Roth, A.E. and Sotomayor, M.,
"The College Admissions Problem Revisited,"
Econometrica, 57, 1989, 559-570.

Econometrica, Vol. 57, No. 3 (May, 1989), 559–570

THE COLLEGE ADMISSIONS PROBLEM REVISITED

BY ALVIN E. ROTH AND MARILDA SOTOMAYOR[1]

The college admissions problem is perhaps the simplest model of many-to-one matching in two-sided markets such as labor markets. We show that the set of stable outcomes (which is equal to the core defined by weak domination) has some surprising properties not found in models of one-to-one matching. These properties may help to explain the success that this kind of model has had in explaining empirical observations.

KEYWORDS: College admissions problem, matching.

1. INTRODUCTION

THE "COLLEGE ADMISSIONS PROBLEM" is the name given to a two-sided matching problem by Gale and Shapley (1962). Colleges have preferences over students and students have preference over colleges; each college C can accept at most some number q_C of students, and each student can enroll in at most one college. The problem is to analyze what kinds of assignments might arise from such a market, with the primary theoretical tool being the set of stable outcomes (which is closely related to, and a subset of, the core) of the resulting game, and, more recently, the dominant strategy and Nash equilibria of the corresponding strategic game. This and related models have recently been employed in both theoretical and empirical studies of labor markets.

The case in which all the q_C equal 1 is called the "marriage problem," and is symmetric between the two sides of the market. It was initially thought that the essential features of the college admissions problem could be captured by treating it as a marriage problem in which each of the q_C positions available at a college C would be treated as q_C distinct individuals. Most of the subsequent theoretical literature concerned with these problems focused on the marriage problem, with the tacit or explicit assumption that results established for the marriage problem would carry over to the college admissions problem through this kind of transformation. That this is not the case was first observed in Roth (1984, 1985a). The second of those papers establishes that some of the important properties of marriage problems do not, as previously believed, carry over to the college admissions problem.[2] The present paper establishes some surprising results about the set of stable outcomes of the college admissions problem that have no parallel in the simpler case of the marriage problem.

[1]This work has been partially supported by grants from the National Science Foundation and the Office of Naval Research, by an Alfred P. Sloan Research Fellowship, and by the CNPq—The Conselho Nacional de Desenvolvimento Cientifico e Tecnologico, Brazil. We are grateful to Shmuel Zamir and George Svetlichny for numerous helpful discussions, and to Charles Blair and David Gale for careful reading of an earlier draft of this paper, and for their suggestions for shortening our proofs. The present proof of Lemma 3 was suggested by Gale.

[2]Treating a college admissions problem like a marriage problem did not adequately take into account that colleges are matched to groups of students, rather than to individuals. Roth (1985a) formulated a model of the college admissions problem that took this into account, and it is this (re)formulation that will be employed here.

1.1. *The Main Results*: *An Example*

Suppose students and colleges possess complete and transitive strict preferences over each other as individuals. That is, students have strict preferences over colleges, and colleges have strict preferences over students. Of course, a college which has, say, 1000 freshman places to fill must evaluate *groups* of students (i.e. whole entering classes), not merely individuals. If colleges' preferences over groups are complete and transitive, we will need to make no further assumption about how colleges evaluate groups of students beyond the simple one that, if a college is faced with two entering classes that differ in only a single student, then it prefers the class containing the more preferred student.

To introduce the main results by an example, suppose one of the colleges, college C, gives an entrance exam and evaluates students according to their scores on this exam, and evaluates entering classes according to their *average* score on the exam. (So even when we assume no two students have exactly the same score, so college C's preferences over individuals are strict, it does not have strict preferences over entering classes, since it is indifferent between two entering classes with the same average score.) Then the set of stable outcomes may contain many allocations, at which college C will have different entering classes. However, for this example our results imply that *no two distinct entering classes that college C could have at stable outcomes will have the same average exam score*. Furthermore, for any two distinct entering classes that college C could be assigned at stable outcomes, we can make the following strong comparison. Aside from the students who are in both entering classes, *every student in one of the entering classes will have a higher exam score than any student in the other entering class*.

More generally, the results say that in a college admissions problem in which all preferences over individuals are strict, no college will be indifferent between any two (different) groups of students that it enrolls at stable outcomes. Furthermore, for every pair of stable outcomes, each college will prefer every student who is assigned to it at one of the two outcomes to every student who is assigned to it in the second outcome but not the first. These results are mathematically unusual, and also have significant implications for the use of these models in understanding empirical economic phenomena.

The manner in which they are mathematically unusual can be understood by the following observation. The first result implies that if a given matching is stable (and hence in the core), and if some college is indifferent between the entering class it is assigned at that matching and a different entering class it is assigned at another matching, then the second matching is *not* in the core. We thus have a way of concluding that an outcome is not in the core, based on the direct examination of the preferences of only *one* agent (the college). Since the definition of the core involves preferences of coalitions of agents, this is surprising. In fact we know of no comparable results concerning the core of a game (except in the trivial case of games which have at most a single core outcome).

The significance of these results for using such a model to study real markets bears on the issues that arise when an organization is modelled as an individual rather than as a collection of agents. While it may often be an acceptable approximation to model individuals within an organization as sharing the same preferences over candidates (e.g. when the information about candidates consists primarily of things like exam scores), more is involved in assuming that they have the same preferences over entering classes. For example, in the study of the hospital-intern labor market reported in Roth (1984, 1986), the agents on one side of the market are modelled as being the programs within a hospital that offer a particular kind of internship. Several physicians are typically associated with each program, and one can imagine that some might evaluate an entering class by paying most attention to the quality of its best members, while others might evaluate entering classes by weighing most heavily the quality of the weakest members. Our results demonstrate that this kind of divergence of preferences cannot arise in connection with comparison of stable matchings, since, for entering classes that could be admitted at stable outcomes, the rank ordering of entering classes in terms of their best member is the same as the rank ordering in terms of their worst member,[3] for example.

2. THE FORMAL MODEL

The first elements of the model are two finite, disjoint sets, $C = \{C_1, \ldots, C_n\}$ and $S = \{s_1, \ldots, s_m\}$, of colleges and students, respectively. The rules of the game are that any student and college may mutually agree that the student will attend the college, any college may choose to keep any of its positions unfilled, and any student may remain unmatched if he wishes. For each college C, there is a positive integer q_C called its *quota*, which is the number of positions it has to offer. An outcome is a matching of students to colleges, such that each student is matched to at most one college, and each college is matched to at most its quota of students. A student who is not matched to any college will be "matched to himself," and a college that has some number of unfilled positions will be matched to itself in each of those positions.

To give a formal definition, we first define, for any set X, an *unordered family of elements of* X to be a collection of elements, *not necessarily distinct*, in which the order is immaterial. So a given element of X may appear more than once in an unordered family of elements of X, which is what distinguishes an unordered family from a subset of X.

DEFINITION 1: A *matching* μ is a function from the set $C \cup S$ into the set of unordered families of elements of $C \cup S$ such that (i) $|\mu(s)| = 1$ for every student

[3]Modulo indifference, as when two potential entering classes have the same individual as the best member, but different worst members. This will not make a difference in the preferences for groups: see Corollary 1.

s and $\mu(s) = s$ if $\mu(s) \notin C$; (ii) $|\mu(C)| = q_C$ for every college C, and if $|\mu(C) \cap S| = r < q_C$ then $\mu(C)$ contains $q_C - r$ copies of C; (iii) $\mu(s) = C$ if and only if s is in $\mu(C)$.

So $\mu(s_1) = C$ denotes that student s_1 enrolls at college C at the matching μ, and $\mu(C) = \{s_1, s_3, C, C\}$ denotes that college C, with quota $q_C = 4$, enrolls students s_1 and s_3 and has two positions unfilled. We will represent matchings graphically; for example,

$$\mu_s \quad \begin{matrix} C_1 & C_2 & (s_4) \\ \\ s_1 s_3 C_1 & s_2 & s_4 \end{matrix}$$

represents a matching at which college C_1, with quota $q_C = 3$, is matched with two students, s_1 and s_3, college C_2 with a quota of 1 is matched with s_2, and s_4 is unmatched.

Each student has preferences over the colleges, and each college has preferences over the students. We will assume these preferences are complete, transitive, and strict, and so may be represented by ordered lists, with $P(C) = s_1, s_2, C, s_3, \ldots$ denoting that college C prefers s_1 rather than s_2, that it prefers either one of them rather than leave a position unfilled, and that all other students are unacceptable, in the sense that it would be preferable to leave a position unfilled rather than filling it with, say, student s_3. Similarly, $P(s) = C_2, C_1, C_3, s, \ldots$ represents the preferences of a student s. We write $C_i >_s C_j$ to indicate that student s prefers C_i to C_j. Similarly, $s_i >_C s_j$ represents college C's preferences $P(C)$ over individual students. College C is *acceptable* to student s if $C >_s s$, and student s is acceptable to college C if $s >_C C$. We will usually abbreviate a preference list to include just the acceptable alternatives.

2.1. *Preferences Over Matchings*

Each student's preferences over matchings correspond to his preferences over his own assignments at the matchings. We cannot yet say this about colleges, because, as noted in Roth (1985a), while we have described colleges' preferences over students, each college with a quota greater than 1 must be able to compare *groups* of students in order to compare matchings, and we have yet to describe the preferences of colleges over groups.

The only assumption we will need connecting colleges' preferences over groups to their preferences over individuals is one insuring that, for example, if $\mu(C)$ assigns college C its 3rd and 4th choice students, and $\mu'(C)$ assigns it its *2nd* and 4th choice, then college C prefers $\mu'(C)$ to $\mu(C)$. Specifically, let $P^{\#}(C)$ denote the preferences of college C over all assignments $\mu(C)$ that it could receive at any matching μ of the college admissions problem. Following Roth (1985a), a college C's preferences $P^{\#}(C)$ will be called *responsive* to its preferences $P(C)$ over individual students if, for any two assignments that differ in only one student, it prefers the assignment containing the more preferred student.

DEFINITION 2: The preference relation $P^{\#}(C)$ over groups of students is *responsive* (to the preferences $P(C)$ over individual students) if, whenever $\mu'(C) = \mu(C) \cup \{s\} \setminus \{\sigma\}$ for σ in $\mu(C)$ and s not in $\mu(C)$ such that $s >_C \sigma$, then C prefers $\mu'(C)$ to $\mu(C)$ (under $P^{\#}(C)$).

We will write $\mu'(C) >_C \mu(C)$ to indicate that college C prefers $\mu'(C)$ to $\mu(C)$ according to its preferences $P^{\#}(C)$, and $\mu'(C) \geqslant_C \mu(C)$ to indicate that C likes $\mu'(C)$ at least as well as $\mu(C)$, where the fact that $\mu'(C)$ and $\mu(C)$ are not singletons will make clear that we are denoting the preferences $P^{\#}(C)$, as distinct from statements about C's preferences over individual students. (Note that C may be indifferent between distinct assignments $\mu(C)$ and $\mu'(C)$.)

Different responsive preferences $P^{\#}(C)$ exist for any preference $P(C)$; e.g. responsiveness does not specify whether a college with a quota of 2 prefers to be assigned its 1st and 4th choice students or its 2nd and 3rd choices. However any preference $P^{\#}(C)$ can be responsive to (at most) a unique preference ordering $P(C)$ over individual students (since $P(C)$ can be derived from $P^{\#}(C)$ by considering C's preferences over assignments $\mu(C)$ containing no more than a single student). The assumption that colleges have responsive preferences connects their preferences for groups of students to their rankings of individuals in a natural way.[4] We will henceforth assume that colleges have complete and transitive preferences over groups that are responsive to their preferences over individuals, and that each agent's preferences over matchings correspond exactly to his (its) preferences over his (its) own assignments at the matchings.

3. STABILITY AND GROUP STABILITY

A matching μ is *individually irrational* if $\mu(s) = C$ for some student s and college C such that either the student is unacceptable to the college or the college is unacceptable to the student. Such a matching will also be said to be *blocked* by the unhappy agent. Similarly, a college C and student s will be said to together block a matching μ if they are not matched to one another at μ, but would both prefer to be matched to one another than to (one of) their present assignments. That is, μ is *blocked by the college-student pair* (C, s) if $\mu(s) \neq C$ and if $C >_s \mu(s)$ and $s >_C \sigma$ for some σ in $\mu(C)$.

[4] Of course, the assumption that colleges *have* preferences over individual students, and not merely over entering classes, is not trivial. It precludes, for example, allowing colleges to express a preference for a geographically diverse student body, as well as the kinds of complementarities that may exist in some labor markets. The work of Kelso and Crawford (1982) suggests that weaker assumptions may be possible. A referee has noted the resemblance of the assumption of responsiveness to that of independence in utility theory.

DEFINITION 3: A matching μ is *stable* if it is not blocked by any individual agent or any college-student pair.

It isn't obvious that this definition, which is the same as the one for the marriage problem, will be adequate for the college admissions problem, since we now might need to consider coalitions consisting of colleges and several students (all of whom might be enrolled simultaneously at the college), or even coalitions consisting of multiple colleges and students. We shall now consider these larger coalitions, and see that nothing is lost by concentrating on simple college-student pairs.

We will say that a matching μ is *group unstable*, or that it is *blocked by a coalition* A, if there exists another matching μ' and a coalition A, which might consist of multiple students and/or colleges, such that, for all students s in A, and for all colleges C in A, (i) $\mu'(s) \in A$ (i.e. every student in the coalition who is matched by μ' is matched to a college in the coalition); (ii) $\mu'(s) >_s \mu(s)$ (i.e. every student in the coalition prefers his new match to his old one); (iii) $\sigma \in \mu'(C)$ implies $\sigma \in A \cup \mu(C)$ (i.e., every college in the coalition draws new students only from the coalition, although it may continue to be matched with some of its 'old' students from $\mu(C)$); and (iv) $\mu'(C) >_C \mu(C)$ (i.e., every college in the coalition prefers its new entering class of students to its old one).

That is, μ is blocked by some coalition A of colleges and students if, by matching among themselves, the students and colleges in A could all get an assignment preferable to μ.

DEFINITION 4: A *group stable* matching is one that is not blocked by any coalition.

We will now see that (when preferences are responsive) this definition of group stability[5] is equivalent to our definition of (pairwise) stability.

PROPOSITION 1: *A matching is group stable if and only if it is stable.*

PROOF: If μ is unstable via an individual student or college, or via a student college pair, then it is clearly group unstable via the coalition consisting of the same singleton or pair. In the other direction, if μ is blocked via coalition A and outcome μ', let C be in A. Then the fact that $\mu'(C) >_C \mu(C)$ implies that there exists a student s in $\mu'(C)\backslash\mu(C)$ and a σ in $\mu(C)\backslash\mu'(C)$ such that $s >_C \sigma$. (Otherwise, $\sigma \geqslant_C s$ for all σ in $\mu(C)\backslash\mu'(C)$ and s in $\mu'(C)\backslash\mu(C)$, and this would imply $\mu(C) \geqslant_C \mu'(C)$, by repeated application of the fact that preferences

[5]The difference between group instability and domination via a coalition A is the requirement in the definition of domination that all members of $\mu'(C)$ be in A rather than in $A \cup_\mu (C)$. (The set of group stable outcomes is the subset of the core equal to the core defined by *weak* domination (c.f. Roth (1985b)).

are responsive and transitive.) So s is in A and s prefers C to $\mu(s)$, so μ is unstable via s and C. *Q.E.D.*

4. THE CONNECTION BETWEEN THE COLLEGE ADMISSIONS PROBLEM AND THE MARRIAGE PROBLEM

The importance of Proposition 1 for the college admissions problem goes beyond that it allows us to concentrate on small coalitions. It says that group stable matchings can be identified using only the preferences P over *individuals*—i.e., without knowing the preferences $P^{\#}(C)$ that each college has over entering classes. This follows since stability and group stability are equivalent. So, for fixed preferences over individuals, the set of stable (or group stable) matchings will not be sensitive to changes in the preferences $P^{\#}(C)$ (so long as these preferences remain responsive to the preferences $P(C)$). This suggests that the college admissions problem may be very similar indeed to the marriage market. This is an issue about which there has been considerable confusion in the literature. We turn now to consider it.

4.1. *The Related Marriage Market*

Consider a particular college admissions problem, with colleges $C = \{C_1, \dots, C_n\}$ having quotas $q_1, \dots, q_n$, and students $S = \{s_1, \dots, s_m\}$. The preferences of students and colleges over individuals are given by $P = \{P(C_1), \dots P(C_n); P(s_1), \dots P(s_m)\}$.

There is a related marriage problem, in which each college C with quota q_C is broken into q_C "pieces," so that in the related marriage problem the agents will be college positions and students, each having a quota of 1. That is, we replace college C by q_C positions of C denoted by $c_1, c_2, \dots, c_{q_C}$. Each of these positions has preferences over individuals that are identical with those of C. Since each position has a quota of 1, we do not need to consider its preferences over groups of students. In the related marriage problem, each student's preference list is modified by replacing C, wherever it appears on his list, by the string $c_1, c_2, \dots, c_{q_C}$, in that order.

There is a natural one-to-one correspondence between matchings in the original college admissions problem and matchings in the marriage problem related to it in this way. A matching μ of the college admissions problem, which matches a college C with the students in $\mu(C)$, corresponds to the matching $\bar{\mu}$ in the related marriage market in which the students in $\mu(C)$ are matched, in the order which they occur in the preferences $P(C)$, with the ordered positions of C that appear in the related marriage market. (That is, if s is C's most preferred student in $\mu(C)$, then $\bar{\mu}(s) = c_1$, while C's second most preferred student in $\mu(C)$ is matched to c_2 and so forth.) This correspondence preserves the stability of the matching. That is, we have the following lemma, whose proof we leave to the reader.

LEMMA 1: *A matching of the college admissions problem is stable if and only if the corresponding matching of the related marriage problem is stable.*

Thus the theorem that the set of stable matchings is nonempty for every marriage problem will immediately generalize to the case of the college admissions problem via Lemma 1. However, since stable matchings can be identified without regard to the preferences of colleges over groups of students, this result does not permit us to directly conclude anything about the preferences of colleges for different (stable or unstable) matchings. Thus results for the marriage problem which compare different matchings will have to be considered again. We will see that many of these theorems do in fact generalize to the college admissions problem, sometimes with additional power.[6]

5. SOME USEFUL RESULTS

The following results will be used in proving our main results. The first was proved independently in Roth (1984) and Gale and Sotomayor (1985), in the context of the marriage problem. (The statement given below for the college admissions problem is immediate via Lemma 1.) The second comes from Roth (1986).

THEOREM 1: *When all preferences over individuals are strict, the set of students enrolled and positions filled in a college admissions problem is the same at every stable matching.*

THEOREM 2: *When all preferences over individuals are strict, any college that does not fill its quota at some stable matching is assigned precisely the same set of students at every stable matching.*

LEMMA 2 (*Decomposition Lemma*):[7] *Let μ and μ' be stable matchings in a marriage problem $(M, W, \boldsymbol{P})$. Let $M(\mu)$ $(M(\mu'))$ be all men who prefer μ to μ' (μ' to μ) and define sets of women $W(\mu')$ and $W(\mu)$ analogously. Then μ' and μ are both one-to-one correspondences between $M(\mu')$ and $W(\mu)$, and between $M(\mu)$ and $W(\mu')$. That is, both μ' and μ match any man who prefers one of the two stable matchings to a woman who prefers the other, and vice versa.*

6. THE MAIN RESULTS

LEMMA 3: *Suppose that colleges and students have strict individual preferences, and let μ and μ' be stable matchings for $(S, C, \boldsymbol{P})$, such that $\mu(C) \neq \mu'(C)$ for some C. Let $\bar{\mu}$ and $\bar{\mu}'$ be the stable matchings corresponding to μ and μ' in the*

[6]But care must be taken, since some results that hold for the marriage problem turn out not to hold for the college admissions problem. (For example, it was shown in Roth (1985a) that the college-optimal stable matching may not be Pareto optimal with respect to the set of colleges, although in the marriage problem the optimal stable matching for each side of the market is Pareto optimal with respect to the agents on that side.)

[7]A version of this result is found in Knuth (1976). A stronger result, involving comparisons of stable matchings in related marriage markets, appears in Gale and Sotomayor (1985).

related marriage market. Then if $\bar{\mu}(c_i) >_C \bar{\mu}'(c_i)$ for some position c_i of C it follows that $\bar{\mu}(c_j) \geqslant_C \bar{\mu}'(c_j)$ for all positions c_j of C.

PROOF: It is enough to show that $\bar{\mu}(c_j) >_C \bar{\mu}'(c_j)$ for all $j > i$. So suppose this is false. Then there exists an index j such that $\bar{\mu}(c_j) >_C \bar{\mu}'(c_j)$, but $\bar{\mu}'(c_{j+1}) \geqslant_C \bar{\mu}(c_{j+1})$. Theorem 1 implies $\bar{\mu}'(c_j) \in S$. Let $s' \equiv \bar{\mu}'(c_j)$. By the Decomposition Lemma $c_j \equiv \bar{\mu}'(s') >_{s'} \bar{\mu}(s')$. Furthermore, $\bar{\mu}(s') \neq c_{j+1}$, since $s' >_C \bar{\mu}'(c_{j+1}) \geqslant_C \bar{\mu}(c_{j+1})$ (where the first of these preferences follows from the fact that for any stable matching $\bar{\mu}'$ in the related marriage market, $\bar{\mu}'(c_j) >_C \bar{\mu}'(c_{j+1})$ for all j). Therefore $c_{j+1} >_{s'} \bar{\mu}(s')$, since c_{j+1} comes right after c_j in the preferences of s' (or any s) in the related marriage problem. So $\bar{\mu}$ is blocked via s' and c_{j+1}, contradicting (via Lemma 1) the stability of μ. *Q.E.D.*

We remark that the proof of Lemma 3 shows that if $\bar{\mu}(c_i) >_C \bar{\mu}'(c_i)$ for some position c_i of C then $\bar{\mu}(c_j) >_C \bar{\mu}'(c_j)$ for all $j > i$. This observation gives an alternative way to prove Theorem 2. To see this, recall that if a college C has any unfilled positions, these will be the highest numbered c_j at any stable matching of the corresponding marriage problem, and by Theorem 1 these positions will be unfilled at any stable matching, i.e. $\bar{\mu}(c_j) = \bar{\mu}'(c_j)$ for all such j, and hence for all j, by the above remark.

Since colleges' preferences over groups are responsive to their preferences over individuals, the following result follows from Lemma 3.

THEOREM 3: *If colleges and students have strict preferences over individuals, then colleges have strict preferences over those groups of students that they may be assigned at stable matchings. That is, if μ and μ' are stable matchings, then a college C is indifferent between $\mu(C)$ and $\mu'(C)$ only if $\mu(C) = \mu'(C)$.*

PROOF: If $\mu(C) \neq \mu'(C)$, then (without loss of generality) $\bar{\mu}(c_i) >_C \bar{\mu}'(c_i)$ for some position c_i of C. By Lemma 3, $\bar{\mu}(c_j) \geqslant_C \bar{\mu}'(c_j)$ for all positions c_j of C, where $\bar{\mu}$ and $\bar{\mu}'$ are the matchings in the related marriage market corresponding to μ and μ'. So $\mu(C) >_C \mu'(C)$, by repeated application of the fact that C's preferences are responsive and transitive. (First compare $\mu'(C)$ with a matching that agrees with $\mu(C)$ on position c_1 (in the related marriage problem) and with $\mu'(C)$ for all other positions, then compare this new matching with a matching that agrees with $\mu(C)$ on positions c_1 and c_2, and with $\mu'(C)$ on all other positions, etc. Responsiveness of preferences determines each pairwise comparison in the resulting chain, and transitivity then assures the desired result.) *Q.E.D.*

Before proceeding, let us pause to consider what we have learned. Consider a college C with $q_C = 2$ and preferences $P(C) = s_1, s_2, s_3, s_4$. Let μ and ν be matchings such that $\mu(C) = \{s_1, s_4\}$ and $\nu(C) = \{s_2, s_3\}$. Then without knowing anything about the preferences of students and other colleges, we can conclude that μ and ν cannot both be stable. Note that this is so even though C may either prefer one of $\mu(C)$ or $\nu(C)$ to the other, or be indifferent between them.

Recall that, for fixed preferences over individuals, the set of stable (or group stable) matchings is not sensitive to changes in colleges' preferences for groups of students (so long as those preferences remain responsive to the colleges' preferences over individuals). Theorem 3 therefore tells us something not only about the preferences each college actually has for groups of students that it may be assigned at stable matchings, but also about all the different preferences for groups that it *could* have, given its preferences over individuals. That is, let $P(C)$ be college C's preferences over individual students. Then Theorem 3 tells us that if μ and μ' are both stable matchings, then no preferences that are responsive to $P(C)$ can be indifferent between the groups of students $\mu(C)$ and $\mu'(C)$. The following Theorem shows why this is so.

THEOREM 4: *Let μ and μ' be stable matchings for $(S, C, \boldsymbol{P})$). If $\mu(C) >_C \mu'(C)$ for some college C, then $s >_C s'$ for all $s \in \mu(C)$ and $s' \in \mu'(C) - \mu(C)$. That is, C prefers every student in its entering class at μ to every student who is in its entering class at μ' but not at μ.*

PROOF: Consider the related marriage market $(S, C', \boldsymbol{P})$ and the stable matchings $\bar{\mu}$ and $\bar{\mu}'$ corresponding to μ and μ'. Let $q_C = k$, so that the positions of C are $c_1, \ldots, c_k$. First observe that C fills its quota under μ and μ', since, if not, Theorem 2 would imply that $\mu(C) = \mu'(C)$. So $\mu'(C) - \mu(C)$ is a nonempty subset of S, since $\mu(C) \neq \mu'(C)$. Let $s' = \bar{\mu}'(c_j)$ for some position c_j such that $s' \notin \mu(C)$. Then $\bar{\mu}(c_j) \neq \bar{\mu}'(c_j)$. By Lemma 3 $\bar{\mu}(c_j) >_C \bar{\mu}'(c_j)$. The Decomposition Lemma implies $c_j >_{s'} \bar{\mu}(s')$. So the construction of the related marriage problem implies $C >_{s'} \mu(s')$, since $\mu(s') \neq C$. Thus $s >_C s'$ for all $s \in \mu(C)$ by the stability of μ, which completes the proof. *Q.E.D.*

To illustrate what Theorem 4 adds to what we already know, consider again a college C with $q_C = 2$ and preferences $P(C) = s_1, s_2, s_3, s_4$. Consider two matchings μ and ν such that $\mu(C) = \{s_1, s_3\}$ and $\nu(C) = \{s_2, s_4\}$. Then the Theorem says that if μ is stable, ν is not, and vice versa.

The following corollary follows immediately from the theorem and the definition of responsive preferences.

COROLLARY 1: *Consider a college C with preferences $P(C)$ over individual students, and let $P^{\#}(C)$ and $P^*(C)$ be preferences over groups of students that are responsive to $P(C)$ (but are otherwise arbitrary). Then for every pair of stable matchings μ and μ', $\mu(C)$ is preferred to $\mu'(C)$ under the preferences $P^{\#}(C)$ if and only if $\mu(C)$ is preferred to $\mu'(C)$ under $P^*(C)$.*

Corollary 1 formalizes our introductory comments about why different individuals (having the same preferences over individual students) can be modelled as a single agent, e.g. college C, in a model of this kind.

The following example will illustrate the results of this section. Let the preferences be given by

s_1: C_5, C_1	C_1: $s_1, s_2, s_3, s_4, s_5, s_6, s_7$
s_2: C_2, C_5, C_1	C_2: s_5, s_2
s_3: C_3, C_1	C_3: s_6, s_7, s_3
s_4: C_4, C_1	C_4: s_7, s_4
s_5: C_1, C_2	C_5: s_2, s_1
s_6: C_1, C_3	
s_7: C_1, C_3, C_4	

and let the quotas be

$$q_{C_1} = 3, \quad q_{C_j} = 1 \qquad (j = 2, \ldots, 5).$$

Then the set of stable outcomes is $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ where

	C_1	C_2	C_3	C_4	C_5
$\mu_1 =$	$s_1s_3s_4$	s_5	s_6	s_7	s_2
$\mu_2 =$	$s_3s_4s_5$	s_2	s_6	s_7	s_1
$\mu_3 =$	$s_3s_5s_6$	s_2	s_7	s_4	s_1
$\mu_4 =$	$s_5s_6s_7$	s_2	s_3	s_4	s_1

Note that these are the only stable matchings, and $\mu_1(C_1) >_{C_1} \mu_2(C_1) >_{C_1} \mu_3(C_1) >_{C_1} \mu_4(C_1)$, for any responsive preferences.

Further consequences of these results can be found in Roth and Sotomayor (1989).

Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

Manuscript received September, 1987; final revision received June, 1988.

REFERENCES

GALE, DAVID, AND LLOYD SHAPLEY (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

GALE, DAVID, AND MARILDA SOTOMAYOR (1985): "Some Remarks on the Stable Matching Problem," *Discrete Applied Mathematics*, 11, 223–232.

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

KNUTH, DONALD E. (1976): *Marriage Stables*. Montreal: Les Presses de l'Universite de Montreal.

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

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

——— (1985b): "Common and Conflicting Interests in Two-Sided Matching Markets," *European Economic Review* (Special issue on Market Competition, Conflict, and Collusion), 27, 75–96.

——— (1986): "On the Allocation of Residents to Rural Hospitals: A General Property of Two-Sided Matching Markets," *Econometrica*, 54, 425–427.

ROTH, ALVIN E., AND MARILDA SOTOMAYOR (1989): *Two Sided Matching Markets: A Study in Game-Theoretic Modelling and Analysis*, Econometric Society Monograph. Cambridge, MA: Cambridge University Press.