JOURNAL OF ECONOMIC THEORY **36**, 277–288 (1985)

The College Admissions Problem Is Not Equivalent to the Marriage Problem*

ALVIN E. ROTH

Department of Economics, University of Pittsburgh,
Pittsburgh, Pennsylvania 15260

Received August 28, 1984; revised December 6, 1984

Two-sided matching markets of the kind known as the "college admissions problem" have been widely thought to be virtually equivalent to the simpler "marriage problem" for which some striking results concerning agents' preferences and incentives have been recently obtained. It is shown here that some of these results do not generalize to the college admissions problem, contrary to a number of assertions in the recent literature. No stable matching procedure exists that makes it a dominant strategy for colleges to reveal their true preferences, and some outcomes may be preferred by all colleges to the college-optimal stable outcome. *Journal of Economic Literature* Classification Numbers: 025, 026, 820. © 1985 Academic Press, Inc.

1. INTRODUCTION

This paper considers the preferences and incentives of agents in two-sided markets involving disjoint sets of agents (e.g., firms and workers). A particularly simple example of the kind of market to be considered here is the labor market for resident physicians studied in [9], in which the two sides of the market are hospitals and graduating medical students. In that market, each student seeks one job, and each hospital seeks a fixed number of students. This market is an example of the "college admissions problem" studied by Gale and Shapley [3], who focused attention on the special case called the "marriage problem," in which every agent seeks a match with exactly *one* agent on the other side of the market.

Here it will be shown that the incentives facing the agents in the college admissions problem are quite different from those found in the marriage problem (see, e.g., [6]). This comes as a considerable surprise, since the

*This paper is a revision of a working paper entitled "Incentives in the College Admissions Problem and Related Two-Sided Markets." This work has been supported by grants from the National Science Foundation and the Office of Naval Research, and by Fellowships from the John Simon Guggenheim Memorial Foundation and the Alfred P. Sloan Foundation.

0022-0531/85 $3.00

Copyright © 1985 by Academic Press, Inc.
All rights of reproduction in any form reserved.

college admissions problem has come to be regarded as virtually equivalent to the marriage problem, and as a result the recent literature on the subject contains a number of misleading statements on precisely this issue. The cause of this confusion turns out to be a subtle kind of incompleteness in the traditional formulation of the college admissions problem.

2. The Marriage Problem

The agents in the marriage problem consist of two disjoint sets $M=\{m_1,..., m_n\}$ and $W=\{w_1,..., w_p\}$ ("men" and "women"). Each man has a complete preference ordering over the set $W \cup \{u\}$, and each woman has a complete preference ordering over $M \cup \{u\}$, where u denotes the possibility of remaining unmarried. That is, each agent can compare the desirability of each of his or her potential assignments, which are the agents from the opposite set and the possibility of remaining unmarried. An agent's preferences are called *strict* if he or she is not indifferent between any two distinct potential assignments. It will be sufficient for the purpose of this paper to only consider the case in which all agents have strict preferences, and this will henceforth be assumed. Let $w_j P(m) w_k$ denote that man m prefers woman w_j to woman w_k, and let $w_j R(m) w_k$ denote that he either prefers w_j to w_k or else is indifferent. (Note that he can only be indifferent if $j=k$, since all preferences are strict.) Similar notation will be used for the preferences of the women, and $P=(P(m_1),..., P(m_n), P(w_1),..., P(w_p))$ will denote the vector of preference orderings of each agent.

An *outcome* of the marriage problem is defined by a function $x: M \cup W \rightarrow M \cup W \cup \{u\}$, such that, for any m in M and w in W, $x(m)=w$ if and only if $x(w)=m$. An outcome x matches a subset of the men with a subset of the women in monogamous marriages, and leaves the remainder of the men and women unmarried. The preferences of the agents over alternative outcomes x and y correspond precisely to their preferences over their potential assignments; i.e., each man prefers x to y if and only if he prefers $x(m)$ to $y(m)$, and similarly for each woman. An outcome x is called *individually rational* if no man or woman prefers u (being unmarried) to the assignment $x(m)$ or $x(w)$, respectively; i.e., if $x(m)\, R(m) u$ and $x(w)\, R(w) u$ for all m in M and w in W.

An outcome x is *unstable* if it is not individually rational or if there exist a man m and a woman w who prefer each other to their assignment at x; i.e., a man m and a woman w for whom $wP(m)\, x(m)$ and $mP(w)\, x(w)$. An outcome x that is not unstable is *stable*. The set of stable outcomes constitutes the core of the game whose rules are that any man and woman may marry if and only if they both agree, and each agent may choose to remain unmarried. The set of stable outcomes with respect to a vector P of preferences is therefore denoted $C(P)$. For any man m the set of *achievable* assignments will be the set $A_m(P) = \{x(m) \mid x \text{ is in } C(P)\}$ of assignments achieved at some stable outcome, and the set of achievable assignments for each woman is defined analogously. (An achievable assignment for an agent may be either an agent from the opposite set, or else the assignment u.)

We can now state the following results about the structure of the set of stable outcomes. (Theorems 1 and 2 are originally found in [3] and Theorem 3 in [6].)

THEOREM 1. *For any vector P of preference orderings, the set $C(P)$ of stable outcomes of the marriage problem is nonempty.*

THEOREM 2. *The set $C(P)$ of stable outcomes of the marriage problem contains an M-*optimal stable outcome x^* *with the property that, for every man m in M, $x^*(m)$ is man m's* most preferred achievable assignment; *i.e., $x^*(m)\ R(m)\ x(m)$ for any other stable outcome x. Similarly, it contains a W-*optimal stable outcome y^* *such that $y^*(w)\ R(w)\ x(w)$ for every woman w and any stable outcome x.*

THEOREM 3. *There does not exist any outcome y that every man prefers to the M-optimal stable outcome x^* in the marriage problem; i.e., for no outcome y is it the case that $y(m)\ P(m)\ x^*(m)$ for all m in M. Similarly, there exists no outcome z preferred by all the women to y^*.*

Theorem 2 establishes the existence of a stable outcome x^* with the surprising property that all the men are in agreement that it is the best stable outcome. By symmetry, there also exists a W-optimal stable outcome y^* with corresponding properties, and it turns out that the optimal stable outcome for one side of the market is the worst stable outcome for every agent on the other side of the market. Theorem 3 states that there does not even exist an unstable outcome that all the men prefer to x^*, although it turns out (cf. Sec. 6 of [6]) that there can exist unstable outcomes that all men like at least as well as x^* and some men prefer. An algorithm to construct the outcome x^* was presented in [3].

Since an agent's preferences are typically known only to himself, we can consider what incentives an agent might have to reveal his true preferences. For the sake of clarity, consider a situation in which some centralized procedure is employed to produce an outcome from any vector P of strict preferences that might be stated by the agents. Such a procedure will be called a *stable matching procedure* if the outcome $x(P)$ it produces is always stable with respect to the stated preferences; i.e., if $x(P)$ is contained in $C(P)$ for any stated preferences P.

Let the *true* preferences of the agents be given by the vector P^* of preference orderings. The adoption of any particular matching procedure gives rise to a strategic game, in which each agent's strategy set is the set of all possible strict preference orderings he might state. We can now ask if the adoption of some stable matching procedure will always give all or some of the agents the incentive to state their true preferences; i.e., if it will make it a dominant strategy for an agent m or w to always state the true preference $P^*(m)$ or $P^*(w)$ rather than some other preference $P(m)$ or $P(w)$. The following two results are from Roth [6].

THEOREM 4. *There exists no stable matching procedure for the marriage problem which makes it a dominant strategy for all agents to state their true preferences.*

THEOREM 5. *The matching procedure that always yields the M-optimal stable outcome $x^*(P)$ for any stated preferences P makes it a dominant strategy for every m in M to state his true preferences in the marriage problem. Similarly, a procedure that always yields $y^*(P)$ makes it a dominant strategy for every w in W to state her true preferences.*

Dubins and Freedman [1] present an extension of Theorem 5 that states that in fact no coalition of men can all misstate their preferences in such a way that every member of the coalition receives a match he (strictly) prefers to his assignment at $x^*(P^*)$.

We can now consider the extent to which these results generalize to the college admissions problems.

3. THE COLLEGE ADMISSIONS PROBLEM

The agents in the college admissions problem consist of two disjoint sets $C=\{c_1,..., c_n\}$ and $S=\{s_1,..., s_p\}$ ("colleges" and "students"). Each college c_i has a *quota* q_i which is the number of students for which it has places. Each student s has a strict preference ordering P(s) over the set $C \cup \{u\}$, and each college c has a strict preference ordering $P(c)$ over the set $S \cup \{u\}$. An *outcome* of the college iadmissions problem is defined by a correspondence $\mathrm{x}: C \cup S \rightarrow C \cup S \cup \{u\}$ such that $|x(s)|=1$ for all s in S, $|x(c_i)|=q_i$ for all c_i in C, and, for any c in C and s in S, $x(s)=c$ if and only if s is an element of $x(c)$. That is, an outcome assigns a subset of the students to a subset of the places and leaves the rest of the students and places unmatched. (If a college with quota q is assigned some number $k<q$ of students at an outcome x, then $q-k$ elements of $x(c)$ are equal to u.) No student is assigned to more than one place, and no college is assigned more than its quota of students.

An outcome x is *individually rational* if for every student s $x(s)\,R(s)u$, and if for every college c and σ in $x(c)$, $\sigma R(c)u$, where R denotes preference or indifference as in Section 2. An outcome x is *unstable* if it is not individually rational or if there exist a college c_i and a student s_j who each prefer one another to one of their assignments; i.e., such that $c_i P(s_j)\,x(s_j)$ and $s_j P(c_i)\sigma$ for some σ in $x(c_i)$. An outcome x that is not unstable will be called *stable*, and the set of stable outcomes with respect to any vector P of preference orderings will again be denoted $C(P)$. A student s and college c will be said to be *achievable* for one another if there is some outcome x in $C(P)$ such that $x(c)$ contains s. This completes the traditional specification of the college admissions problem, and it is easy to see that the marriage problem is the special case of the college admissions problem that arises when each college has a quota of 1 (i.e., when $q_i = 1$ for every c_i in C).

Gale and Shapley [3] observed that the algorithm discussed for the marriage problem could be modified for the college admissions problem. Theorems 1 and 2 can then be extended to the college admissions problem, as follows.[1]

THEOREM 1*. *For any vector P of preference orderings. the set $C(P)$ of stable outcomes of the college admissions problem is nonempty.*

THEOREM 2*. *The set $C(P)$ of stable outcomes of the college admissions problem contains a C-*optimal stable outcome x^* *with the property that, for every c in C with quota q, $x(c)$ contains college c's q most preferred achievable students if the number k of students achievable for c is at least q, and otherwise it contains all of c's achievable students (and leaves $q-k$ positions unmatched). The set $C(P)$ also contains an S-*optimal stable outcome y^* *with the property that $y^*(s)\,R(s)\,x(s)$ for any student s and stable outcome x.*

When we try to extend Theorem 3, however, we see that the formulation of the college admissions problem given above is not complete enough to allow the theorem to be meaningfully stated. This is because we cannot state whether a college prefers one outcome to another until we have specified its preferences over *outcomes*, and so far we have only specified its preferences over *students*. Even if we continue to consider only the case in which a college's preferences over outcomes are determined entirely by its

[1] Note that Theorem 2* reduces to Theorem 2 when all quotas equal 1, but that it has some additional power in the general case. In the set of stable outcomes, a college need make no tradeoffs between achievable students. The theorem implies, for example, that if a college c with a quota of $q=2$ positions is matched with its first and third choice students at one stable outcome, and with its second and fourth choice students at another stable outcome, then at the outcome x^* it is matched with its first and second choice students.

own assignment of students, each college whose quota is greater than 1 will have to make comparisons of *sets* of students, and the preferences of colleges over such sets have yet to be specified.

It has nevertheless often been asserted in the literature that even when quotas may be greater than 1, the college admissions problem may be treated as a straightforward extension of the marriage problem. In the introduction to a paper whose formal treatment dealt with the marriage problem, I made an incautious comment of this sort (see [6]), as did Dubins and Freedman [1] in a concluding section of their paper that was intended to extend to the college admissions problem their results on the marriage problem. Gale and Sotomayor [4] sketch a formal proof of the equivalence of the two problems that will be useful to refer to later. They say the following about the college admissions problem:

> As in previous treatments of the problem, we begin by reducing it to the special case in which each institution has a quota of one. This is done by the following device: we replace institution A by q_A copies of A denoted by $A_1, A_2,..., A_{q_A}$. Each of these A_i has preferences identical with those of A but with a quota of 1. Further, each applicant who has A on his preference list now replace A by the string $A_1, A_2,..., A_{q_A}$ in that order of preference. It is now easy to verify that the stable matchings for the original problem are in natural one-to-one correspondence with the stable matchings of this modified model. With this modification, the model becomes completely symmetric in the applicants and institutions. To reflect this, we make the usual change of scenario to that of the "stable marriage problem" in which instead of applicants and institutions, we consider men and women and think of the matchings as (monogamous) marriages.

Gale and Sotomayer go on to review a number of results for the marriage problem, including Theorems 1, 2, 3, and 5 of the previous section. However we will see below that, although the conclusions of Theorems 1, 2, and also 4 apply to the college admissions problem, the conclusions of Theorems 3 and 5 do not.

The problem cannot be fixed simply by completing the specification of the model by specifying the preferences of colleges for sets of students. When the model is completely specified in this way, the conclusions of Theorems 3 and 5 will be false, so long as the preferences of colleges for sets of students are related to their preferences for individual students in a plausible way. Specifically, let $P^{\#}(c)$ denote the preference relation of college c over all assignments $x(c)$ that it could receive at some outcome x of the college admissions problemm. A college c's preferences $P^{\#}(c)$ will be called *responsive* to its preferences $P(c)$ over individual assignments if $y(c)\,P^{\#}(c)\,x(c)$ whenever $y(c)$ is obtained from $x(c)$ by replacing some student s_j (or u) in $x(c)$ with a preferred student s_k who is not in $x(c)$; i.e., whenever $y(c)=x(c)\cup\{s_k\}\backslash\{\sigma\}$ for σ in $x(c)$ and s_k not in $x(c)$ such that $s_k P(c)\sigma$. That is, a college c has responsive preferences over assignments if, for any two assignments that differ in only one student, it prefers the assignment containing the more preferred student. For example, if $x(c)$ assigns college c its 3rd and 4th choice students, and $y(c)$ assigns it its 2nd and 4th choice students, the college c prefers $y(c)$ to $x(c)$ if its preferences are responsive.

PROPOSITION 1. *When colleges have responsive preferences, the conclusion of Theorem* 3 *is false for the college admissions problem: there may exist outcomes that all colleges strictly prefer to the C-optimal stable outcome.*

The proof will be by means of an example.

Proof. Consider the problem consisting of three colleges $C=\{c_1, c_2, c_3\}$ and four students $S=\{s_1, s_2, s_3, s_4\}$. College c_1 has a quota of $q_1=2$, and both other colleges have a quota of 1. Each of the colleges prefers each of the students to leaving a position unmatched, and colleges c_1 and c_2 both prefer lower-numbered students to higher-number students, so their true preferences P^* are given by $s_1 P^*(c_1)\, s_2 P^*(c_1)\, s_3 P^*(c_1)\, s_4 P^*(c_1) u$, and $s_1 P^*(c_2)\, s_2 P^*(c_2)\, s_3 P^*(c_2)$ $s_4 P^*(c_2) u$. The preference ordering of the third college is given by $s_3 P^*(c_3)\, s_1 P^*(c_3)\, s_2 P^*(c_3)\, s_4 P^*(c_3) u$. The preference orderings of students s_1 through s_4 are given by $c_3 P^*(s_1)\, c_1 P^*(s_1)\, c_2 P^*(s_1) u$; $c_2 P^*(s_2)\, c_1 P^*(s_2)\, c_3 P^*(s_2) u$; $c_1 P^*(s_3)))\, c_3 P^*(s_3)\, c_2 P^*(s_3) u$; and $c_1 P^*(s_4)$ $c_2 P^*(s_4)\, c_3 P^*(s_4) u$. This information is summarized in Table I.

It is straightforward to verify that the C-optimal stable outcome in the set $C(P^*)$ is the outcome $x=x^*(P^*)$ such that $x(c_1)=\{s_3‘\, s_4\}$, $x(c_2)=\{s_2\}$, and $x(c_3)=\{s_1\}$. That is, $x^*(P^*)$ gives college c_1 its 3rd and 4th

TABLE I

P^*:	$P^*(c_1)$	$P^*(c_2)$	$P^*(c_3)$	$P^*(s_1)$	$P^*(s_2)$	$P^*(s_3)$	$P^*(s_4)$
	s_1	s_1	s_3	c_3	c_2	c_1	c_1
	s_2	s_2	s_1	c_1	c_1	c_3	c_2
	s_3	s_3	s_2	c_2	c_3	c_2	c_3
	s_4	s_4	s_4	u	u	u	u
	u	u	u				
Quotas:	$q_1=2$	$q_2=1$	$q_3=1$				

$$x^*(P^*)=x=[(c_1; s_3, s_4); (c_2; s_2), (c_3; s_1)]$$
$$y=[(c_1; s_2, s_4); (c_2; s_1), (c_3; s_3)]$$

Notes. Proposition 1. All colleges prefer y to $x^*(P^*)$. Proposition 2. $\{y\}=C(P')$, where P' results from c_1 misstating its preferences, so no stable matching procedure makes it a dominant strategy for each college to state its true preferences.

choice students, and gives colleges c_2 and c_3 each their second choice student.

Consider now the feasible outcome y such that $y(c_1) = \{s_2, s_4\}$, $y(c_2) = \{s_1\}$, and $y(c_3) = \{s_3\}$. The outcome y gives colleges c_2 and c_3 each their first choice student, so they both strictly prefer it to $x^*(P^*)$. Let us now examine college c_1, which is assigned its 3rd and 4th choice students at $x^*(P^*)$ and its 2nd and 4th choice students at y. Since college c has responsive preferences, it strictly prefers y to $x^*(P)$. Thus every college strictly prefers y to $x^*(P^*)$. This completes the demonstration that the conclusion of Theorem 3 is false in the college admissions problem when colleges have responsive preferences over outcomes.

The next section is devoted to discussing these matters with the additional precision that will be needed in order to properly consider the incentives facing the agents in the noncooperative game that arises when any stable matching procedure is adopted. It will be shown that when colleges have responsive preferences, the conclusion of Theorem 5 is also false; in fact there do not exist *any* stable matching procedures that make it a dominant strategy for colleges to state their true preferences.

4. The College Admissions Problem Revisited

In this section we will consider a specification of the college admissions problem in which all agents (both colleges and students) will have preferences over all possible outcomes. Only with such a specification can the problem be formulated as a well defined game.

As in the previous section, let there be two disjoint sets of agents C and S, with each c_i in C having a quota of q_i and a strict preference relation $P(c_i)$ on $S \cup \{u\}$, and each s_j in S having a strict preference relation $P(s_j)$ on $C \cup \{u\}$. Denote the vector of such preferences by $P = (P(c_1),...,P(c_n), P(s_1),..., P(s_p))$. An outcome x, an individually rational outcome, a stable outcome, and the set $C(P)$ of stable outcomes, are all defined precisely as before.

In addition, each college c_i has a preference relation $P^{\#}(c_i)$ on the set $\{x(c_i) \mid x \text{ is an outcome}\}$ of feasible assignments the college could receive. (So colleges have preferences defined over entire "entering classes," as well as over individual students.) The preferences of the agents over different outcomes x and y correspond precisely to their preferences over their own assignments at x and y; i.e., a college c_i prefers x to y if and only if $x(c_i) P^{\#}(c_i)\, y(c_i)$, and a student s_j prefers x to y if and only if $x(s_j)\, P(s_j)\, y(s_j)$. Denote by $P^{\#}$ the vector of preferences $P^{\#} = (P^{\#}(c_1),..., P^{\#}(c_n), P(s_1),..., P(s_p))$, which defines the preferences of the agents over all feasible outcomes.

Note that, since the college admissions problem reduces to the marriage problem in the special case that all $q_i = 1$, the conclusion of Theorem 4 carries over immediately when we consider the (now well-defined) game that arises from the adoption of a stable matching procedure that acts on agents' stated preferences.

THEOREM 4*. *There exists no stable matching procedure for the college admissions problem that makes it a dominant strategy for all agents to state their true preferences.*

Theorem 4* follows immediately from Theorem 4 and the fact that the marriage problem is a special case of the college admissions problem, so that if no procedure exists that fills the requirements of the theorem for a special case, then certainly no procedure exists that fills the requirements for the general case. For the case of responsive preferences, we will see that Proposition 2 below considerably strengthens the conclusions of Theorem 4*.

Let each college c's preferences $P^{\#}(c)$ over entering classes be *responsive* to its preferences $P(c)$ over students as defined in Section 3.Note that many different responsive preference orderings $P^{\#}(c)$ exist for any preference P(c), since, for example, responsiveness does not specify whether a college with a quota of 2 prefers to be assigned its 1st and 4th choice students instead of its 2nd and 3rd choice students. However, the preference ordering $P(c)$ over individual students can be derived from $P^{\#}(c)$ by considering a college c_i's preferences over assignments $x(c_i)$ containing no more than a single student (and $q_i - 1$ copies of u).

It is now straightforward to verify that the set $C(P)$ of stable outcomes (which depends only on the vector P) is equal to the core defined by *weak* domination (see [10]) of the cooperative game in which the preferences of the agents are given by the vector $\mathrm{P}^{\#}$, and whose rules are that any college and student may be matched with each other if they both agree, but no college may agree to be matched with more than its quota of students, no student may agree to be matched to more than one college, any student is free to remain unmatched, and any college is free to keep any of its positions unfilled.

Now consider the noncooperative game that arises when a stable matching procedure is adopted; i.e., a procedure that, for any stated preferences P, produces an outcome x in $C(P)$. Since the set $C(P)$ of stable outcomes is entirely determined by P, we can consider stable matching procedures that work by asking each college c to state its preference ordering $P(c)$ over individual students, rather than stating its full preference $P^{\#}(c)$ over groups of students. (Since $P(c)$ is completely determined by $P^{\#}(c)$, it can be thought of as a summary of the full preferences.) We can therefore consider the noncooperative game in which the strategy set of each college c consists of all the possible strict preferences $P(c)$ it might state, and similarly for each student. (The proposition below would be unchanged if we considered the game in which the strategy set for each college c was the set of all possible full preferences $P^{\#}(c)$ it might state.) It is shown below that the conclusion of Theorem 5 does not generalize to the college admissions problem: if P^* is the vector of true preferences (i.e., if, for each college c, $P^*(c)$ is the correct summary of college c's true full preferences), then, no matter what stable matching procedure is employed, it is *not* a dominant strategy for each college c to state $P^*(c)$.

PROPOSITION 2. *When colleges have responsive preferences, the conclusion of Theorem 5 is false for the college admissions problem: in fact no stable matching procedure exists that makes it a dominant strategy for each college to state its true preferences.*

The proof makes use of the same example used to prove Proposition 1.

Proof. Let the sets C and S of agents, the quotas, and the true preferences P^*, be those of the example used in the proof of Proposition 1 (see Table I). Then when all agents state their true preferences, the set $C(P^*)$ of stable outcomes contains a *unique* outcome, which is the outcome $x = x^*(P^*) = y^*(P^*)$. Thus any stable matching procedure must select the outcome x, and so college c_1 receives the assignment $x(c_1) = \{s_3, s_4\}$; i.e., it is assigned its 3rd and 4th choice students. Suppose now that college c_1 were to state instead the (false) preference ordering $P'(c_1)$ given by $s_2 P'(c_1)\, s_4 P'(c_1)\, u P'(c_1)\, s_1 P'(c_1) s_3$, and that all other agents were to state their true preferences, so that the vector of stated preferences is $P' = (P'(c_1), P^*(c_2), P^*(c_3), P^*(s_1), \ldots, P^*(s_4))$. It is straightforward to verify that the set $C(P')$ of stable outcomes with respect to P' also contains a unique outcome, which is the outcome $y = x^*(P') = y^*(P')$ described in the proof of Proposition 1. Thus any stable matching procedure must select the outcome y. Since c_1 thus receives the assignment $y(c_1) = \{s_2, s_4\}$ which it prefers to $x(c_1) = \{s_3, s_4\}$, it does better by stating $P'(c_1)$ than by stating its true preferences $P^*(c_1)$. This completes the proof.

Since the college admissions problem, unlike the marriage problem, is not symmetric between the two sides of the market, we also need to consider how things look from the student side of the market. For this, the related marriage problem constructed by Gale and Sotomayor (quoted above) will prove useful, since, unlike the colleges, the students retain their identity in the related problem, and so their preferences in the two problems are the same. This fact will make the proof of Theorem 3* and 5*, given below, immediate.

THEOREM 3*. *In the college admissions problem, there does not exist any outcome z that every student prefers to the S-optimal stable outcome y^*; i.e., for no outcome z is it the case that $z(s)\ P(s)\ y^*(s)$ for all s in S. However the corresponding result does not hold for the C-optimal stable outcome x^* when colleges have responsive preferences.*

THEOREM 5*. *The matching procedure that always yields the S-optimal stable outcome $y^*(P)$ for any stated preferences P makes it a dominant strategy for every s in S to state his true preferences in the admissions problem. However, when colleges have responsive preferences, no stable matching procedure makes it a dominant strategy for every c in C to state its true preferences.*

The proof of the first parts of Theorems 3* and 5* is immediate from the similar results for the marriage problem, through the Gale and Sotomayor construction, while the last sentence of each of these two theorems is simply a restatement of Propositions 1 and 2.

A final remark is in order about the Nash equilibria of the game arising from the adoption of some stable matching procedure such as one yielding the C-optimal stable outcome in terms of the stated preferences. In the marriage problem, since it is a dominant strategy for agents on one side of the market to state their true preferences, it is natural to study the Nash equilibria that arise when only the agents on the other side of the market misrepresent their preferences. It was shown in [7 and 2] (using slightly different formulations of the strategy sets of the agents) that these Nash equilibria result in outcomes that are stable with respect to the true preferences of the agents. In the college admissions problem, the situation is somewhat more complex, since a procedure yielding the C-optimal stable outcome does not make it a dominant strategy either for colleges or for students to state their true preferences. It should be noted, however, that every individually rational outcome x, whether stable or not, can be achieved as a Nash equilibrium in which each agent states that he prefers only his assignment at x to being unmatched.

In terms of the general theory of two-sided matching markets, the results presented here shed some new light on the family of markets of which the marriage problem, the college admissions problem, the labor markets studied in [5], and those studied in [8], are increasingly general examples. While Theorems 1, 2, and 4 generalize to all of these models (see [8]), the results presented here show that Theorems 3 and 5 do not generalize even to the college admissions problem, and hence not to the still more general models. This suggests that the existence of dominant-strategy stable procedures for one side of the market, exhibited in Theorem 5, may be less intimately connected than might have been supposed with the existence, exhibited in Theorem 2, of optimal stable outcomes for each side of the market.

REFERENCES

1. L. E. DUBINS AND D. A. FREEDMAN, Machiavelli and the Gale–Shapley algorithm, *Amer. Math. Montly* **88** (1981), 485–494.
2. D. GALE, Ms. Machiavelli and the Gale–Shapley algorithm, mimeo, 1983.
3. D. GALE AND L. SHAPLEY, College admissions and the stability of marriage, *Amer. Math. Monthly* **69** (1962), 9–15.
4. D. GALE AND M. SOTOMAYOR, Some remarks on the stable matching problem, mimeo, (1983).
5. A. S. KELSO, JR. AND V. P. CRAWFORD, Job matching, coalition formation, and gross substitutes, *Econometrica* **50** (1982), 1483–1504.
6. A. E. ROTH, The economics of matching: Stability and incentives, *Math. Oper. Res.* **7** (1982), 617–628.
7. A. E. ROTH, Misrepresentation and stability in the marriage problem, *J. Econ. Theory* **34** (1984), 383–387.
8. A. E. ROTH, Stability and polarization of interests in job matching, *Econometrica* **52** (1984), 47–57.
9. A. E. ROTH, The evolution of the labor market for medical interns and residents: A case study in game theory, *J. Polit. Econ.* **92** (1984), 991–1016.
10. A. E. ROTH, Common and conflicting interests in two-sided matching markets, *European Economic Review* (1985), forthcoming.