sigma \backslash \{B\}$ remain the same for α_L and α_H. Of course, $\alpha_L \rightarrow BX$ denotes the fact that α_L assigns B to X.

Assuming that everything else, including B's test scores, remain the same, we change B's preferences by moving X up one rank without changing the relative ordering of any other college. Let the new preference ordering be denoted p′(B). Note that X is now in the τth position in p′(B).

We have:

$$\begin{array}{c} \underline{p'(B)} \\ C_1 \\ C_2 \\ \vdots \\ C_{\tau\text{-}1} \\ X \\ C_\tau \\ \vdots \end{array}$$

Let α'_L denote the application of SODA with p(B) replaced by p′(B). (Two different combinations of test score and preferences arise: [p(B), α_L], [p′(B), α'_L].) Suppose α'_L assigns B to a college that ranks above X in p(B) (for instance, C_τ). Then B could manipulate α_L by reporting p′(B) instead of p(B). But this contradicts Theorem 1. For the previous supposition we view p(B) as the true preference of B; for following supposition consider $p'(B)$ as B's true preference ordering. Suppose that α'_L assigns B to a college (say Y) that ranks below X in $p(B)$. (See possible preference ordering below.) Then Y must rank below X in $p'(B)$ which means that B can manipulate α'_L by reporting $p(B)$, also a contradiction of Theorem 1. (This is still under the assumption that $p(B)$ assigns B to college X). Therefore α'_L assigns B to X.

<u>p(B)</u>	<u>p′(B)</u>
C_1	C_1
C_2	C_2
$\vdots$	$\vdots$
$C_{\tau-1}$	$C_{\tau-1}$
C_τ	X
X	C_τ
Y	Y
$\vdots$	$\vdots$

Note that if B manipulates her preferences by placing a lower ranked school (according to her true preferences) above X it is possible that B will get into that less preferred college. It is only in the case that B manipulates her preferences by reporting a higher ranked school below X that the manipulation is inconsequential. In other words, viewing $p(B)$ as B's true preferences: if B reports X, the lower ranked school, above C_τ, B does not change the likelihood that she will get matched with C_τ; however, if B were to report that Y ranked above X, B may get into a less preferred college. This concept is embedded in Theorem 1.

Now consider α_H: Note that $\pi_B = \tau$ for $p'(B)$. Therefore, by the induction hypothesis α'_H must assign B to college X or a college that ranks higher than X in $p'(B)$. (Of course α'_H is the history of SODA with everything the same as α'_L except that B has a higher score in the test used by college X.)

Finally, suppose that α_H assigns B to a college that ranks below X in p(B). Then B can manipulate at α_H by reporting p′(B) because a college that ranks at least as high as X in p′(B) ranks at least as high as X in p(B). But this contradicts Theorem 1. (We have now explored four combinations of test scores and preferences: [p(B), α_L], [p′(B), α'_L], [p(B), α_H], and [p′(B), α'_H].) Thus, if B ranks college X ranks in position $\tau+1$ in her preference ordering, and B is assigned to X when B receives the low score, then B will be assigned to X or better when B receives a high score. □

Note that the proof of Step 2 is valid whether the difference between α_L and α_H lies in B's score on T_X or on some other college's test. (All of the detail concerning the effect on B's assignment of different scores is embedded in the proof of Theorem 1.)

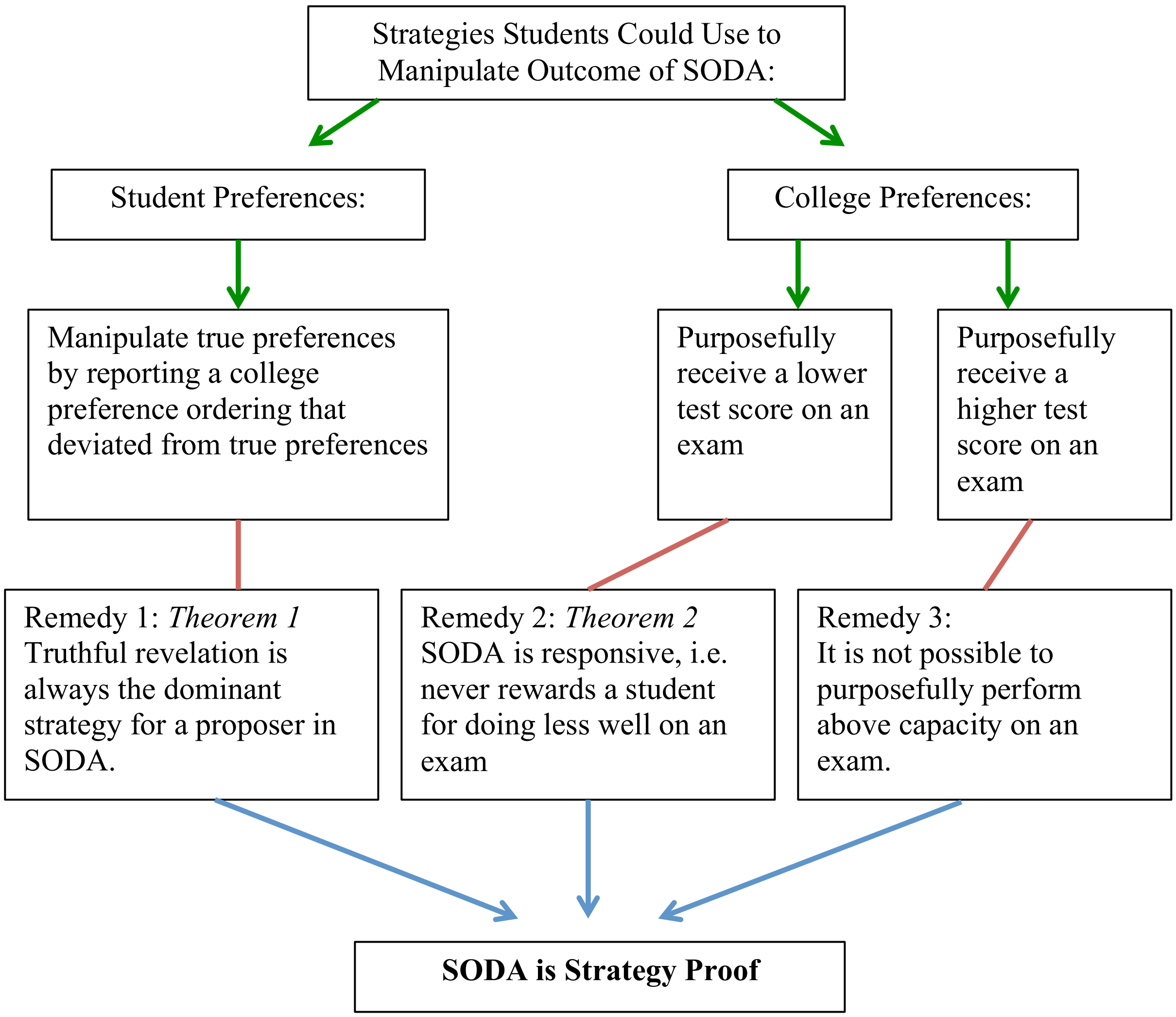
3.4 A Proposition: Is SODA Strategy Proof?

Within the structure of "optimal" deferred acceptance algorithms, the "proposer" cannot benefit my misrepresenting his or her preferences, as proven in Theorem 1. In any simple many-to-one, two-sided matching problem, the recipient can manipulate his or her preferences to obtain a more preferable matching, as explored in Chapter 1. However, colleges do not have that luxury in the college admissions problem as we have defined it, as a central clearing house determines student rankings based on test scores used by each college. Because colleges' preferences are completely determined by the test scores of the students, by definition schools cannot manipulate their preferences.

This leaves students as the only potential agents to manipulate preferences. Figure 3.4 demonstrates the potential strategies a student could employ to manipulate the outcome of a particular application of SODA. A student can either change the outcome of SODA by falsely reporting her preferences or by altering the preferences of the colleges (by manipulating their test scores). Theorem 1 proves that students have no incentive to manipulate their college preferences (See Remedy 1 in figure 3.4); SODA is fundamentally strategy proof for students in the sense that truthful revelation of preferences is the dominant strategy. However, students can also change the outcome by "manipulating" the preferences of the colleges in one of two ways: (i) purposefully receiving a lower test score or (ii) purposefully receiving a higher test score. We can reject both possibilities. No student has an incentive to perform below capacity on an exam because SODA is responsive (never rewards a student for performing worse on an exam), thus negating the strategy of receiving a lower test score in order to obtain a more favorable assignment (Remedy 2). And (Remedy 3) by definition, a student cannot perform beyond his or her capacity on an exam, thus eliminating the possibility of receiving a higher test score to change the outcome.

While this section does not comprise a formal proof, it provides an intuitive framework for the proposition that SODA is strategy proof when the test scores of the students completely determine the preferences of the college. Theorem 1 and Theorem 2 are necessary, but not sufficient, for proving that SODA is strategy proof in the case of any number of students, any number of colleges, and any number of tests.

Figure 3.4: "Strategy Proofness" of SODA



Chapter 4 – A Comparative Analysis of SODA's Properties

This chapter compares SODA to other matching mechanisms based on the following criteria: stability, multiple tests, fairness, responsiveness, and strategy proofness. As this chapter will explore, SODA is the only algorithm among the mechanisms we examine that is stable, uses multiple tests, is responsive, and (as we proposed in Chapter 3) is strategy proof. This chapter commences by defining each property and explaining its importance grounded in the possibility of a real-life application. In section 4.2, we evaluate relevant matching mechanisms according to the set criteria, providing simple examples of either fulfillment of or violations of the properties. This chapter does not consist of formal proofs but rather is intended to provide a more comprehensive, *comparative* look at SODA.

4.1 The Criteria

Stability:

If there exists a student and a school who mutually prefer each other to their current matches, then that student-college pair can upset the matching by blocking. If no pair wants to block, then that matching is stable. Formally, there is no match such that some student A would rather be paired with college X than college Y, with which she is currently matched, and college X would rather be paired with student A than student B, a student who is currently assigned to X. If there is no such blocking pair, regardless of the specification of student and college preferences, then we say the mechanism is stable.

Stability ensures the success of a given algorithm because no coalition, either one student-college pair or many pairs, has an incentive to block in order to achieve a more preferred outcome. While we have thus far explained stability within the mathematical constructs of matching mechanisms, the property of stability is crucial for the success of any real-world application. Without stability, an assignment procedure will unravel. In reality, where agents are free to make agreements outside of a mechanism, stability is necessary in order to ensure participation. As explained by the Royal Swedish Academy of Sciences, "from an economic point of view, stability formalizes an important aspect of idealized frictionless marketplaces," going on to say that "if individuals have unlimited time and ability to strike deals with each other, then the outcome must be stable, or else some coalition would have an incentive to form and make its members better off" (*Stable Allocations and the Practice of Market Design*, 2012). In the case of a voluntary college admissions process similar to that of NIMP, lack of stability would translate to lack of involvement, jeopardizing the legitimacy of the process itself. Thus we choose stability as the first desirable property of any algorithm assigning students to colleges.

Multiple Tests:

We require that the algorithm allow for the use of multiple tests. The use of multiple tests is a definitional property. According to the parameters of an algorithm, it will either use one test (as in the case of the Serial Choice Algorithm with a single exam) or multiple tests (as in the case of SODA). In practice, however, allowing for multiple tests in and of itself does not constitute satisfaction of this property; we must also take into account the way that the test(s) interact with the procedure.

Keeping in mind that the college admissions problem is an abstract representation of a real-life decision, this decision may include as many "tests" as there are institutions. "Tests" in the economic representation may translate to quantifiable aspects of students as a school strives to build a dynamic incoming class. The use of multiple tests is also vital for the assumption that colleges' preferences over students are strict. For instance, imagine two students receive exactly the same test score and have the same quantitative academic record (GPA)—this is where the importance of a qualitative "knife-edge" phenomenon of true preferences comes into play. College Preferences are based solely on test scores in this *theoretical* exploration; however, in reality, college admissions are based on much more. If colleges were to assign a points system to different admirable characteristics or faculties of a student – test scores (such as the SAT or ACT), extracurricular activities, personal interviews, recommendations, entrance essays, varying outside interests, or work experience, for instance – the "knife-edge" phenomenon discussed in Chapter 1 would then emerge. Distinctive characteristics of a student would play into each weighted points section, determining the quantitative score of the student on a generalizable scale. It is reasonable to think of a student's T_X score as the quantification of both qualitative and quantitative aspects on a weighted scale determined by college X.

In addition, even if we take the concept of a test as face value, multiple tests are important to appeal to a wider range of potential applicants. Because of educational background, expertise, natural inclination, demographic, interest in a certain area of academics, and/or a host of other reasons, a student may prefer one test to another. For instance, the SAT and ACT have key differences causing some students to gravitate towards one or the other. According to the Princeton Review, the SAT has a stronger emphasis on vocabulary than the ACT, the ACT tests science while the SAT does not, and the ACT tests more advanced math concepts (*The SAT vs. the ACT*, 2015).

Depending on ones abilities, a student may perform better on the ACT or the SAT, but this depends on having the opportunity to choose. Moreover, even if only one test were used, such as the SAT, there would in practice be multiple tests employed if colleges weighted the components (math, reading, and writing) differently.

Fairness:

In order for an algorithm to be fair, a student who receives a higher test score than another student on the test used by their mutually preferred school cannot be denied admission to that school if the student with the lower test score has been accepted. Formally, when test scores generate college preferences, an assignment is fair if we cannot find a student S and a college C such that S prefers C to the college assigned to S and S has a higher score on the test used by C than some student assigned to C. As mentioned in Chapter 2, when test scores are perfectly consistent with college preferences, stability and fairness are interchangeable.

People have a natural inclination towards fairness. From the first grade playground to compensation in the workplace, the intrinsic need to feel as if a process is fair follows us. This concept can be demonstrated by the ultimatum game in experimental economics. Each player is given a certain amount of money (or chips, coins, points, etc.) to split between him or herself and another participant. It is a one-time offer – a "take it or leave it" scenario. The recipient of the offer is more likely to reject the sum if the amount offered is deemed "unfair" (Kagel, Kim, & Moser, 1996). Although it would be in each player's best interest to take *any* sum of money offered, the concept of fairness overshadows this rational choice. This seemingly illogical decision demonstrates the psychological power of fairness in our behavior. While most economic theories weigh fairness against other desirable properties, in the case of college admissions, participation would dwindle in the absence of fairness.

The above example also hits on a more natural understanding of fairness: justifiability. If certain students or colleges have an unjustifiable advantage or disadvantage, the algorithm violates fairness on a visceral level. Acknowledging this, while we first and foremost evaluate the algorithm according to the formal definition, the evaluation touches on the colloquial definition as well.

Responsiveness:

Responsiveness, used interchangeably with positive responsiveness throughout the paper, formalizes the notion that an algorithm should reward students for performing to the best of their capability. As introduced earlier, an mechanism that assigns students to colleges is responsive if it never assigns a student to a college that she prefers to the one that she would be assigned if she had earned a higher score on any single exam, assuming no change in any student's preferences and no change in any other test score.

The connection between incentives and behavior is at the core of economics; it is also central to responsiveness. The right incentive structure is necessary to elicit the desired behavior – in this case maximal effort on *all* exams. If a particular algorithm does not always reward a student for performing her best on an exam, her motivation to do so will presumably diminish. The belief that one's actions affect his or her future is positively correlated with motivation; the converse is encompassed by the psychological phenomenon "learned helplessness", where subjects detach actions from outcomes, thus diminishing the motivation to escape aversive situations and undermining future motivation for related stimuli (Maier & Seligman, 1976). This phenomenon demonstrates the importance of the connection between actions and outcomes, and it relates to positive responsiveness in the sense that a mechanism must ensure students feel as if their effort will yield a positive outcome in order to incentivize high performance on exams.

Responsiveness also encompasses an aspect of fairness; if a student can benefit by receiving a lower test score on a test, it is unfair to the students who receive higher test scores for that student to be assigned a more preferred college. Because it is arguably unlikely for a student to try to manipulate the outcome by receiving a lower test score because of lack of information about other students' test scores (asymmetric information), habitually performing to the best of one's ability, pride, and so on, we focus on this aspect as a more serious concern.

Responsiveness also addresses allocative efficiency. The purpose of any matching algorithm is to generate good matches – that is, matches that are beneficial to both parties to some extent. For college admissions, students who perform better should be matched with better schools (schools that the high preforming students presumably prefer).

Strategy Proofness:

In the context of the college admissions problem, an algorithm is strategy proof if neither the school nor the student has an incentive to deviate from truthful revelation of preferences; as examined in Chapter 3, if we define colleges' preferences as equivalent to the ranking of student test scores, colleges do not have the capability to misrepresent preferences using SODA and CODA. However, for two-sided, many-to-one games in which both sides are able to manipulate preferences, there exists no stable mechanism that guarantees truthful revelation as a dominant strategy for both parties (Roth 1985). This research deviates from the bulk of the existing literature on the college admissions problem by using test scores to eliminate the possibility of colleges manipulating preferences (in the context of SODA and CODA).

Transparency and strategy proofness are deeply intertwined. If truthful revelation is the dominant strategy for all agents, it ensures that the system runs smoothly. As discussed in Chapter 1, truthful revelation acts to level the playing field. If agents, students or parents in the case of college admissions, are able to benefit by deviating from truthful revelation, it creates an unfair advantage for those who figure out how to successfully maneuver the system. Looking to the Boston mechanism, “naively truth-telling students (or parents)” suffered the most because the algorithm is not strategy proof (Ehlers & Klaus, 2012). As with any education proposition, equal opportunity is central. Strategy proofness guarantees that no student is at a systematic disadvantage because he or she lacks the social or financial capital necessary to manipulate the game. This is key.

4.2 Other Algorithms

Student Serial Choice (SSC with multiple tests):

The rules of the game: There are tests T_1, T_2, …., T_K where K is the total number of entrance exams. The tests are ordered randomly, with the number correlating to the arbitrary ordering. The student with the highest score on T_1 goes first and chooses her most-preferred college. The student with the highest score on T_2 then goes and chooses her most-preferred college. The pattern continues until the student with the highest test score on T_K chooses her most preferred college. If a college reaches its capacity of q_C students at any point, it is taken off the menu of choices. The process begins again as the student with the second highest test score on T_1 chooses her most preferred college among the remaining schools. The student with the next highest score on T_2 then chooses her most preferred college, and so on until the student with the second highest score on T_K chooses her most preferred school. This pattern continues as the algorithm rotates through the students with the next highest scores on the randomly ordered tests. When a student chooses a school, the assignment is final. The algorithm terminates when all students have been assigned to a college or all colleges have reached their capacity.

Stability: Imagine there are colleges C_1, C_2, C_3, and C_4 where each has a capacity of two students. Suppose that C_2 uses T_4 (similar to how one school focused on the quantitative score rather than the verbal score in Chapter 2's example), and student B receives the highest test score on T_4. Student B is fourth to choose her most preferred college according to the random ordering of tests, and her most preferred college is C_2. Suppose that the students with the highest test scores on tests T_1 and T_2 both choose C_2. C_2 has then reached its capacity of two students and is taken off the menu. Student B must then choose the school ranked just below C_2 on her preference list. However, C_2 and B mutually prefer each other to their current matches. There exists a student-college pair, C_2 and B, which could benefit by blocking, thus violating stability.

Multiple Tests: This algorithm consists of multiple tests; however, the arbitrary ranking of tests undermines the importance of this property. Additionally, the facts that colleges do not have a say in final assignments and that each college's preferred test is not taken into account diminishes the use of multiple tests.

Fairness: The formal definition of fairness that "we cannot find a student S and a college C such that S prefers C to the college assigned to S and S has a higher score on the test used by C than some student assigned to C" focuses on the ranking of student's test scores on the specific test used by that college. In this sense, the ranking is fair. Because the algorithm rotates through the highest scores, then the second highest scores, and so on, it is not possible for a student with a lower score on the same test to be assigned to a preferred college first. However, this algorithm violates fairness in the colloquial sense: the arbitrary ordering of tests is unjustifiable. The students who take test T_1 have a systematic advantage over all other students who took any other test. Students who take T_2 similarly have an advantage all students who took any test other than T_1, and so on. This structural disadvantage is not based on merit, and thus violates a fundamental definition of fairness.

Responsive: SSC with multiple tests is responsive. A student can never benefit by doing worse on an exam because doing so bumps her down to the next rotation of school assignments; she would never receive admission to a more preferred college by receiving a lower test score on an exam. While we do not provide a formal proof here, the intuition is sufficient. If each student's "turn" to choose schools is according to the highest test score in the first rotation, the second highest in the next, and so on, a student cannot benefit by getting a lower score on an exam because his or her turn would come later, and he or she may have a smaller list of schools from which to choose. Satisfying the formal definition of fairness, however, does not account for the fact that the tests are arbitrarily ranked (covered in fairness).

Strategy Proof: No student can benefit by deviating from truthful revelation because an assignment is final once a student chooses her (most preferred) school. If a student instead chose a school that was not her most preferred, she would be punishing herself by deviating from truthful revelation. Schools cannot misrepresent their preferences, as their preferences are not taken into account. While the mechanism itself is strategy proof, we must remember that there exists a coalition that could benefit by cooperating outside of this procedure. Because the algorithm is not stable, unless participation is mandatory, students and colleges can benefit deciding to make arrangements outside of the mechanism. A strategy for some would then be to ditch the algorithm all together. The violation of stability discussed above trumps the benefit of within-mechanism strategy proofness.

College Serial Choice (CCS with multiple tests):

The rules of the game: C_1, C_2, … , C_m are the m colleges, which have been ordered randomly with the subscript corresponding to the institution's position in the order. The college C_1 goes first, choosing the student with the highest score on the test used by C_1. College C_2 goes next and chooses the student with the highest score on the test used by C_2 among the remaining students, and so on until each college has chosen a student. The process then begins again as C_1 chooses the student with the highest remaining test score on the test used by C_1. C_2 then chooses the student with the highest remaining test score on its preferred test, and so on. This pattern continues as the algorithm rotates through the randomly ordered colleges each choosing their most preferred student among the remaining students. The assignments are final, and when a college reaches its capacity, it ceases to participate.

Stability: Imagine there are colleges C_1, C_2, and C_3 each with room for one student. Colleges C_1 and C_3 admit students based on their test scores on test X (T_X). C_2 admits students based on test Y (T_Y). Suppose a student B receives the highest test score on T_X, and B's most preferred college is C_3. Because C_1 chooses its most preferred student first, the mechanism assigns B to C_1. C_2 chooses the student with the highest test score on T_Y. C_3 would then choose the student with the second highest test score on T_X. However, this leaves a student-college pair that could benefit by blocking. There exists a student B who would rather be paired with C_3 than her current match C_1, and a college C_3 who prefers B to the student with the second highest test score on T_X. Thus, CSC is not stable.

Multiple Tests: Yes, this allows for multiple tests.

Fairness: Fairness mirrors stability. According to the formal definition, we can find a student and a college such that the student prefers the college to the one assigned to her by the mechanism, and that student has a higher test score on the test used by the college than some other student who is assigned to that college. The example used above for the violation of stability can be used to demonstrate a violation of fairness. In addition to violating the strict definition of fairness, this algorithm violates a fundamental understanding fairness: the random ordering of colleges is not justifiable. Once again, the ordering of schools is not earned; devoid of a merit-based system to determine the ordering of schools, CSC is not fair in both formally and intuitively.

Responsiveness: CSC is not responsive. Given the situation where there are three colleges, C_1, C_2, and C_3 each with room for one student, and colleges C_1 and C_3 admit based on T_X while C_2 admits based on T_Y. Suppose student B prefers C_3 to C_1. It is possible for student B to get accepted into a more preferred school by receiving a lower test score on T_X. Let α_H be the application of CSC in which B receives a higher test score T_X. Let α_L be the application of CSC where the only change is that student B receives a lower test score on T_X. Suppose that in α_H student B receives the highest test score on T_X. C_1 would go first and choose B. Suppose that in α_L B's test score changes such that there is one student A with a test score that surpasses that of student B's lower T_X score. B would then have the second highest T_X score. C_1 chooses the student with the highest test score on T_X, student A. C_2 chooses the student with the highest test score on T_Y. Then C_3 chooses the student among those remaining with the highest test score on T_X, or student B. Student B, who prefers C_3 to C_1 receives a more preferable assignment in α_L than α_H, thus violating responsiveness.

Strategy Proof: From the perspective of the schools, they cannot benefit by misrepresenting their preferences because the assignments are final, and they are always choosing their most preferred student among those remaining. If schools were to deviate from truthful revelation, they would be choosing students whom they prefer less than another possible assignment. Thus, the dominant strategy for schools is truthful revelation. The students, however, could manipulate college preferences by purposefully receiving a lower test score in order to get a preferred matching. While this is highly unlikely considering asymmetry of information and the personal pride that results from performing one's best, it is possible. Thus, CSC is not strategy proof.

Arbitrary Assignment (AA with multiple tests):

The rules of the game: Students and colleges are randomly matched. No college can be matched with more than its capacity of students. Student preferences and test scores are not taken into account.

Stability: Suppose there are three schools C_1, C_2, and C_3 each with room for one student. Suppose among the students there exists a student B whose most preferred college is C_2, and student B is C_2's most preferred student. Because the assignments are random, it is possible that B is assigned to C_1. C_2 and B would then benefit by blocking, thus violating stability. Although it is possible that the random matching is stable, the algorithm does not guarantee stability.

Multiple Tests: Yes, there are multiple tests; however, the fact that the algorithm does not take into account test scores entirely diminishes the importance of test scores, let alone the opportunity to take multiple tests.

Fairness: Arbitrary Assignment is neither fair according to the formal definition nor fair according to a fundamental understanding. Similar to stability, imagine three schools C_1, C_2, and C_3 each with room for one student. Suppose C_2 prefers student B to student A because B receives a higher test score on the test used by C_2. Suppose that student B similarly prefers C_2 to any other college. It is possible that by random assignment, student A is assigned to C_2, and student B is assigned to C_1. Student B received a higher test score than student A and prefers C_2 to her current match; thus the matching is unfair. In addition, Arbitrary Assignment is not based on merit. Just as the random ordering of colleges and test scores is not justifiable for SSC and CSC, respectively, the random assignment of students to colleges is not justifiable.

Responsive: While a student cannot benefit by receiving a lower test score, one cannot benefit by receiving a higher test score either. The algorithm neither rewards nor punishes students for receiving a lower test score because test scores hold no weight in the outcome. It is possible for a student to be assigned to a preferred school with a lower test score because the assignment is random. Thus, AA is not responsive.

Strategy Proofness: Yes, this algorithm is strategy proof. An agent cannot benefit by deviating from truthful revelation, nor can an agent be punished for deviating from truthful preferences. Because preferences and test scores are not taken into account, the algorithm requires no strategy and, thus, is necessarily strategy proof.

Serial Choice Algorithm (SCA with a single test):

The rules of the game: A single test is used to determine student rankings. College preferences are entirely determined by the students' test scores on this single test. The student with the highest test score declares her most preferred college and is matched with that college. The assignment is final. The student with the second highest test score then chooses her most preferred college and is matched with that college. Going in order from highest to lowest test score, each student chooses his or her most preferred college of those remaining. Each choice is final if the school accepts the proposal. Once a college has reached its capacity of q_C students, students can no longer be matched with that school. In other words, when a school is full, it remains full, and it is no longer an available option. The algorithm terminates when every school reaches its capacity or every student has been matched with a school.

Stability: SCA always yields a stable matching. Students propose to colleges in order of their test scores, highest to lowest. Because the test scores determine college preferences, each time a student chooses a college, it is the college's most preferred student among those remaining, and by definition, it is the student's most preferred school among those remaining. If B prefers college C but is assigned to college D then C must have been unavailable when it was B's turn to choose. This means that all the students assigned to C have higher test scores than B, and thus B and C cannot block. In fact, the SCA with a single test yields the *only* stable outcome. Thus SCA=SODA=CODA with a single test.

Let S_1, S_2, … , S_n denote the students, with S_1 choosing before S_2, who chooses before S_3, and so on. Let C_i be the college assigned to S_i by the SCA. (Of course we can have $C_i = C_j$ even if $i \neq j$.) Every stable outcome must have S_1 matched with C_1 because C_1 is S_1's most-preferred school, and S_1 is C_1's most-preferred student. Suppose that every stable assignment must have S_i assigned to C_i for all $i \leq t$. Then every stable assignment must have S_{t+1} assigned to C_{t+1}. All of the colleges preferred to C_{t+1} by S_{t+1} must be full of students preferred to C_{t+1} by those colleges. In fact, those colleges must be full of students that are assigned to those colleges at every stable outcome. Therefore, every stable assignment must assign S_{t+1} to C_{t+1}. Otherwise S_{t+1} would be assigned to a college D that ranks below C_{t+1} in the preference ordering of S_{t+1}, and C_{t+1} would prefer S_{t+1} to any student S_j that took the place of C_{t+1} because we would have to have $j > t+1$ if the outcome were stable. (And C_{t+1} would prefer to enroll S_{t+1} to having any empty seat.)

Multiple Tests: The major, and only, downfall of this algorithm according to these properties is that it uses a single test.

Fairness: Stability and fairness are interchangeable when the test scores of the students determine the preferences of the college (as was fully explained in Chapter 2). Mirroring the explanation of stability above, it is impossible for a student with a lower test score to be assigned to her preferred college over a student with a higher test score who prefers the same college because the student with the higher test score chooses first. (And she will presumably choose the remaining college that she most prefers.)

Responsiveness: This algorithm is the most blatantly responsive. Students are strictly rewarded for getting higher test scores because they have priority over students who receive lower test scores. Receiving a lower test score "would only result in a lower ranking, and a later choice, perhaps from a smaller list of schools" (Campbell, 2006). In other words, students have the incentive to get the highest possible test score in order to rank higher, choose earlier, and have the largest list of schools from which to choose.

Strategy Poof: Yes, SCA is strategy proof. As covered in each of the above explanations, the student has the incentive to get the highest test score possible, and it is the dominant strategy for the student to truthfully reveal her most preferred school at the time of her choice. These two components constitute student-side strategy proofness. Can a college benefit by misrepresenting its preferences? Or, in the context of this algorithm, can a school benefit by denying acceptance to a student who proposes? No. By denying a student's proposal, the school would be saving a seat for a later priority student with a lower test score. Thus no student or college can benefit from misrepresenting preferences, and no student can benefit by lowering her test score. In addition, the application of

SODA and CODA generate the same outcome; in other words, there is only one stable matching, and it is strategy proof.

4.3 – Summarizing the Algorithms

Pulling from the information in Chapters 2, 3, and 4, the below chart summarizes the algorithms according the set criteria. As previous chapters have shown, SODA is the only algorithm that is stable, uses multiple tests, is fair, is positively responsive, and may be strategy proof.

	Stable	**Multiple Tests**	**Fair**	**Responsive**	**Strategy Proof**
SODA	✖	✖	✖	✖	✖
CODA	✖	✖	✖		
SSC		✖		✖	✖
CSC		✖			
AA		✖			✖
SCA	✖		✖	✖	✖

Chapter 5 – Conclusion

Since the deferred acceptance algorithm's formal introduction by Gale and Shapley in 1962, its application to a theoretical marriage market has transcended into the reality of kidney transplants, school choice, the National Residency Matching Program, and college admissions –the focus of this research. This paper provided evidence that the Student Optimal Deferred Acceptable algorithm is characterized by a set of desirable properties including stability, fairness, and positive responsiveness for any number of students, colleges, and tests. Using the framework of the marriage problem discussed in Chapter 1, Chapter 2 explored a simple example of the case of two students, two colleges, and two tests; Chapter 3 built upon this, proving that SODA is responsive in the general case. Providing a comparative context, Chapter 4 examined alternative algorithms according to the criteria of stability, multiple tests, fairness, responsiveness, and strategy proofness. Among the mechanisms we explore, SODA is the only one using multiple tests that always yields an outcome that is stable, fair, responsive, and strategy proof. However, we have fallen into Gale and Shapley's (1962) trap as:

> *The reader who has followed us this far has doubtless noticed a trend in our discussion. In making special assumptions needed in order to analyze our problem mathematically, we necessarily moved further away from the original college admission question, and eventually… abandoned reality altogether and entered the world of mathematical make believe.*

The most conspicuous assumption is that using test scores to determine college preferences is justified based on merit, and merit alone. The responsiveness property of the SODA algorithm addresses allocative efficiency, fairness, and incentives within the bounds of the mechanism; however, we must consider the intersection between allocative efficiency and the reproduction of social class in terms of test scores. As long as test scores are factored into the college admissions process, regardless of their weight among other characteristics, this issue needs to be addressed.

Socioeconomic status and standardized test scores are frighteningly positively correlated. When SAT component scores are broken down by family income, scores increase significantly with each $20,000 increase income. The mean Critical Reading, Mathematics, and Writing scores for students in the $0-$20,000 income range are 435, 462, and 429, respectively; for students in the "more than $200,000" family income bracket, the Critical Reading, Mathematics, and Writing scores are 565, 586, and 563, respectively (College Board, 2013). This is hardly surprising considering the intersecting inequalities that feed into disparities in academic achievement. So what does this mean in terms of SODA? Rich kids get assigned to more preferable schools. Allocatively efficient algorithms seek to create good matches with "better" students attending "better" schools; however, if preference orderings are strictly based on standardized test scores, the mechanism blatantly violates equal access to education, thus fueling the opportunity gap.

The "allocatively efficient" matching mechanism, devoid of historical, demographic, and socio-economic considerations then becomes an exclusionary procedure for the reproduction of social class via education. Mathematically sound does not constitute morally sound. In future research, we need to add measures to the SODA algorithm to address the structural disadvantages of certain students; this research acts starting point.

Works Cited

Roth, A. E., (1982). The Economics of Matching: Stability and Incentives. *Mathematics of Operations Research*, 7(4), 617-628. doi: 10.1287/moor.7.4.617

Abdulkadiroglu, A. (2013). School Choice. *Oxford Handbook of Market Design*, (Z. Neeman, M. Niederle, A. E. Roth, and N. Vulkan, eds). Oxford University Press, forthcoming.

Abdulkadiroglu, A., Pathak, P.A., & Roth, A. E., (2005). The New York City High School Match. American Economic Review, Papers and Proceedings, 95(2), 364-367.

Abdulkadiroglu, A., Pathak, P.A., & Roth, A. E. (2009). Strategy-proofness versus Efficiency in Matching with Indifferences: Redsigning the NYC High School Match. *American Economic Review*, 99(5), 1954-1978.

Abdulkadiroglu, A., Pathak, P.A., Roth, A. E., & Somnez, T. (2005) The Boston Public School Match. *American Economic Review, Papers and Proceedings,* 95(2), 368-371.

American Mathematical Society (2015). *Feature Column: School Choice.* Retrieved August 2014 from http://www.ams.org/samplings/feature-column/fcarc-school-choice

Campbell, D.E., (2006). Matching. *Incentives: Motivation and the Economics of Information* (2nd ed.), (pp. 467-512). New York, NY: Cambridge University Press.

Dubins, L.E., & Freedman, D.A. (1981). Machiavelli and the Gale-Shapley algorithm. *American Mathematical Monthly*, 88, 485-94.

Ehlers, L. & Klaus, B. (2012). Strategy-Proofness makes the Difference: Deferred-Acceptance with Responsive Priorities. *University of Montreal.* Retrieved from http://www.cireqmontreal.com/wp-content/uploads/cahiers/15-2012-cah.pdf

Gale, D. & Shapley, L. S. (1962). College Admission and the Stability of Marriage. *The American Mathematical Monthly,* 69(1), 9-15. Retrieved from file:///Users/cshira/Honors%20Fellowship/Lit%20Review/College%20Admission%20and%20the%20Stability%20of%20Marriage.pdf

Great Schools Partnership: The Glossary of Education Reform. (2014). *Opportuntiy Gap.* Retrieved April 16, 2015, from http://edglossary.org/opportunity-gap/

Kagel, J.H., Kim, C. & Moser, D. (1996). *Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs,* 13(1), 100-110.

doi:10.1006/game.1996.0026

Maier, S.F. & Seligman, M.E. (1976). Learned Helplessness: Theory and Evidence. *Journal of Experimental Psychology*, 105(1), 3-46. Retrieved from http://dx.doi.org/10.1037/0096-3445.105.1.3

Mongell, S. & Roth, A. E. (1991). Sorority Rush as a Two-Sided Matching Mechanism. *American Economic Review,* 81, 441-464.

Roth, A.E. & Sotomayor, A.O. (1999). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* New York, NY: Cambridge University Press.

Roth, A.E. (1985). Common and conflicting interests in two-sided matching markets. *European Economic Review*, 27, 75-96.

Stable Allocations and The Practice of Market Design (2012). *The Royal Swedish Academy of Sciences*. Retrieved from http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2012/advanced-economicsciences2012.pdf

The College Board. (2013) *2013 College-Bound Seniors: Total Group Profile Report*. Retrieved April 15, 2015, from College Board SAT Data.

The Match: National Residency Matching Program. (2015) *Couples in the Match.* Retrieved April 16, 2015 from http://www.nrmp.org/match-process/couples-in-the-match/

The Princeton Review. (2015). *The SAT vs. the ACT.* Retrieved April 14, 2015 from http://www.princetonreview.com/college/sat-act

The Prize of Economic Sciences 2012 (2014). *Nobelprize.org.* Retrieved from http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2012/