New Physicians: A Natural Experiment in Market Organization

ALVIN E. ROTH

The National Resident Matching Program is a centralized clearinghouse through which new medical graduates in the United States obtain their first positions. The history of this market, from the market failures that the centralized system was designed to address, to the present, is discussed, and a hypothesis about the behavior of such markets is presented. New evidence is then presented from a set of similar centralized markets in the United Kingdom. Because some of these latter markets have failed, while others have succeeded, they provide a natural experiment that permits the hypothesis to be tested. The new evidence also suggests directions in which modifications of existing procedures might be considered.

MEDICAL SCHOOL GRADUATES IN THE UNITED STATES seek their first employment through a centralized labor market called the National Resident Matching Program (NRMP). This centralized market was established in the 1950s in response to persistent failures to organize the market in a timely and orderly way by decentralized means. Similar centralized labor markets, inspired by similar market failures, have been used in some regions of the National Health Service in the United Kingdom, where new medical graduates seek pre-registration positions comparable to U.S. residencies. Because different regions have used different procedures for organizing the market, the British system presents a natural experiment that allows these procedures to be compared with each other and with the U.S. market (*1*). Because some of the centralized procedures used in Britain have failed and been abandoned, whereas others have succeeded, this natural experiment also presents an opportunity to test the hypothesis put forward to explain the success and longevity of the U.S. market (*2*).

The centralized markets discussed here have the same outward form: each student submits a list of positions whose order is intended to reflect his preferences, and the person responsible for filling each set of positions likewise submits a ranked list of students. The markets differ in the algorithms then used to match students to positions—that is, they differ in what matching will be produced for a given set of submitted preference lists. The evidence suggests that these differences have a profound effect on the incentives that agents may have to try to circumvent the centralized market (and on what kinds of preference lists they submit). The differences between the kinds of matchings produced by centralized markets that have succeeded and those that have failed suggest constraints that must be faced in the design of algorithms to perform this kind of market function. (In this respect, this work is on the relatively unexplored interface of economics and operations research.) These differences also permit us to draw inferences about the nature of the matchings achieved in decentralized markets that experience similar success or failure.

The body of theory to which this evidence applies most directly concerns "two-sided matching markets," in which agents of one kind (for instance, workers) are matched with agents of another kind (employers) (*3*, *4*). Another two-sided matching market briefly discussed is the process by which sororities recruit members on U.S. college campuses (*5*).

The U.S. Market

From around 1900 to 1945, competition among hospitals for new interns and residents forced the date of appointment ever earlier, until medical students and hospitals were concluding agreements for post-graduation employment up to 2 years before graduation. This was costly in a variety of ways for both students and hospitals, and the date of appointment was finally brought under control in 1945 through intervention by the medical schools. There followed a period in which the market was very disorderly, with students being called upon to make increasingly prompt decisions whether to accept offers. (By 1949, a grace period of 12 hours had been rejected as too long.) The market was characterized by chaotic, last-minute recontracting, with students seeking to improve on the positions they had been firmly offered (and had sometimes accepted) by contacting the hospitals they preferred, and with hospitals sometimes pressuring students into premature decisions in order to be able to contact promptly students on their waiting lists. A centralized clearinghouse was proposed only when other attempts to organize the market had been exhausted.

This centralized market is still in operation (*6*). From 1952, following the introduction of the centralized matching procedure, there was a high degree of voluntary, orderly participation, with about 95% of U.S. medical school graduates entering the match and ultimately being offered and accepting the position with which they were matched. However, beginning in the mid-1970s, with a growing number of married couples in need of two positions in the same vicinity, high percentages of these obtained them outside of the centralized match; thus, the overall rate of participation, while still high, began to drop.

Previously, I proposed a hypothesis (*2*) to account for the transition from chaotic recontracting to orderly voluntary participation that took place in 1952, and the transition from uniformly high rates of participation among medical school graduates prior to the 1970s, to the defection of married couples in the late 1970s and

The author is the A. W. Mellon Professor of Economics at the University of Pittsburgh, Pittsburgh, PA 15260.

early 1980s. To describe the hypothesis, it is first necessary to briefly describe a model phrased in terms of the preferences of each student for positions and of each hospital program for students (*7*).

Because a hospital program typically employs more than one student, a description of a hospital's preferences must include how it evaluates alternative groups of students. For the purposes of this article, it is sufficient to suppose merely that if two groups of students differ only by a single individual, the hospital prefers the group containing the higher ranked individual. A student will be called "unacceptable" to a hospital if the hospital would prefer to keep a position vacant rather than fill it with that student, and a hospital is unacceptable to a student if, rather than accept one of its positions, the student would prefer to remain unmatched (and seek employment in a secondary market).

An outcome of the market is a matching of students and hospitals, such that no hospital is assigned more students than it has positions, and no student is assigned more than one position. A matching is called "unstable" if some student is matched to an unacceptable hospital, some hospital is matched to an unacceptable student, or if some student and hospital who are not matched to one another would both prefer to be matched together (*8*). Notice that an unstable matching gives some agent or pair of agents an incentive to find a different match. A matching that is not unstable in this way is called "stable" (or pairwise stable) (*9*). At a stable matching, each student finds that any hospitals he prefers to the one he is matched with do not return the favor.

The hypothesis proposed previously (*2*) is based on the demonstration that the 1952 matching algorithm produced a stable matching (in terms of any preferences that were submitted), and that the procedure used to assign married couples two jobs in the same vicinity was particularly prone to produce unstable matchings (*10*). Thus, the "stability hypothesis" applied to this market is that the chaotic conditions prior to 1952 reflected the instabilities then present in the market, that the success of the centralized procedure was due to the stability of the matching it produced, and that the decline in participation among married couples in the 1970s was because they once again found instabilities. Note that a student who has been offered or had proposed to him a specific job (or a couple who was matched with a pair of jobs) has only to make a few phone calls to determine if any preferred hospitals would be willing to offer a position, so the problem of determining if there are any exploitable instabilities is not a difficult one.

Of course, even though the stability hypothesis seems to account for the major developments in this market, the real explanations might lie elsewhere; for example, maybe any centralized market organization would have solved the problems experienced prior to 1951, and perhaps the experience of married couples has less to do with instabilities than with the difficulties young couples have in making decisions. Another kind of question about the stability hypothesis concerns the incentives that agents may have to submit rank-orderings that differ from their true preferences. The algorithm used in the U.S. market has the property that opportunities may arise both for students and hospitals to obtain a more preferred matching by submitting a preference list different from their true preferences rather than by submitting their true preferences (*3, 11*). If agents may have reason not to submit their true preferences, the fact that the algorithm produces a matching stable with respect to the submitted preferences does not ensure that the matching is stable with respect to the true preferences (that is, the preferences according to which agents search for and accept alternative opportunities).

One approach to addressing this question is to consider whether the agents in the market have the kind of information about one another's preferences needed to profitably submit rank orderings different from their true preferences. If not, submitted preferences might approximate true preferences sufficiently to produce stable outcomes. And if agents were sufficiently well informed so that they could submit the preference lists that would achieve their best available matches, given the lists submitted by everyone else, then under some circumstances it is possible to demonstrate that the resulting matching would be stable with respect to the true preferences (even though these would differ from the submitted preferences) (*3*). My point in raising these matters here is to indicate that, despite the (mathematical) demonstration of the stability properties of the U.S. system algorithm, and the degree to which stability and its absence seem to explain the major changes in the history of the market, there is room to doubt the stability hypothesis as an explanation of the behavior of the U.S. market. The question remains largely an empirical one, which gives further reason to make additional observations of the kind considered next.

The British Markets

A medical school graduate in Britain is eligible only for provisional registration with the General Medical Council. For full registration a doctor must complete separate medical and surgical pre-registration positions. An outcome of this market is thus a matching of students and consultants (supervising physicians and surgeons) such that no consultant is assigned more students than he has positions, and no student is assigned more than two positions, one medical and one surgical.

Most positions are filled on a regional basis, with graduates of a medical school going to a hospital in the same region (*12*). These regional markets are two orders of magnitude smaller than the U.S. market, ranging from approximately 100 to 300 positions.

That students seek two positions, rather than one as in the United States, makes for some important differences between the U.K. markets and the U.S. market. Both sides of the market must have preferences not just over individuals but over sets; that is, consultants have preferences over groups of students, and students have preferences over pairs of jobs. Thus, agents' preferences cannot be modeled as simple rank orderings, even though, as in the United States, students and consultants are asked to submit rank orderings of one another in these markets. (For example, because of the registration requirements, students prefer one medical and one surgical position to any other combination, regardless of their preferences for individual positions.) Therefore, the mathematical models used to develop the theory of these markets make weaker assumptions about preferences than were appropriate for the U.S. market. The main assumption is that consultants regard students and students regard positions more as substitutes than as complements, in that a consultant who is willing to hire a given student as part of some group of students should remain willing to hire him even if some other member of that group becomes unavailable (*13*). Under this assumption it can be shown (*1*) that the set of stable matchings is nonempty for any preferences, where, as before, a matching is stable if no matches are unacceptable and if no student or consultant who are not matched to one another would both prefer to be matched together (*14*).

The centralized pre-registration markets in Britain arose in reaction to problems that emerged in the 1960s. The markets in the various regions of the National Health Service were previously run in a decentralized way, with students responsible for finding positions on their own, and consultants responsible for filling the positions under their supervision. Competition among students for desirable positions and among consultants for desirable house officers eventually led to these positions being filled earlier and

earlier in the students' education, just as had happened in the United States. When a Royal Commission on Medical Education (1965–1968) investigated the problems confronting the pre-registration market, the organization of the U.S. market presented an obvious alternative. Following the Royal Commission's recommendations, many regions introduced centralized matching procedures, but different regions used different algorithms to determine the match from the submitted preferences. (Some, but by no means all, of these centralized schemes were implemented by computer.) It appears that over a dozen regional matching procedures were introduced, but, in contrast to the U.S. experience, only a few survived to the present. Most failed to solve the problems that motivated their introduction and were abandoned.

There are eight matching algorithms for which I have been able to obtain sufficiently precise descriptions to determine whether they produce stable matchings. [All but one (*15*) was computerized.] Two of these always produce stable matchings (*1*), and both of these have controlled the unraveling of appointment dates and survived to the present. The six remaining schemes are based on algorithms that may frequently produce instabilities. Only two of these have survived (and these are in the two smallest markets); the other four have been abandoned. Table 1 summarizes these results and also includes the U.S. market and the "preferential bidding system" used by U.S. sororities (*5*), which are the other two centralized matching procedures whose rules I have so far been able to learn in sufficient detail to determine whether they produce stable matchings (*16*). Of ten matching schemes so far observed, four produce stable matchings, and all four of these are still in use. Six produced unstable matchings, and four of these are no longer in use (*17*).

Because ten is a relatively small sample, and because Table 1 groups together markets that are vastly different in size, and procedures that differ in more ways than the stability of their outcomes, it may be illuminating to briefly describe one of the failed procedures and how it failed.

The scheme introduced in Newcastle in 1967 (*18*) used the product of the student's ranking of the consultant and the consultant's ranking of the student as the basis for a "priority" for that student to be employed by that consultant. The first step of the algorithm was to make all the first priority matches, after which consultants with unfilled positions and students still needing jobs were scanned to identify any second priority matches, and so on. If a consultant and student each ranked one another first [a "(1, 1) match"], they had a priority of 1. If the consultant ranked the student first but the student ranked the consultant second [a "(1, 2) match"], they had a priority of 2, as did a consultant who ranked a student second but was ranked first by the student [a "(2, 1) match"]. Ties were broken in the student's favor (*19*).

For example, consider six consultants, each of whom has only one position to fill, and six students, each of whom needs only one position. (The example does not depend on this simplification.) The rank orderings of the agents are as follows:

C_1: s_1, . . .	s_1: C_1, . . .
C_2: s_1, s_3, s_2, s_4, s_5, s_6	s_2: C_2, C_1, C_3, C_4, C_5, C_6
C_3: s_3, s_4, . . .	s_3: C_4, C_3, . . .
C_4: s_4, s_3, . . .	s_4: C_3, C_4, . . .
C_5: s_1, s_2, s_5, s_3, s_4, s_6	s_5: C_1, C_2, C_5, C_3, C_4, C_6
C_6: s_2, s_5, . . .	s_6: C_5, C_2, . . .

The Newcastle algorithm makes the matches: C_1s_1 (1, 1); C_3s_4 and C_4s_3 (2, 1); C_2s_2 (3, 1); C_5s_6 (6, 1); C_6s_5 (2,6). This outcome is unstable because C_5 and s_5 are one another's third choices, but are each matched to their sixth choice.

So far, the example has been analyzed as if the agents all state their true preferences. Before examining the incentives that agents may have to do otherwise, consider the way in which this matching procedure failed and was abandoned. The following description, for example, is from John Anderson, the Postgraduate Dean at Newcastle (*20*): "Shortly before the scheme was discarded we found that in up to 80% of cases students and consultants only used the computer to indicate a first preference. . . . The main reason for the abandonment of the scheme, therefore, was that there were problems in getting students and consultants to participate in an orderly way, and this led to those who rigidly observed the requirements of the scheme to be penalised."

Table 1. Stable and unstable algorithms.

Market	Stable	Still in use
United States	Yes	Yes
Edinburgh (1969)	Yes	Yes
Cardiff	Yes	Yes
Sororities	Yes	Yes
Cambridge	No	Yes
London Hospital	No	Yes
Birmingham	No	No
Edinburgh (1967)	No	No
Newcastle	No	No
Sheffield	No	No

To understand this phenomenon, consider now the incentives which this procedure gives to the agents. To make the above example clear, suppose consultants C_1 through C_5 are in the most desirable teaching hospital, and C_6 is in a much less desirable hospital. Similarly, suppose students s_1 through s_5 are all top graduates, while s_6 has a much less distinguished record.

Then in the example, C_5 is disappointed to learn that his new junior house officer will be s_6, all the more when he learns that s_5, whom he liked better, is unhappy with his own appointment and would have preferred to work for C_5. In a market in which positions could be formally filled by private arrangement (as in the U.S. market), this situation might have led C_5 to decline to hire s_6 and to offer his position instead to s_5. If, in addition, the market were as large and impersonal as the U.S. market, then s_5 might be tempted to accept the offer even if he had already made some sort of preliminary commitment to C_6. This would generally not have been an option in Great Britain, where authority was (and is) more centralized, so that consultants and students would have to abide by the outcome of the matching algorithm once it was completed (*21*).

But suppose C_5 resolves not to suffer the same fate the next year. He, therefore, approaches one of the good students in the next year's class, before the formal match, and suggests that they agree to be matched, which they will accomplish by ranking one another first in the formal match (*22*). The student, mindful of the experience of s_5 the previous year, is receptive. Now consider the situation in the formal match, when a number of positions have been prearranged to be (1, 1) matches. Suppose students t_1, t_2, and t_3 have made such arrangements with consultants C_3, C_4, and C_5, but consultant C_2, not knowing this, submits his true rank ordering, t_1, t_2, t_3, t_4, t_5, . . ., and t_4 submits his true rank ordering C_3, C_4, C_5, C_2, ... Although C_2 does not know it, t_4 is his highest ranking student who is actually available, and C_2 is t_4's most preferred available consultant. But since the product of their rankings is 16, C_2 could well end up with his 15th choice student.

Thus, when some matches have been arranged beforehand, other students and consultants stand to do poorly if they do not also prearrange their matches. And matches can be prearranged, even when the rules of employment require students and consultants to participate in the centralized procedure, because a student and consultant who rank one another first will be matched regardless of what anyone else does. This goes a long way toward explaining both

why a high percentage of appointments were soon arranged in advance under this system, and why this worked to the disadvantage of those who tried to arrange employment through the formal match procedure.

The Newcastle market is representative in this respect of the four unstable procedures in Table 1 that are no longer in use. Because it produced unstable matchings, some students and consultants had mutual interests in circumventing the formal procedure. And the market environment and matching procedure allowed them to communicate these interests to one another and effectively prearrange matches privately. Furthermore, as more agents went outside the system, the greater was the incentive for other agents to do so, so that the market quickly began to unravel.

If the sample consisted only of the four stable procedures that are still in use and the four unstable procedures that are no longer in use, it would have provided very strong support for the simplest form of stability hypothesis, along the lines of "stable procedures work and unstable ones do not." The two unstable matching procedures that are still in use (Cambridge and London Hospital) are therefore worth special notice, because they suggest that the situation may be more complex. These are the two smallest of the U.K. markets in the sample, and each involve the graduates of a single medical school and the teaching hospital associated with that school. The participants in these two markets may be effectively compelled by social pressure of various sorts (*1*) to comply with the match procedures, and this may be a sufficient explanation for their continued use (*23*). But it is worth noting that of the ten stable and unstable procedures in this sample, these two are the only procedures in which a student and consultant who each rank one another as first choice will not necessarily be matched together. This makes it more difficult for a student and consultant to prearrange a match, and together with the social pressures that can be brought to bear in a small market in which participants know each other, this may be a factor in preventing these unstable procedures from unraveling.

Concluding Remarks

There is now considerable evidence that stable centralized procedures allow markets to be organized more successfully than do unstable procedures. Unstable matchings give some market participants incentives to circumvent the formal procedures. Open questions remain concerning the extent to which there may be unstable procedures that in some environments nevertheless prevent the interested parties from acting on these incentives and whether there may be environments in which even stable procedures are prone to unravel.

Note that the centralized markets studied here do not involve central planning as it is most usually understood, because these markets have been designed to be sensitive to the preferences expressed by the participants, rather than to achieve the independent objectives of a planner. What is centralized is not the objective, but the market mechanism itself. This kind of centralization is something that occurs more often than is generally recognized: for example, the stock market contains institutions like the New York Stock Exchange, governed by explicit rules about how and when trades may be transacted. How such centralized market institutions arise, and the functions they serve, are subjects that deserve more attention than they have so far received in the economics literature, which often treats markets as arising spontaneously.

Another natural question is, what practical lessons about the design of centralized matching markets can we draw from this evidence? It seems clear that a designer of a centralized procedure to replace a decentralized market that has failed for reasons symptomatic of instability would be well-advised to choose a stable procedure. In fact, different stable procedures have different strategic properties (*3*), so more specific advice may be possible. But what, if anything, can we say about modifying existing centralized procedures? This is probably the question that I am asked second most often by participants in the NRMP. (The most often asked question concerns how to approach the existing system strategically.)

Let me state right away that I think that modifications of procedures that are working fairly well should be contemplated with caution, because a procedure like the NRMP is a vast improvement over the chaotic conditions that preceded it, and because we have seen that unstable centralized procedures are themselves prone to unravel. So the first priority should be preserving the stability of the system. [This puts serious constraints on certain kinds of "fine-tuning" that might otherwise seem desirable, for example, to attempt to change the distribution of residents to rural hospitals (*3*).] And modifications that increase stability may become important as instabilities develop (as in the case of married couples).

But beyond this (and here I venture cautiously), one of the features of these markets that is burdensome to many participants is the fact that there are strategic decisions to make in deciding what preference list to submit. Although there are aspects of this that are inevitable consequences of stability (*3*), one issue that arises is that any stable procedure makes it important for hospital programs to be highly ranked by the students they desire (and vice versa). There is thus an incentive for program directors to engage in various activities designed to influence the rank orderings submitted by students. One possibility, regarded as unethical, is for a director to ask a student whether he will rank the program first, and to say that if the student agrees to do so the director will in turn rank the student in his first group of choices (*24*). This is a problem for two reasons. First, although it is arguably justifiable for students to respond to such a breach of conduct by coolly asserting that they will rank the program first (and then proceeding to rank programs as they wish), many students will be understandably reluctant to begin their professional careers with an act of duplicity. And the misinformation introduced in the market when students do change their preference lists causes avoidable distortions. A possible solution is for hospitals' preference lists to be made available to students before they must submit their own. This practice has developed for other reasons at Cardiff, where it does not interfere with the operation of the market (*25*).

In closing, modern economic theory is increasingly game-theoretic in nature, which is to say that it focuses on the rules by which markets are organized. Centralized markets present a natural place to begin an empirical investigation concerned with rules, because important parts of the rules are formally codified, and can therefore be observed with precision. But a natural direction in which to extend this research is to decentralized entry-level labor markets, some of which exhibit the kinds of failures that can now be interpreted as symptoms of instabilities, and others that do not.

It is noteworthy that the simple idea of pairwise stability formulated by Gale and Shapley (*4*) has turned out to have so much empirical power. It is the kind of theoretical work that merits high scientific recognition.

REFERENCES AND NOTES

1. For a more detailed presentation, see A. E. Roth, *Am. Econ. Rev.*, in press.
2. A. E. Roth, *J. Polit. Econ.* **92**, 6 (1984).
3. The modern theory of these markets is described by A. E. Roth and M. A. O. Sotomayor [*Two-Sided Matching: A Study in Game-Theoretic Modelling and Analysis* (Econometric Society Monograph Series, Cambridge University Press, Cambridge, 1990)]. Much of the theory traces its origins to the work of Gale and Shapley.
4. D. Gale and L. Shapley, *Am. Math. Mon.* **69**, 1 (1962).

5. S. Mongell and A. E. Roth, *Am. Econ. Rev* , in press.
6. Roughly 20,000 positions are now offered annually in the U.S. market, where students each seek one position and hospital programs seek multiple students. After students are interviewed by hospitals, both students and hospitals submit preference lists. (Salaries are part of the job description and are simply one of the factors that determine students' preferences over hospitals.)
7. Because preferences cannot be directly observed, it is desirable to make the weakest assumptions possible about them. (In these markets the preference lists that agents submit are observable, but the matching algorithm may give agents incentives to submit lists that do not faithfully reflect their preferences.) The assumptions made here, although not terribly strong, nevertheless assume that hospitals have preferences over groups of students that bear a certain relation to preferences over individual students. These assumptions can be considerably weakened, and in some cases this is necessary to accommodate choices by students or hospitals that do not reflect an underlying preference over individual students or positions at all (*1, 3*). Throughout, preferences are assumed to be complete and transitive orderings.
8. Specifically, the hospital must prefer the student to one of the students it is matched with, or, if it has some unfilled positions, it must prefer the student to leaving a position unfilled.
9. The notion of (pairwise) stability was formulated by Gale and Shapley (*4*). In their model of one-to-one matching, pairwise stable matchings have the further property that no larger group of agents can collectively improve their situation, and in that model the set of stable matchings coincides with the set called the "core," which is a standard game-theoretic concept.
10. Prior to the mid-1980s, couples participating in the match were required to specify one of their members as the "leading member," and to submit a rank ordering of positions for each member of the couple; that is, a couple submitted two preference lists. The leading member was then matched to a position in the usual way, the preference list of the other member of the couple was edited to remove distant positions, and the second member was then matched if possible to a position in the same vicinity as the leading member. It is easy to see why instabilities would often result. Consider a couple whose first choice is to have two particular jobs in Boston, and whose second choice is to have two particular jobs in New York. Under the couples algorithm, the leading member might be matched to his or her first choice job in Boston, whereas the other member might be matched to some undesirable job in Boston. If their preferred New York jobs ranked this couple higher than students matched to those jobs, an instability would now exist. Recent changes in the way married couples are handled may somewhat reduce this problem, but when there are married couples in the market, the set of stable matchings may be empty (*2, 3*).
11. In 1951, an algorithm for the U.S. market was proposed that gave students clear incentives to state rank orderings different from their true preferences. It was replaced for this reason by the 1952 algorithm, which was claimed in the literature distributed to participants never to give such incentives to either students or hospitals. In fact, this property is incompatible with stability (*3*).
12. But in London, where there are many more graduates of local medical schools than local pre-registration positions, medical schools commonly have arrangements with hospitals elsewhere.
13. Even this assumption is sometimes too strong to fit the facts and must be relaxed. For example, at some point in the history of the market operated in Cardiff, students could hold no more than one of their positions at the teaching hospital. So the first choice of some student might be to have his medical position at the teaching hospital and his surgical position elsewhere, but if that medical position were unavailable his second choice might be to have his surgical position at the teaching hospital. That is, he would become unwilling to accept the surgical position that was part of his first choice package, even though it was still available, because the medical position which went with it was no longer available. The role of this kind of assumption was first explored by A. Kelso and V. Crawford [*Econometrica* **50**, 6 (1982)].
14. One important difference between the U.K. and American markets is that in the U.K. markets, even when a matching is (pairwise) stable, so that no student and consultant can together arrange to do better than a given matching, there might be a larger coalition, consisting of many consultants and students, who by rearranging job assignments could obtain preferred assignments for all members [C. Blair, *Math. Oper. Res.* **13**, 619 (1988); (*3*)]. Needless to say, identifying and organizing large coalitions may be more difficult than making private arrangements between two parties, and it will become clear that the set of stable matchings is still of primary concern.
15. Included here is a matching scheme that was tried and subsequently abandoned at Sheffield; it is not included among those formally analyzed in (*1*) because the match was done by a committee, whose exact procedures cannot therefore be determined. But A. D. Clayden and J. Parkhouse [*Brit J Med. Ed.* **5**, 5 (1971)] report a computer program designed, in their words, "to mimic the manual allocation," and for my purposes here, I take that to be the system used.
16. The two Edinburgh procedures were adopted sequentially: the unstable procedures were adopted in 1967 and replaced by the stable procedures in 1969. The preferential bidding system used by sororities does not produce a matching for some possible inputs, but on the four campuses studied (*5*), in which the number of interested new members does not exceed the available positions, students have an incentive to submit preference lists containing only their first choice sorority. When students do so (and high percentages of the preference lists in that data do contain only a single choice), the resulting matching is stable.
17. More precisely, the four procedures I have called stable can all be shown to have the property that the matchings they produce are stable with respect to whatever preference lists may be used as the input. The six procedures I have called unstable do not have this property; although they may sometimes produce stable matchings, there are preferences for which they will not.
18. A. G. Leishman and R. P. Ryan, *Lancet* **ii**, 459 (1970).
19. At least initially—a later modification was to reverse this method of tie-breaking (D. A. Shaw, personal communication).
20. J. Anderson, personal communication.
21. But notice that if C_5 had submitted a preference list on which s_5 was his first choice, they would have been matched, as would also have been the case if s_5 had submitted a preference list on which C_5 was his first choice. So there are incentives for both students and consultants to carefully consider what preference lists to submit, because submitting their true preferences does not always yield the most preferred outcome.
22. In these relatively small markets both parties to such an agreement can be confident it will be carried out, because a consultant with a reputation for not delivering on his promises will soon find it difficult to attract good junior house officers, and a junior physician is reluctant to incur the enmity of a senior physician in the region in which he hopes to practice. The situation may differ in larger, more impersonal markets.
23. Stability as formulated here plays little role in procedures in which the participants can be compelled to accept the resulting match. A familiar example is the procedure by which U.S. football teams draft college players. That procedure is not stable in the sense discussed here, but does not need to be, because players must play for the team that drafts them.
24. There is anecdotal evidence that directors do this and, furthermore, that they do not always live up to their part of the bargain. Although I have not tried to gather systematic evidence on this point, the incentives for just this kind of behavior are large.
25. One way to exercise caution in modifying an existing procedure is to have evidence from another market that the modification does not introduce new problems. In this respect, the emerging tools of experimental economics may be of use in allowing some kinds of modifications to be explored on a small scale under laboratory conditions [A. E. Roth, Ed., *Laboratory Experimentation in Economics. Six Points of View* (Cambridge University Press, Cambridge, 1987)].
26. Supported by the Alfred P. Sloan Foundation and NSF. I am also indebted to numerous British physicians, surgeons, and medical administrators who have taken the time to correspond with me on these matters, and sometimes to unearth old records. I would be remiss not to mention J. Anderson, T. J. Bayley, P G. Bevan, K. C. Calman, S. C. Farrow, J. Fraser, F. J. Goodwin, T. M. Hayes, K. Johns, J H. Lazarus, D. McInnes, G. A. Mogey, R. Mulligan, K. M. Parry, R. P. Ryan, D. A. Shaw, D. M. Taylor, H. R. A. Townsend, and N. D. Wright. And I have received helpful comments from S. Mongell, J. Ochs, and J. Prasnikar.