Econometrica, Vol. 54, No. 2 (March, 1986), pp. 425-427

NOTES AND COMMENTS

ON THE ALLOCATION OF RESIDENTS TO RURAL HOSPITALS: A GENERAL PROPERTY OF TWO-SIDED MATCHING MARKETS

BY ALVIN E. ROTH[1]

EACH YEAR, A HIGH PROPORTION of medical students seeking employment as residents in American hospitals (upon graduating from both American and foreign medical schools), and American hospitals seeking to employ these students as residents, voluntarily participate in a centralized job matching clearinghouse now known as the National Resident Matching Program (NRMP).[2] A number of hospitals, particularly rural hospitals, fail each year to fill as many positions as they have available, and find that a high percentage of the positions they do fill are filled by foreign medical school graduates. It has been suggested that changes in the manner in which the clearinghouse treats hospitals and students might alter this situation.[3] However it was shown in Roth [**4**] that any two outcomes that are *stable*—the relevant equilibrium notion[4] for this kind of market—fill the same number of positions at any hospital. Since that paper also showed that the clearinghouse procedure yields a stable outcome, any change in procedure that preserves this property would thus have no effect on the perceived *numerical* maldistribution of physicians among hospitals.

Here it is shown that any hospital that fails to fill all of its positions at some stable outcome will not only fill the same *number* of positions at any other stable outcome, but will fill them with exactly the same residents. Thus, while the staffs of other hospitals are determined by which of the multiple equilibria of such a market is reached, the situation of hospitals whose positions are not all filled remains unaffected. The maldistribution of physicians, and particularly of American educated physicians, is therefore a property of equilibria of this kind of market, and not an artifact of the particular equilibrium presently selected.

The formal model:[5] The agents in the hospital-intern market consist of two disjoint sets $H=\{h_1,\ldots,h_n\}$ and $S=\{s_1,\ldots,s_p\}$ (hospitals and students). Each hospital h_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 $H\cup\{u\}$, and each hospital h has a strict

[1] 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.

[2] This clearinghouse was called the National Intern Matching Program (NIMP) when it was first established in 1951. It operates by having students and hospitals submit rank-orderings of one another indicating their preferences, which are then used to produce a matching of students and hospitals. The history of the market failures that preceded the establishment of this central clearinghouse, and the operation of the market and the clearinghouse since it was established, are discussed and analyzed in Roth [**4**].

[3] In particular it has been suggested that reversing the role of hospitals and students in the clearinghouse procedure would aggravate this situation: see Sundarshan and Zisook [**10**], or the discussion of their claim in Roth [**4**].

[4] Perhaps the phrase "solution concept" would more clearly connote that stable outcomes are considered here at the level of modelling of cooperative game theory, rather than as a strategic (Nash) equilibrium of the kind considered in noncooperative game theory, when games are modelled by their more detailed strategic form. Indeed, we will see that the set of stable outcomes is closely related to the core of the game. It was further argued in Roth [**4**] that any clearinghouse procedure in which participation is voluntary, as is the case with the NRMP, must yield stable outcomes if it is to enjoy a high rate of participation, since otherwise it gives students and hospitals an incentive to arrange matches outside of the system.

[5] See Roth [**4**] for a more detailed discussion of the market and how it is modelled.

preference ordering $P(h)$ over the set $S \cup \{u\}$, where u denotes the possibility of remaining unmatched.[6] Let $h_j P(s) h_k$ denote that student s prefers hospital h_j to hospital h_k, and let $h_j R(s) h_k$ denote that he either prefers h_j to h_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 hospitals, and $P = (P(h_1), \ldots, P(h_n), P(s_1), \ldots, P(s_p))$ will denote the vector of preference orderings of each agent for agents on the other side of the market. An *outcome* of the market is defined by a correspondence $x: H \cup S \to H \cup S \cup \{u\}$ such that $|x(s)| = 1$ for all s in S, $|x(h_i)| = q_i$ for all h_i in H, and, for any h in H and s in S, $x(s) = h$ if and only if s is an element of $x(h)$. 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 hospital h with quota q is assigned some number $k < q$ of students at an outcome x (i.e., if $|x(h) \cap S| = k < q$), then $q - k$ elements of $x(h)$ are equal to u. No student is assigned to more than one place, and no hospital is assigned more than its quota of students.[7]

An outcome x is *individually rational* if for every students s, $x(s) R(s) u$, and if for every hospital h and σ in $x(h)$, $\sigma R(h) u$. An outcome x is *unstable* if it is not individually rational or if there exist a hospital h and a student s who are not matched at x ($x(s) \neq h$) and who each prefer one another to one of their assignments, i.e., such that $hP(s)x(s)$ and $sP(h)\,\sigma$ for some σ in $x(h)$. An outcome x that is not unstable is *stable*, and the set of stable outcomes with respect to any vector P of preference orderings will be denoted $S(P)$. It was shown in Roth [7] that, when hospitals have responsive preferences, the set of stable outcomes equals the core (defined by weak domination) of this market when the rules are that any student and hospital may sign an employment contract if they both agree. The set $S(P)$ of stable outcomes is always nonempty.

THEOREM: *Let h be a hospital with quota q, and let x be a stable outcome such that $x(h)$ contains fewer than q students; i.e. $|x(h) \cap S| < q$. Then, for any other stable outcome y, $y(h) = x(h)$.*

LEMMA 1 (Roth [**4**]):[8] *The number of positions filled by each hospital is the same at every stable outcome. That is, if for some stable outcome x and hospital h, $|x(h) \cap S| = k$, then for any other stable outcome y, $|y(h) \cap S| = k$ also. (The set of students who are assigned positions is also the same at every stable outcome.)*

For any hospital h with quota q, let its *choice set* from any set T of acceptable[9] students be h's q most preferred students in T if $|T| \geq q$, and be the entire set T if $|T| \leq q$. Denote this choice set by $C_h(T)$.[10]

[6] See Roth [**4, 5**] for a discussion of this market that includes nonstrict preferences.

[7] Students' preferences over outcomes correspond precisely to their preferences over hospitals, so that student s prefers outcome x to outcome y if and only if he prefers hospital $x(s)$ to hospital $x(y)$. The preferences of hospitals over outcomes are necessarily related to their preferences over students in a more complex way, since a hospital h with quota $q > 1$ receives different *sets* of students and vacancies at different outcomes. (The fact that hospitals may have quotas greater than 1 makes this problem different from the "marriage problem" in which both sides of the market have quotas of 1. That problem, first formally studied by Gale and Shapley [**1**], was for some time thought to be mathematically equivalent to this more general problem, but this turns out not to be the case in a number of important respects (see Roth [**6**]). However it will be sufficient for the purpose of this paper to consider only the hospitals' preferences over individual students, so long as their preferences over outcomes have the property that, for any two outcomes that assign some hospital two sets of students that differ in only one student, the hospital prefers the outcome that gives it the more preferable student. (Such preferences are called *responsive* in Roth [**4, 6**].) It will henceforth be assumed that hospitals have responsive preferences.

[8] An independent proof of Lemma 1 also appears in Gale and Sotomayor [**2**].

[9] A student s is acceptable to hospital h if h prefers s to leaving a position unmatched, i.e. if $sP(h)u$.

[10] For simplicity of notation, we will define an outcome w below by expressions of the form $w(h) = C_h(T)$, instead of the more precise $\{w(h) \cap S\} = C_h(T)$. When $|T| < q$, the remaining positions in $w(h)$ are of course unmatched, i.e., equal to u.

LEMMA 2:[11] *Let x and y be outcomes in $S(P)$. Then there is a feasible outcome w at which each hospital h is matched with its choice set from $\{x(h) \cup y(h)\}$, i.e., $w(h) = C_h(x(h) \cup y(h))$ for every hospital h. Furthermore, w is a stable outcome; i.e., it is contained in $S(P)$.*

PROOF OF THE THEOREM: Let x and y be stable outcomes, and let h be a hospital with quota q such that $|x(h) \cap S| = k < q$. Let w be the stable outcome constructed in Lemma 2, so that $w(h) = C_h(x(h) \cup y(h))$. Then, by Lemma 1, $|x(h) \cap S| = |y(h) \cap S| = |w(h) \cap S| = k$. But, by the definition of choice sets $C_h(T)$, we know that $|w(h) \cap S| = \min[q, |\{x(h) \cup y(h)\} \cap S|$. Therefore the total number of students in $x(h) \cup y(h)$ is equal to the number of students in either $x(h)$ or $y(h)$ alone; therefore $x(h) = y(h)$.

Thus the perceived maldistribution of residents among hospitals remains the same at every stable outcome. Any unstable outcome that might be proposed gives students and hospitals (specifically, those students and hospitals with respect to which the outcome is unstable) the incentive to find mutually preferable matches. Since the history of the market prior to the adoption of a procedure yielding stable outcomes shows that agents in this market are prepared to act on these incentives when given the opportunity (c.f. Roth [**4**]), this maldistribution seems unlikely to be changed by any system that does not involve some element of compulsion, or some change in the relative numbers of available positions and eligible students.

University of Pittsburgh

Manuscript received January, 1985; *revision received April,* 1985.

REFERENCES

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

[**2**] GALE, DAVID, AND MARILDA SOTOMAYOR: "Ms. Machiavelli and the Stable Matching Problem," *American Mathematical Monthly*, 92(1985), 261-268.

[**3**] KNUTH, DONALD E.: *Mariages Stables.* Montreal: Les Presses de l'Universite de Montreal, 1976.

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

[**5**] ———: "Misrepresentation and Stability in the Marriage Problem," *Journal of Economic Theory*, 34(1984), 383-387.

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

[**7**] ———: "Common and Conflicting Interests in Two-Sided Matching Problems," *European Economic Review*, 27(1985), 75-96.

[**8**] ———: "Conflict and Coincidence of Interest in Job Matching: Some New Results and Open Questions," *Mathematics of Operations Research*, 10(1985), 379-389.

[**9**] SHAPLEY, LLOYD S., AND MARTIN SHUBIK: "The Assignment Game I: The Core," *International Journal of Game Theory*, 1(1972), 111-130.

[**10**] SUNDARSHAN, ALEX, AND SIDNEY ZISOOK: "National Resident Matching Program," *New England Journal of Medicine*, 305(1981), 525-526.

[11] A version of the following property, called the "consensus property" in Roth [**8**], was noted for the marriage problem by Knuth [**3**], who attributed it to John Conway, and was also observed in the sidepayment model of Shapley and Shubik [**9**].