

A Natural Experiment in the Organization of Entry-Level Labor Markets: Regional Markets for New Physicians and Surgeons in the United Kingdom

Author(s): Alvin E. Roth

Source: *The American Economic Review*, Vol. 81, No. 3 (Jun., 1991), pp. 415-440

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/2006511>

Accessed: 25-01-2019 15:45 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*American Economic Association* is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

# A Natural Experiment in the Organization of Entry-Level Labor Markets: Regional Markets for New Physicians and Surgeons in the United Kingdom

By ALVIN E. ROTH\*

*The histories of seven regional markets for new physicians and surgeons in the United Kingdom are considered. Like the American market, these markets have experienced failures that led to the adoption of centralized market mechanisms. Because different regions employ different centralized mechanisms, these markets provide a test of the hypothesis that the success of the American market is related to the fact that it produces matches which are stable in the sense that no two agents mutually prefer to be matched to one another than to their assigned partners. Even in the more complex U.K. markets, this kind of stability plays an important role. Centralized markets that produced unstable matches in environments in which agents could act upon instabilities fared no better than the decentralized markets they replaced. (JEL C78, D00, J41, J44)*

In this paper, I seek to analyze a unique natural experiment that has emerged over the last 20 years in the United Kingdom, concerning the organization of entry-level labor markets. The markets in question are those in which newly graduated medical students seek their first hospital positions, called preregistration house-officer positions. These positions are closely comparable to first-year intern positions in American hospitals (although there are some important differences, which will be

discussed). The natural experiment arises because these markets are organized differently in different regions of the National Health Service. These different markets allow investigation of how market behavior is influenced by market organization. This diverse set of markets also invites comparison with the American market for interns and allows the hypothesis advanced in Roth (1984a) about the behavior of that market to be tested and refined.<sup>1</sup>

The particular forms of market organization in the United Kingdom that are the subject of this paper arose in reaction to problems that emerged in the 1960's. Prior to the mid-1960's, the preregistration house-officer markets in the various regions of the National Health Service were largely run in a decentralized way, with students responsible for finding positions on their own and with consultants (as senior physicians and surgeons are called) responsible for filling the positions under their supervision. Competition among students for desirable positions and among consultants for desirable house officers eventually led to

\*Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260. In the course of gathering information about the markets described here, I have been helped by numerous British physicians and medical administrators, who have taken the time to correspond with me at length on these matters and sometimes to unearth old records. Among those who have taken pains to help me, I would be remiss not to mention Drs. 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. Moge, R. Mulligan, K. M. Parry, R. P. Ryan, D. A. Shaw, D. M. Taylor, H. R. A. Townsend, and N. D. Wright. Of course, they are in no way responsible for any errors or omissions in my account of these markets, nor for the conclusions reached in this paper. I am also grateful to Dr. Susan Mongell, who helped assemble the published literature on this topic. This work has been supported by the National Science Foundation and the Alfred P. Sloan Foundation.

<sup>1</sup>Some of my conclusions about this comparison were earlier described briefly in Roth (1990).

these positions being filled earlier and earlier in the students' education. The situation is described as follows by A. Doig and G. Munday (1969):

The filling of preregistration house-officer posts by private arrangements between consultants and students is considered unsatisfactory by most students and by an increasing number of consultants. This practice allows consultants to offer appointments to promising students early in their clinical years and often causes subsequent regret on the part of those who accept when they are later offered more desirable posts; broken agreements cause considerable inconvenience and irritation to consultants. [p. 1250]

J. Alexander-Williams and Ivor Stephenson (1973), writing about the situation in Birmingham prior to the mid-1960's, describe things similarly.

Our graduates who wished to take posts in the teaching hospital insisted on promises earlier and earlier in their undergraduate careers as a result of the many more attractive posts in the region being promised years in advance to graduates imported from other regions. [p. 605]

Doig and Munday further note that

...an initial attempt was made to effect improvement by simply standardising the dates for the receipt of applications and the offering of posts; this procedure had the unfortunate effect of telescoping many of the difficulties and frustrations which had previously occurred over a year or more into a much shorter period. [1969 p. 1250]

This early history is strikingly similar to that of the American market for interns from 1900 to 1951, described in Roth (1984a). By 1944, the standard date at which internship appointments were made in the American market had advanced to two full years before the internship was to begin. In 1945, standardized appointment dates nearer the beginning of employment were

successfully established through the intervention of the medical schools. However, problems manifested themselves in the waiting period between the time offers of internships were made and the time students were required to accept them. Roth (1984a) describes these as follows:

...a student who was offered an internship at, say, his third choice hospital, and who was informed he was an alternate (i.e., on a waiting list) at his second choice, would be inclined to wait as long as possible before accepting the position he had been offered, in the hope of eventually being offered a preferable position. Students who were pressured into accepting offers before their alternate status was resolved were unhappy if they were ultimately offered a preferable position, and hospitals whose candidates waited until the last minute to reject them were unhappy if their preferred alternate candidates had in the meantime already accepted positions. Hospitals were unhappier still when a candidate who had indicated acceptance subsequently failed to fulfill his commitment after receiving a preferable offer. In response to pressure originating chiefly from the hospitals, a series of small procedural adjustments were made in the years 1945–51. [p. 994]

These adjustments primarily involved shortening the time during which students were allowed to hold offers without either accepting or rejecting them. By 1949, a time of only 12 hours had been rejected as too long, and by 1951 it was widely recognized both that there were serious problems in the last stage of the matching process and that these could not be resolved by compressing this last stage into shorter and shorter time periods.

At this juncture, a centralized (and computerized<sup>2</sup>) matching procedure was proposed and adopted in the American market.

<sup>2</sup>Initially the procedure was carried out with card-sorting machines.

Students and hospital programs continued to interview one another as before, but then students submitted to a central clearinghouse a list of hospital programs ranked in order of preference, and hospital programs submitted a list of students. On the appointed day, students were notified of the hospital program to which they had been matched, hospital programs were notified of the students with which they had been matched, and both hospitals and students were encouraged to sign the necessary contracts with one another. Note that at every stage of the procedure, from the decision to submit preference lists to the decision whether to accept the match ultimately proposed, this matching procedure was instituted as a *voluntary* procedure. It is therefore particularly notable that it achieved very high rates of voluntary participation, with the proposed matchings accounting for the large preponderance of jobs filled in this market for many years following its inception. This matching procedure has survived for many years; the algorithm used today to accomplish the match is still essentially the one first used in 1952.

When a Royal Commission on Medical Education (1965–68) issued a report on the problems confronting the preregistration market, the organization of the American market presented an obvious alternative. Many medical schools and their affiliated regional hospitals 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 more than a dozen regional matching procedures were introduced, but in contrast to the American experience, only a few have survived to the present. Most failed to solve the problems that motivated their introduction and were abandoned.

The principal goal of the present paper is to investigate what common properties might distinguish those centralized procedures that failed and were abandoned from those that succeeded and are still in use today. This will also present an opportunity

to test and refine, in a different environment, the hypothesis advanced in Roth (1984a) about the success of the centralized procedure introduced in the American market.

It was shown in Roth (1984a) that the algorithm employed in the American market produces a *stable* matching in terms of the submitted preferences, where a matching is called stable if no student or hospital is matched to an unacceptable mate and if there is no student and hospital not matched to one another who would both prefer to be matched to each other than to (one of) their partner(s) in the matching. The next section recounts briefly how the major developments in the American market can be explained in terms of stability or the lack of it, which leads naturally to the hypothesis that the success of the American procedure is intimately related to the stability of the matching it produces. Therefore, one goal of the present study is to provide a test of this “stability hypothesis” as it applies to the U.K. markets. Another goal is to consider the incentives agents may have to submit other than their true preferences and the effects this may have.

The regional markets in the United Kingdom provide a stringent test of the stability hypothesis for two reasons. First, administrative authority is more centralized in those markets than in the American market. In some of the regions in which a centralized matching procedure was tried in the United Kingdom, for the period in which it was in effect it was the only way to formalize employment agreements. In contrast to the American market, students in such regions could not decline a job with which they had been matched and arrange instead another job in the same region, and consultants could not decline a student with whom they had been matched in favor of another participant in the match.<sup>3</sup> However, as will be seen, it was sometimes possible for students and consultants to circumvent the match by

<sup>3</sup>However, students could sometimes arrange jobs in other regions.

coordinating with one another before the formal match. The second reason why instabilities might have less consequence in the U.K. markets than in the American market is that the U.K. markets are smaller than the American market by a factor of about 100. Consequently, there are many reasons why, even in regions in which other avenues to formalize employment agreements existed in principle, students and consultants might feel obliged to "play by the rules" and accept a suggested match, even if a better one might have been available. Such reasons might not apply in the much larger and more impersonal American market.

Even in the U.K. markets, the stability of matchings plays an important role, as will be seen below. Perhaps even more important than the formal stability conditions is whether agents can anticipate (or detect *ex post*) and act upon any instabilities that may occur. The centralized markets that were unstable in this way fared no better than the decentralized markets they replaced.

This paper concentrates on seven different procedures by which regional markets in the United Kingdom have been organized. These seven are all of those for which I have been able to obtain sufficiently precise descriptions of the matching algorithm (all seven were computerized). Of these seven, two are based upon modifications of an algorithm that always produces stable matchings, and both of these have controlled the unravelling of appointment dates and survived to the present. The five remaining schemes are based upon algorithms that may frequently produce instabilities. Only two of these have survived (and these are in the two smallest markets); the other three have been abandoned.

The organization of this paper is as follows. Section I discusses the American market. Section II describes the chief differences between the American market and the U.K. markets and describes a formal model suitable for the U.K. markets. The main result of this section is that, even at a stable matching, in these markets there may exist higher-order instabilities of a kind not present in the American market. Section III

begins the analysis of particular market organizations in the United Kingdom, starting with three defunct schemes, which all had particularly severe problems due to the fact that agents had incentives to submit other than their true preferences and to form pairwise coalitions prior to the formal match. Section IV discusses the two markets organized around stable matching procedures and the modifications that have been made in each, and Section V considers the two other surviving schemes and the particular markets in which they operate. Section VI considers some of the difficulties presented by the fact that there is no real national market in the United Kingdom for preregistration positions, but only a collection of regional markets, and conclusions are presented in Section VII. Some of the necessary formal apparatus and proofs will be deferred to the Appendix.<sup>4</sup> (See Roth and Marilda Sotomayor [1990] for a full account of the literature.)

## I. The American Market

In the American market, students each seek one position, while each hospital seeks some number of students.<sup>5</sup> The size of the market has approximately doubled since 1952: today roughly 20,000 positions are offered annually.

Interns' salaries are part of the job description and are not negotiated by each hospital and intern. Therefore, salaries will not play an explicit role in the model but will simply be one of the factors that determine the preferences that students have over hospitals. Similarly, hospitals can rank students. Because a hospital typically employs more than one student, a full description of a hospital's preferences must include how it

<sup>4</sup>Because of the way the paper is organized, it may be easiest to read the Appendix last, rather than referring to it whenever a proof is deferred to the Appendix.

<sup>5</sup>Strictly speaking, the agents on the institution side of the market are hospital programs rather than hospitals, since different internship programs within a hospital are separately administered and students apply to specific programs.

evaluates alternative groups of students. A simple assumption connecting a hospital's preferences for groups of students to its ranking of individual students is that, if two groups of students differ only by a single individual, the hospital prefers the group containing the higher-ranked individual. Preferences of this sort are said to be *responsive* (to the hospital's ranking of individual students). (More general assumptions are possible, as will be discussed in the next section.) A student is *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 to more than one position. A matching is *stable* if no student is matched to an unacceptable hospital, no hospital is matched to an unacceptable student, and no student and hospital who are not matched to one another would both prefer to be matched together. (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.)

In Roth (1984a), I undertook to explain three episodes in the history of this market. From 1945 through 1951, 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 students on their waiting lists promptly. This behavior persisted despite repeated attempts by various medical organizations to establish more orderly norms of behavior. However, from 1952, following the introduction of the centralized matching procedure, until the mid-1970's, there was a very high degree of voluntary orderly partic-

ipation, with in the neighborhood of 95 percent of American medical school graduates participating in the matching procedure and ultimately being offered and accepting the position they were matched with. However starting in the mid-1970's, as increasing numbers of married couples sought to obtain two positions in the same vicinity, the rate of participation in the match began to drop, with high percentages of married couples seeking and finding positions outside of the centralized match. So there are two transitions to explain: 1) the transition from chaotic recontracting to orderly voluntary participation that took place in 1952 and 2) the transition from uniformly high rates of participation among medical school graduates prior to the 1970's to the defection of married couples in the late 1970's and early 1980's.

The stability hypothesis is based on the demonstration in Roth (1984a) that the 1952 matching algorithm produces a stable matching (in terms of any preferences that may be submitted) and that the procedure used to assign married couples two jobs in the same vicinity was particularly prone to produce unstable matchings.<sup>6</sup> Note that a student who has been offered or has had proposed to him a specific job (or a couple which has been matched with a pair of jobs) has only to make a few phone calls to determine whether any of his preferred hospitals would be willing to offer him a position, so the problem of determining whether there are any exploitable instabilities is not a difficult one. 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

<sup>6</sup>In recent years, changes in the way married couples are handled by the match show some signs of ameliorating this problem. However, the problem is not completely tractable, since it was shown in Roth (1984a) that, when there are married couples in the market, the set of stable matchings may be empty. It should also be noted that other developments in the match during the 1970's contributed to the decline in participation during that period.

due to the stability of the matching it produced, and that the decline in participation among married couples in the 1970's was because they once again found instabilities.

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, it might be postulated that *any* centralized market organization would have solved the problems experienced prior to 1951 and that the experience of married couples has less to do with instabilities of the kind dealt with here than with the difficulties young couples have in making decisions. Therefore, it is particularly illuminating to be able to consider the kind of natural experiment, involving both stable and unstable centralized matching mechanisms, that one finds in the United Kingdom.

Before moving on to the U.K. markets, it will be useful to make two somewhat technical observations relevant both to those markets and to the American market. The first is that agents may have incentives to submit rank-orderings to the centralized matching procedure that differ from their true preferences.<sup>7</sup> It was shown in Roth (1982) that no stable matching mechanism exists that makes it a dominant strategy for all agents to state their true preferences, and it was shown in Roth (1985a) that no stable procedure can make it a dominant strategy for all hospital programs with more than one position to rank students in order of their true preferences. In the American market, stating true preferences is not a dominant strategy for either students or hospitals.

This raises another question about the stability hypothesis: if agents may have reason not to submit their true preferences to the centralized matchmaker, then the fact that the matching algorithm produces a matching that is stable with respect to the

submitted preferences does not assure that the matching is stable with respect to the true preferences (i.e., the preferences according to which agents search for and accept alternative opportunities).

One approach to addressing this question is to consider what will happen when the rank-orderings that agents submit are in equilibrium, even though they may be different from the true preferences. It was shown in Roth (1984b) that, when algorithms equivalent to the one used in the American market are employed, every Nash equilibrium in undominated strategies yields a matching that is stable with respect to the true preferences (as well as to the stated preferences) in the special case arising when all hospital programs have only a single position to fill. In the (actual) case, when hospitals fill multiple positions, fewer strategies are dominated, and so only weaker results have so far been obtained: there are equilibria at which the resulting matching is stable with respect to the true preferences, and other equilibria such that it is not (Roth, 1985a). Another approach is to ask whether the agents in the market have the kind of information about one another's preferences needed to profit from submitting rank-orderings different from their true preferences. (These information requirements are considerable; cf. Roth, 1989.) If not, submitted preferences might approximate true preferences sufficiently to produce stable outcomes. Viewed in this way, the question remains largely an empirical one, which gives further reason to make additional observations of the kind considered in this paper.

The second technical observation concerns the significance of defining stability in terms of individual agents and pairs of agents, without reference to larger coalitions of agents, such as coalitions consisting of a hospital and several students (all of whom might be employable by the same hospital) or coalitions consisting of multiple hospitals and students. When the rules of the market are that any student and hospital may sign a contract with one another if they both agree, the following result from Roth (1985b) says essentially that nothing is

<sup>7</sup>In 1951, an algorithm for the American market was proposed that gave agents clear incentives to state rank-orderings different from their true preferences. For this reason, it was replaced by the 1952 algorithm, which was claimed to give agents no such incentives.

lost by not considering such coalitions explicitly.

**PROPOSITION 1:** *The set of stable outcomes equals the core defined by weak domination<sup>8</sup> of the market in which hospitals employ multiple students but students take only one position.*

It will be seen that the conclusions of Proposition 1 do not carry over to the U.K. markets considered next, as a consequence of the fact that most students seek two positions in those markets.

## II. The U.K. Markets

A medical school graduate in the United Kingdom is eligible only for provisional registration with the General Medical Council. To become eligible for full registration, a doctor must complete 12 months in a pre-registration position, typically six months in a medical position and six months in a surgical position. These positions are supervised by different consultants and are arranged separately, so each graduating medical student must typically find two pre-registration positions. An outcome of this market is thus a matching of students and consultants 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 given medical school going to one of the associated teaching hospitals or to other hospitals in the

same region.<sup>9</sup> Despite their regional character, these markets have a centralized component absent from the American market; the Department of Health and Social Security sets targets for the number of preregistration house-officer posts for each English Regional Health Authority. These regional markets are two full orders of magnitude smaller than the American market: for 1988 the largest English region was the West Midlands, with just over 300 positions, and the smallest was East Anglia, with just over 100 positions.

The fact that students seek two positions, rather than one as in the American market, means that a different model must be used to represent the market. The simplest modification is to assume that students have separate rank-orderings over medical consultants and over surgical consultants. As in the model of the American market, consultants have rank-orderings over students, but now *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. Again, the simplest assumption is that these preferences are responsive to the rank-orderings of individuals, as defined in the previous section. As before, a matching is stable if no matches are unacceptable and if no student and consultant who are not matched to one another would both prefer to be matched together. However, even when the assumptions of the model are kept closely comparable to those for the American market in this way, the fact that matchings are many-to-two (i.e., consultants have more than one student, and students need two jobs) rather than many-to-one has the consequence that stable matchings need no longer be in the core of the market or even be Pareto efficient. That is, in this market there is no

<sup>8</sup>An outcome  $x$  of a market is *dominated* if there is some coalition  $S$  of agents that, by trading among themselves, can obtain allocations they all prefer to  $x$ . The outcome  $x$  is *weakly dominated* if such a coalition  $S$  can obtain allocations that all its members like at least as well as  $x$  and that at least one member strictly prefers to  $x$ . The core is the set of outcomes that are undominated, and the core defined by weak domination is the subset of the core consisting of outcomes that are not even weakly dominated. Any core outcome is Pareto optimal.

<sup>9</sup>However, in London, where there are many more graduates of local medical schools than local preregistration positions, medical schools commonly have arrangements with hospitals elsewhere.



parallel to Proposition 1. Instead, one has the following result, whose proof is in the Appendix:

**PROPOSITION 2:** *In the many-to-two matching model with responsive preferences, the set of stable matchings is nonempty for all preferences, but a stable matching need not be in the core and need not even be Pareto optimal.*

The proposition implies that, even when 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 its members. Needless to say, identifying and organizing large coalitions may be more difficult than making private arrangements between two parties, and it will become clear in what follows that the set of stable matchings is still of primary concern.

There are still further generalizations that must be made to the model to allow it to represent faithfully the variety of special constraints found in the various regional markets. In particular, the assumption that agents have responsive preferences must sometimes be relaxed to allow more complex connections between rank-orderings of individuals and preferences over groups.

On the student side of the market, these complexities arise because students require one of each of two types of jobs. However, perhaps the most unusual example of the sort of complicated preferences I have in mind arises on the consultant side of the market administered from Edinburgh. Following what I gather must have been traditional practice<sup>10</sup> before the introduction of a centralized matching scheme, surgeons have the option of indicating that they do not wish to employ more than one female preregistration house officer at a time. Therefore, a consultant surgeon could conceivably submit a rank-order list in which his top four choices, say, were female, but

nevertheless indicate that he wishes to employ at most one of these at any one time.

It is easy to see that these preferences are not responsive to the rank-ordering of individuals as discussed above, since the surgeon in question would prefer to employ his first and fifth choices rather than his first and second choices. Similarly, students clearly prefer one medical and one surgical position to any other combination, regardless of their preferences for individual positions. It will therefore sometimes be convenient to model agents' preferences for sets of alternatives directly, without explicitly considering their rank-orderings of individual students or positions.

Thus, faced with a set  $S$  of student applicants, a consultant  $C$  can determine which subset of  $S$  he would most prefer to hire. We call this  $C$ 's choice from  $S$ , and denote it by  $Ch_C(S)$ . That is, for any set  $S$  of students,  $C$ 's choice set is  $Ch_C(S) = S'$  such that  $S'$  is contained in  $S$  and  $S'$  is (at least weakly) preferred to any other subset  $S''$  of  $S$ . It will be convenient (but not essential) in what follows to assume that all agents have strict preferences, so that the choice set is unique. Thus, a consultant's preferences are given by a rank-ordering of sets of students,  $S_1, S_2, \dots, S_k, \Phi, \dots$ , with  $S_1$  being his first-choice set of house officers and so forth (and with any unacceptable set being less preferred than the empty set). These preferences determine the choice function (i.e., for any set  $S$ ,  $Ch_C(S)$  is  $C$ 's most preferred subset  $S'$  of  $S$ ). A student's preferences over (pairs of) positions can be represented analogously.<sup>11</sup> The constraints mentioned above are consistent with the preferences meeting the following condition:

*Definition:* An agent  $A$  (a consultant or a student) has *substitutable preferences* over sets of alternatives (i.e., sets of students or pairs of jobs) if, for any set  $T$  that contains

<sup>10</sup>Particularly among urological surgeons.

<sup>11</sup>In this representation of preferences, the requirement that students have one medical and one surgical position would be represented by having sets consisting of a pair of two medical jobs, for instance, be unacceptable to the student.

distinct elements  $t$  and  $t'$ , if  $t$  is in  $\text{Ch}_A(T)$  then  $t$  is in  $\text{Ch}_A(T - t')$ .

If a consultant, for example, has substitutable preferences, then if his preferred set of employees from  $T$  includes student  $t$ , so will his preferred set of employees from any subset of  $T$  that still includes  $t$ . (This follows by repeated application of the definition.) Therefore, the consultant regards student  $t$  and the other students more as substitutes than as complements and continues to want to employ  $t$  even if some of the other students in his choice set become unavailable. Note that responsive preferences are substitutable, so this condition on preferences is a generalization of what has been considered so far.<sup>12</sup> One can similarly generalize the definition of stability as follows.

Consider a matching  $\mu$  which assigns to each consultant  $C$  the set of students  $\mu(C)$  and to each student  $s$  the set of (no more than two) jobs  $\mu(s)$ . The matching  $\mu$  is stable if there is no student  $s$  who would prefer to reject one of the jobs in  $\mu(s)$ , no consultant  $C$  who would prefer to reject one of the students in  $\mu(C)$ , and no student  $s$  and consultant  $C$  who are not matched to one another but who would prefer to be. That is,  $\mu$  is stable if there is no student  $s$  such that  $\mu(s) \neq \text{Ch}_s(\mu(s))$ , no consultant  $C$  such that  $\mu(C) \neq \text{Ch}_C(\mu(C))$ , and no student  $s$  and consultant  $C$  such that  $s$  is not in  $\mu(C)$  [and so  $C$  is not in  $\mu(s)$ ] but such that  $s$  is contained in  $\text{Ch}_C(\mu(C) \cup s)$  and  $C$  is contained in  $\text{Ch}_s(\mu(s) \cup C)$ . It is possible to state the following proposition, which is proved in the Appendix:

**PROPOSITION 3:** *In many-to-two matching, when all agents have substitutable preferences, the set of stable matchings is nonempty.*

<sup>12</sup>Substitutable preferences in two-sided matching were first studied by Alexander Kelso and Vincent Crawford (1982). Charles Blair (1988) showed that, in many-to-many matching with substitutable preferences, the core could be empty even though the set of stable matchings is not. (Proposition 2 shows that this is related to many-to-many matching, not merely to complex preferences.)

In summary, although there are clear similarities between the American market for interns and the regional markets in the United Kingdom for preregistration house officers, there are also some important differences. Some of these have to do with the fact that the U.K. regional markets are both more centralized and much smaller than the American market. Other differences have to do with the fact that students in the U.K. markets seek two jobs. As a result of the latter differences, preferences in the U.K. markets cannot always be modeled as simply as in the American market and require the more general model described in this section.

In considering the operation of the various regional markets, the more general model will be required to establish what is a stable matching. However, when the procedures that produce unstable matchings are considered, it will be possible to demonstrate the instabilities even within the confines of the simpler models.

### III. Matching by Priority: Newcastle, Birmingham, and Edinburgh (1967)

This section considers three closely related matching schemes, all developed in the late 1960's, and all subsequently abandoned.<sup>13</sup> In each of these schemes, a student's ranking of a particular consultant was combined with the consultant's ranking of that student to produce a "priority" for that student to be employed by that consultant. The three schemes differ in the way in which this priority was determined, and each will be discussed below with examples. In each scheme, the overall matching of students with consultants was determined by making the individual matches of students with consultants in order of priority. That is, the first step of each of the three algorithms was to make all of the first-priority matches. Then consultants with unfilled positions and students still needing jobs were scanned to

<sup>13</sup>The Edinburgh scheme was replaced around 1969 by another centralized system, considered in the next section.

identify any second-priority matches, and so on.

The schemes introduced in Newcastle in 1967 (A. G. Leishman and R. P. Ryan, 1970) and in Birmingham in 1966 (Alexander-Williams and Stephenson, 1973) were almost identical. They each used the product of the student's ranking of the consultant and the consultant's ranking of the student as the basis for the priorities. 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"]. The two schemes differed in how they broke ties: in Birmingham, ties were broken in the consultant's favor, so that a (1,2) match would have a higher priority than a (2,1) match. In Newcastle, ties were broken in the student's favor.<sup>14</sup>

In the scheme introduced in Edinburgh in 1967, priorities were lexicographic in consultants' preferences. That is, (1,1) matches were the first priority, followed by (1,2), (1,3), (1,4), and so forth. Only when all consultants' first choices had been exhausted were other matches [(2,1), (2,2), (2,3), etc.] considered. The following example will illustrate the similarities and differences among these three schemes and also prove the following proposition about them:

**PROPOSITION 4:** *Each of these schemes may produce unstable matchings.*<sup>15</sup>

**Example 1.** For simplicity, consider six consultants, each of whom has only one position to fill, and six students, each of whom

needs only one position. (It will be clear that the example does not depend on this simplification.) The rank-orderings of the agents are as follows.

$C_1: s_1, \dots$	$s_1: C_1, \dots$
$C_2: s_1, s_3, s_2, s_4, s_5, s_6$	$s_2: C_2, C_1, C_3, C_4, C_5, C_6$
$C_3: s_3, s_4, \dots$	$s_3: C_4, C_3, \dots$
$C_4: s_4, s_3, \dots$	$s_4: C_3, C_4, \dots$
$C_5: s_1, s_2, s_5, s_3, s_4, s_6$	$s_5: C_1, C_2, C_5, C_3, C_4, C_6$
$C_6: s_2, s_5, \dots$	$s_6: C_5, C_2, \dots$

Then the Birmingham algorithm makes the following matches (the priority is indicated in parentheses after each set of matches):  $C_1s_1$  (1,1),  $C_3s_3$  and  $C_4s_4$  (1,2),  $C_2s_2$  (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 in the Birmingham match they are not matched to each other, but are each matched to their sixth choices.

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 also unstable with respect to  $C_5$  and  $s_5$ .

The Edinburgh (1967) algorithm makes the following matches:  $C_1s_1$  (1,1),  $C_3s_3$  and  $C_4s_4$  (1,2),  $C_6s_2$  (1,6),  $C_5s_5$  (3,3),  $C_2s_6$  (6,2). This outcome is also unstable, but with respect to  $C_2$  and  $s_2$ , who would each prefer one another to their assigned partners (they are each matched to their sixth choices).

So far, the example has been analyzed as if the agents all state their true preferences. Before considering the incentives that agents may have to do otherwise, it will be illuminating to examine the history of these matching systems after their introduction and how they failed and were abandoned.

The various accounts I have received of the demise of these systems all agree on the main events.<sup>16</sup> The following description is

<sup>14</sup>At least initially. A later modification was to reverse this method of tie-breaking. I am indebted to Dr. D. A. Shaw, the Dean of Medicine at Newcastle, for this observation (pers. comm., 13 May 1987).

<sup>15</sup>A stronger result, namely that *any* priority matching scheme will sometimes produce unstable matchings, is proved in Proposition 10 in the Appendix.

<sup>16</sup>I have been able to obtain the least information about the demise of the Edinburgh system, which by 1969 had already been replaced. I have been able to obtain much more detailed accounts, from multiple sources, of the experiences at Newcastle and Birmingham.

from a letter by Dr. John Anderson, the Postgraduate Dean at Newcastle (pers. comm., 6 May 1987):

To understand why our computerized scheme was discarded [in 1981], you should know that in the Northern Region there are 202 recognised posts (this target is set by the DHSS [Department of Health and Social Security]) in 26 approved hospitals. Each year we normally graduate a maximum of 130 students, so that we regularly have a shortfall of at least 70 and usually more, since a number of our graduates obtain pre-registration posts in other regions. We are therefore a major importing region and each six months fill between 50 and 60 posts with graduates of other regions. However, we have never filled more than 185 posts and this means that up to 20 pre-registration posts are regularly unfilled. Sometimes Senior House Officers will be appointed to these posts, but every six months there is a small number of posts that are left unfilled.

This is the background to our problems, and this imbalance between local graduates and posts explains why the computerized scheme failed. Understandably, consultants in the periphery of the region were anxious to fill their posts as quickly as possible and often entered into private arrangements with undergraduates. ...the practice of making private arrangements outwith the computer match scheme gradually spread to the Teaching Hospitals. Those who stuck rigidly to the scheme often found that they were left without any housemen to appoint, as there was no way of preventing these private arrangements and no sanctions could be introduced against those who operated outside the scheme.

In the late 70s and early 80s an increasing number of problems cropped up, mainly concerning conflicts between private arrangements and the formal application procedure. There was a feeling that the computer scheme was an impersonal mechanism which inhibited personal contact between students and consultants and

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 penalized.

The experience in Birmingham was similar, but there the centralized procedure, which was initiated in 1966 for a limited group of hospitals, failed after a few years, was resumed on a larger scale in 1971, failed once more, was restarted again around 1978, and was finally abandoned again around 1981. Of the initial experience, Alexander-Williams and Stephenson (1973) say:

Perhaps the most important cause of failure was the lack of enthusiasm by the consultant staff who did not wish to be deprived of the right to choose their junior staff and so, suspicious that the matching programme might allocate them someone whom they did not want, still tended to make a promise to the first acceptable student who approached them. While there are obviously no objections to first choices being mutually agreed, it soon became known among the undergraduates that certain posts were already promised and so began once again the unseemly struggle that the matching programme was designed to avoid. The breakdown of the scheme was to the disadvantage of the diffident student or the one who waited until he had 'surveyed the field.' [p. 606]

Dr. P. G. Bevan, the director of the board of graduate clinical studies at the University of Birmingham Medical School, writes as follows about the most recent attempt (pers. comm., 2 November 1984):

The main cause of failure for this Matching Plan was the fact that too many Consultants did not abide by the conditions and promised their job to

one particular Student.... In view of this we finally abandoned the Matching Plan three years ago.

A recurrent theme in these accounts is that, after the centralized matching had been in use for a short while, increasing numbers of jobs began once again to be privately arranged in advance between consultants and students and that this worked to the detriment of those who tried to participate in the scheme without prior arrangement.

To understand this phenomenon, consider now the incentives which these priority procedures give to the agents. For this purpose, consider again Example 1. (In the example, there are equal numbers of students and positions, but it will be clear that the behavior described below could exist at least as easily when there is an imbalance between the two.) To make the example clear, suppose consultants  $C_1$ – $C_4$  are in the most desirable teaching hospital,  $C_5$  is in the next most desirable regional hospital, and  $C_6$  is in a relatively undesirable rural hospital. Similarly, suppose students  $s_1$ – $s_5$  are all top graduates of their medical school, while  $s_6$  has a less distinguished record.

Then, under the Birmingham or Newcastle system,  $C_5$  is gravely disappointed to learn that his new junior house officer will be  $s_6$ , all the more when he learns that student  $s_5$ , whom he liked reasonably well, is quite unhappy with his own appointment and would have preferred to work for  $C_5$ . If  $C_5$  had submitted a rank-ordering 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 rank-ordering on which  $C_5$  was his first choice. The example shows there may be incentives for both students and consultants to submit rank-orderings different from their true preferences.

Furthermore, these priority-ranking systems allow such incentives to build upon one another, so that as more agents adapt their submitted rank-orderings to improve their matches, the greater is the incentive for other agents to do so. To see this, suppose  $C_5$  in Example 1 resolves not to suffer the same fate the following year. He there-

fore approaches one of the good students in the next year's class, in advance of the formal match, and suggests that they mutually agree to be matched, which they will accomplish by ranking one another first in the formal match.<sup>17</sup> The student, chastened by 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, which is  $t_1, t_2, t_3, t_4, t_5, \dots$ , and  $t_4$  submits his true rank-ordering  $C_3, C_4, C_5, C_2, \dots$ . 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. However, since the product of their rankings is 16,  $C_2$  could well end up with his 15th-choice student. So when some matches have been prearranged, those not in the know, students as well as consultants, stand to do very poorly if they do not also prearrange their matches. Furthermore, when an agent's top  $n$  choices have all arranged to indicate only a first preference in the formal match (as in the above quote from Anderson), then the agent can do no better in the match than to reach such an agreement himself with his  $n + 1$ st choice. Therefore, the following proposition, which applies as well to the Edinburgh (1967) system,<sup>18</sup> has been proved:

**PROPOSITION 5:** *It is not a dominant strategy for any agent to submit his true preferences in these priority matching systems.*

<sup>17</sup>In the context of these relatively small markets, both parties to such an agreement can be confident that it will be carried out, since 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.

<sup>18</sup>Under the Edinburgh (1967) system,  $C_2$  in Example 1 could have improved his match by ranking  $s_2$  first, and  $s_2$  could have improved his match by ranking  $C_6$  as unacceptable.

*Furthermore, there are multiple equilibria at which all agents submit only a first choice.*

In fact, Proposition 5 does not capture the full strength of what has been proved. Under priority matching, a student and consultant who rank one another first will be matched regardless of what the other agents do. So the problems of coordination that afflict many equilibria do not arise here: pairs of agents may secure their part of the equilibrium by private arrangement. These results thus go a long way toward explaining both why a high percentage of appointments were soon arranged in advance under these systems and why this worked to the disadvantage of those who tried to arrange employment through these priority-based formal match procedures.<sup>19</sup>

#### **IV. Stable Matching: Edinburgh (1969) and Cardiff**

This section considers two very closely related systems, both built around the same stable matching algorithm. That the two systems are closely related is not an accident: the Cardiff system was adapted around 1971 from the system initiated in Edinburgh around 1969 to replace the Edinburgh (1967) priority matching system. However the two markets are rather different, with Cardiff regularly having more positions than local graduates and with Edinburgh regularly having more graduates than positions.<sup>20</sup> Both systems remain in operation today,

having achieved and maintained high rates of orderly participation.

It is also not an accident that the basic algorithm produces stable matchings, but it is a curious bit of intellectual history. In the 1960's and 1970's the architects of the various matching systems in the United Kingdom knew of the experience of the American markets since 1952, but it was not then known that the American algorithm produced a stable matching. However, unconnected with any of these markets, David Gale and Lloyd Shapley (1962) formulated a simple model of one-to-one matching and defined the set of stable matchings, which in that model equals the core of the game. They also developed a "deferred-acceptance" algorithm (described below) for producing stable matchings and observed that it could be applied as well to problems of many-to-one matching.<sup>21</sup> Two British computer scientists, D. G. McVitie and L. B. Wilson further explored this algorithm, and apparently through their work, the algorithm came to the attention of Dr. H. R. A. Townsend, the author of the Edinburgh (1969) matching procedure, known as PRAMS (for Pre-Registration Appointment Matching Scheme).<sup>22</sup> He used Gale and Shapley's deferred-acceptance algorithm as the basis for PRAMS, adapting it to the requirements of the preregistration market and to the local conditions in Edinburgh. The Edinburgh PRAMS algorithm was subsequently adopted in Cardiff, where it was

<sup>19</sup>It appears that something very similar (if not identical) to a priority matching scheme was tried and subsequently abandoned at Sheffield, but I have not included that system among those formally analyzed here because the match was done by a committee, whose exact procedures therefore cannot be determined. However, A. D. Clayden and James Parkhouse (1971 p. 9) report a computer program designed in their words "to mimic the manual allocation," and that program implements a priority algorithm like the Edinburgh (1967) system, except in giving lexicographic priority to students' preferences rather than to consultants' preferences.

<sup>20</sup>The Edinburgh system is furthermore open to students from medical schools in other regions (Sir James Fraser, pers. comm., 7 May 1987).

<sup>21</sup>Gale and Shapley (1962) concentrated primarily on the one-to-one matching problem, and for many years thereafter one-to-one and many-to-one matching came to be regarded as essentially equivalent. That they are not, in ways that are important for the subject at hand, was first observed in Roth (1985a).

<sup>22</sup>Townsend (1981) cites McVitie and Wilson (1971) in the PRAMS.80 manual, which documents the current (as of 1984) version of his computer program, and which he graciously sent to me (pers. comm., 23 November 1984; see also McVitie and Wilson, 1970a,b). Townsend himself is a clinical neurophysiologist, who at the time also held a part-time appointment in Edinburgh's department of Machine Intelligence and Perception, and it was in that capacity that he undertook the design of the PRAMS system (pers. comm., 27 July 1989).

further adapted to local conditions. Before discussing the adaptations introduced in Edinburgh and Cardiff, I will consider first the unmodified<sup>23</sup> deferred-acceptance algorithm for many-to-one matching.

Step 1: Each consultant  $C$  with  $q_C$  positions makes offers to his  $q_C$  highest-ranked acceptable students (or to all of them if there are fewer than  $q_C$ ).

...

Step  $k$ : (i) Each student  $s$  rejects all but the highest-ranked of the acceptable offers he has received in steps 1 through  $k-1$ . (ii) Each consultant  $C$  who has received  $r \geq 1$  rejections in part (i) of step  $k$ , and who now has  $q_C - r$  (nonrejected) offers outstanding, makes offers to his  $q_C - r$  highest-ranked acceptable students among those who have not yet rejected him. (iii) The algorithm stops at any step  $k = T$  at which no rejections are issued, and the resulting matching places each student with the consultant (if any) whose offer he has not rejected and leaves unmatched any student not holding an offer.

Since each student holds at most one unrejected offer at any step of the algorithm and since no consultant makes an offer twice to the same student, the algorithm stops and produces a matching. This matching is stable, because if some consultant  $C$  would prefer a different group of students than he receives, then if he has responsive preferences<sup>24</sup> he will have proposed to those stu-

dents at an earlier step of the algorithm and been rejected. This implies that these students prefer the positions they get from the algorithm to  $C$ , so the matching is stable.

Note that there is another version of the algorithm in which students make applications for positions and each consultant refuses all but the best  $q_C$  acceptable offers from among those he has received. While the Edinburgh matching scheme adapts the consultant-proposing version of the algorithm, the present Cardiff version adapts the student-proposing version.<sup>25</sup> Both versions produce stable matchings in many-to-one markets.

The markets to which this algorithm was adapted involve many-to-two matchings, and as has already been seen from Proposition 2, this changes the market in important ways. Furthermore, some of the adaptations in both Edinburgh and Cardiff imply that the preferences of the agents are not responsive. Consider the following constraints:

- 1) Each student desires no more than one medical and one surgical position.
- 2) Edinburgh surgeons may specify that they will employ no more than one female house officer in any six-month period.
- 3) In Cardiff (for some of the period under consideration, but not presently), at most one of a student's two positions could be a teaching-hospital position.<sup>26</sup>

<sup>23</sup>Except to allow that some students and consultants may find some matches unacceptable.

<sup>24</sup>Responsive preferences, which play an essential role in the argument for the case of many-to-one matching, were introduced in Roth (1985a). Earlier treatments of many-to-one matching had argued from analogy with the case of one-to-one matching and had not considered that  $C$ 's preferences must be based on comparisons of groups of students. A notable exception is the paper of Kelso and Crawford (1982), which considers preferences defined directly over groups, without reference to an underlying preference over individuals.

<sup>25</sup>Over the years, the computer code used in Cardiff has undergone a number of programming and procedural changes designed to cope with increasing numbers of positions and with changes in available computers. The current system goes under the name of PASHA (Preferential Allocation System for House Appointments), and during 1973–1982 it went under the name of CHAMP (Computerized House Appointments Matching Plan), during which time the consultant-proposing version of the algorithm was implemented. I am indebted to Dr. Kelvyn Johns for documentation of the various systems (pers. comm., 6 November 1984 and 23 June 1989).

<sup>26</sup>These descriptions abstract somewhat from the actual adaptations. In Edinburgh, a surgeon may actually only specify that he does not wish to employ more than two female house officers, since his positions for two consecutive six-month periods are allocated simul-

Constraints 1 and 2 obviously concern agents' preferences (i.e., certain sets are unacceptable), and constraint 3 can be modeled as part of students' preferences, since it has the effect that students will not (because they cannot) choose two teaching-hospital positions, either during the formal match or in any postmatch exploration of potential instabilities. The Edinburgh adaptation of the algorithm to constraints 1 and 2 is straightforward: students may hold up to two offers at any step in the algorithm but no more than one medical and one surgical, and they must reject the rest. In the Cardiff adaptation to constraints 1 and 3, students could apply at any step of the algorithm to no more than one medical or surgical position and to no more than one teaching-hospital position.<sup>27</sup> Subject to these constraints, students' and consultants' offers and rejections within the algorithms are otherwise governed by their submitted rank-orderings of individual positions, as in the deferred-acceptance algorithm described above. (In Edinburgh, students submit two preference lists, one for surgical positions and one for medical. In Cardiff, they submit one preference list containing both kinds of positions.<sup>28</sup>

---

taneously. It then remains to schedule the house officers so that two female house officers are not assigned to the same six-month period. Scheduling may present other difficulties as well, some of which may involve "higher-order" instabilities, which will be briefly discussed later. In Cardiff, when constraint 3 was in effect, if some teaching-hospital positions were left unfilled by the initial run of the algorithm, these positions were then "unmarked," so that a student who already had one teaching-hospital position became eligible to fill them.

<sup>27</sup>This might involve, for example, a student applying at the first step of the algorithm to his first-choice medical position in the teaching hospital and to his fifth-choice surgical position. If he were subsequently rejected by the medical position, he might wish to apply to a surgical position in the teaching hospital, which might be his second-choice position. The algorithm gave him the opportunity to do so, which involved withdrawing his application from his fifth-choice surgical position. Thus, unlike the original deferred-acceptance procedure, applications could be withdrawn as well as rejected.

<sup>28</sup>This allows medical and surgical teaching-hospital positions to be compared, as they must be to imple-

**PROPOSITION 6:** *Student preferences satisfying constraint 1 and consultant preferences satisfying constraint 2, but otherwise responsive to a simple rank-ordering, are substitutable. Student preferences satisfying constraints 1 and 3 and otherwise responsive to a simple rank-ordering need not be substitutable.*

Together with Proposition 3, Proposition 6 (which is proved in the Appendix) establishes that stable matchings continue to exist in the Edinburgh market and in the current Cardiff market. The following proposition (also proved in the Appendix) states that the algorithms adapted as described for the Edinburgh market and for the present Cardiff market (without constraint 3) in fact produce stable matchings:

**PROPOSITION 7:** *The consultant-proposing deferred-acceptance algorithm adapted to preferences that obey constraints 1 and 2 but are otherwise responsive to a simple rank-ordering produces a stable matching, as does the student-proposing algorithm adapted to preferences that obey constraint 1.*

There are some complications that have so far been passed over, the chief of which involves scheduling the jobs each student has been assigned into the August and February starting periods in a way that gives each consultant the right number of house officers in each period. Occasionally (although apparently rarely) the necessary scheduling may be infeasible, as when a number of individuals require a job at a particular time, as a consequence of having arranged their other assignment outside of the region. These situations are resolved at the discretion of the PRAMS committee, typically by editing a job from the preference lists of one of the students (Townsend, 1981 p. 42). Of course this may produce an instability, but it is not one that could be

---

ment condition 3. Prior to 1973, winter positions in Cardiff were matched before summer positions, but since 1973 they have been matched simultaneously.



predicted in advance, nor can it be acted on following the match, given the rules by which jobs are formally assigned. Although students are invited to indicate the order in which they prefer to fill their medical and surgical positions, these preferences are not honored if they interfere with a feasible scheduling, so there may be some instabilities involving exchange of time slots among groups of students and consultants. I call these instabilities "higher-order" to indicate that they involve coalitions larger than a student and consultant. In the case of scheduling changes, they involve at least two students and two consultants. As noted in Proposition 2, such higher-order instabilities may result from other causes as well, but such higher-order instabilities are likely to be of lesser importance, since they are so difficult to act on. A case in point is the allocation of married couples, which poses the same theoretical difficulties in the United Kingdom as in the United States, but which seems to have had less practical consequence. Whereas an American couple needs to find two jobs in the same location, in the United Kingdom a couple may need to find four jobs in two locations (in two time periods), and the difficulties of identifying a set of such jobs that both the couple and the relevant consultants prefer to those allocated by the match are formidable.

The situation in Cardiff was even more complex during the years in which students were constrained to hold no more than one teaching-hospital position. Aside from the higher-order instabilities just discussed, the combination of constraint 3 with constraint 1 on the preferences may sometimes have the consequence that no stable matchings exist (see Example 4 in the Appendix). In such circumstances, as well as in certain others<sup>29</sup> the Cardiff scheme would produce matchings with some instabilities. As near as I can determine (and I do not know how

to make this precise), these instabilities arose relatively rarely.<sup>30</sup> What is certain is that they could not be predicted in advance, and so neither the Cardiff system (with or without constraint 3) nor the Edinburgh system allows mutually beneficial "prematch" agreements of the kind discussed in the previous section.<sup>31</sup>

#### V. Matching by Linear Programming: Cambridge and The London Hospital Medical College

This section considers two closely related matching schemes which do not produce stable matchings in terms of the stated preferences but which are still in use. The first was developed in 1973 at the London Hospital and its Medical College, and the second was developed in 1978 at the University of Cambridge School of Clinical Medicine. Both schemes involve the linear-programming assignment algorithm.

The London Hospital scheme takes as input the rank-orderings of students and consultants.<sup>32</sup> Numerical weights are assigned to choices. These are summed for each potential student-consultant pair.

<sup>30</sup>I have no information suggesting that constraint 3 was discarded *because* of its potential to produce instabilities; apparently, the facility for automatically accommodating married couples was also eliminated at around the same time, for reasons of simplicity (K. Johns, pers. comm., 23 June 1989).

<sup>31</sup>Note, however, that in these systems also, students and consultants may assure themselves of being matched by ranking one another first. In this connection, a noteworthy feature of the Cardiff system is that consultants may make their preference list publicly available before the students submit their own preference lists, and in recent years most of them have apparently done so (K. Johns, pers. comm., 3 August 1989). Thus, students will often know where they are in the consultants' rankings before having to submit their own.

<sup>32</sup>The students are all graduates of the London Hospital Medical College, and the consultants include all those offering house-officer posts at the London Hospital, together with a group of affiliated hospitals. In a recent year there were approximately 40 posts in London and 70 in regional hospitals. I am indebted to Dr. F. J. Goodwin, the Postgraduate Sub-Dean at the London Hospital Medical College for this information (pers. commun., 31 October, 1984).

<sup>29</sup>The full computer code for Cardiff PASHA (as of 1984, including the code for handling constraint 3) is more than 2,000 lines long and creates many special situations, such as those involving filled and unfilled teaching-hospital positions.

A. R. Shah and S. C. Farrow (1976) report that for the first five uses of the algorithm (once in 1973 and twice in 1974 and in 1975) choices 1, 2, 3, and 4 were given weights of 20, 14, 9, and 5, respectively. Thus (1,1) matches received a weight of 40, (1,2) and (2,1) matches each received a weight of 34, and so forth. The resulting weights for each potential student–consultant pair form the basis for the linear-programming assignment problem of matching students to consultants so as to maximize the sum of the matches. Shah and Farrow report that:

The general procedure is for the computer solution to be submitted to the sub-committee of the academic board. On each occasion it has formed a very satisfactory basis for the final solution, although several hand adjustments have been made.<sup>33</sup> [p. 477]

The Cambridge scheme was first implemented in 1978 for posts beginning in February and August 1979. (Students, who normally graduate in December, may apply for positions beginning in February, in August, or for one of each.) Students rank consultants as A, B, C, or unacceptable. “Students may make only two A choices for medicine, and only one of these may be for a medical job at Addenbrooke’s, the main teaching hospital. Similarly for surgery. As many B and C choices as they wish are allowed” (D. M. Taylor, pers. comm., 5 June 1987). Consultants rank students similarly (but after learning how the students ranked them, and without a constraint on the number of A rankings.) As in the London scheme, weights are given to each potential match, but in the Cambridge scheme these weights are lexicographic in consultants’

preferences, so consultant–student (A,A) matches have the highest weight, followed by (A,B), (A,C), (B,A), and so forth. As in the London scheme, the matching giving the highest total weight forms the basis of the matching of students and consultants. This matching is also subject to some adjustment (e.g., students are permitted to have only one of their two jobs at Addenbrooke’s Hospital). Example 2 proves the following proposition:

**PROPOSITION 8:** *Both of these schemes may produce unstable matchings.<sup>34</sup> Furthermore, they may fail to make (1,1) matches.*

*Example 2.* For simplicity, consider three consultants, each of whom has only one position to fill, and three students, each of whom needs only one position. Their rank orderings are as follows:

$C_1: s_1, s_2, s_3$	$s_1: C_1, C_2, C_3$
$C_2: s_1, s_3, s_2$	$s_2: C_1, C_3, C_2$
$C_3: s_3, s_1, s_2$	$s_3: C_3, C_2, C_1$

The unique stable matching is  $\mu$  such that  $\mu(C_1) = s_1$ ,  $\mu(C_2) = s_2$ , and  $\mu(C_3) = s_3$ . [This follows since  $(C_1, s_1)$  and  $(C_3, s_3)$  are both (1,1) matches and so must be included in any stable matching.] However, the London Hospital scheme (with weights as given above) gives  $\mu$  a weight of 98, while the highest total weight (108) is achieved by the unstable matching  $\nu$  with  $\nu(C_1) = s_2$ ,  $\nu(C_2) = s_1$ , and  $\nu(C_3) = s_3$ . Similarly the Cambridge scheme [with (A,A) matches worth 8, (A,B) worth 7, and so on, down to (C,C) worth 0 for this example] chooses  $\nu$  (which has a total weight of 20, compared to  $\mu$  with total weight of 16).

In this example, no individual agent can improve his outcome by changing his stated preference.<sup>35</sup> More to the point, even when

<sup>33</sup>Shah and Farrow (1976 p. 477) also note that, to adjust the match, the preferences of different individuals may be weighted differently. “[I]n July 1974 the initial solution led to several applicants, who had completed 6 months, not being allocated a second appointment. This made it necessary to rerun the programme with a reduction in weight of newly qualified applicants. This second run achieved the designed outcome.”

<sup>34</sup>That the London scheme may produce unstable matchings was noted by Townsend (1977).

<sup>35</sup>Actually, this depends on how unacceptable matches are weighted. When unacceptable matches are part of the matching with the highest total weight, they correspond to unmatched students and positions.

all pairs of agents but one organize themselves into (1, 1) matches (as in the example), they cannot be sure of being matched (unless they both rank all other options as unacceptable). Thus, private arrangements are more difficult to make and are less certain than in any of the other matching schemes considered here. Nevertheless, the following parallel to Proposition 5 for priority matching systems applies as well to these linear-programming systems; the proof is left to the reader:

**PROPOSITION 9:** *It is not a dominant strategy for any agent to submit his true preferences in these linear-programming matching systems. Furthermore, there are multiple equilibria at which all agents submit only a first choice.*

While these systems have some important differences from the failed priority matching systems [as exemplified by the fact that (1, 1) matches are not assured of forming], they have enough significant similarities so that one may wonder how to account for the longevity of the London Hospital and Cambridge schemes. One hypothesis is that the environments in which the markets are conducted differ significantly from other environments: each of these two schemes is for the graduates of a single medical school, on the small end of the range of markets considered here.<sup>36</sup> Thus, there may be social and other kinds of pressures that make it difficult to circumvent the formal matching scheme. In this regard S. C. Farrow (pers. comm., 23 October 1984) writes:

Candidates are not obliged to accept allocated posts and have been known to decline; this is of course frowned upon by the scheme and they try hard

to dissuade people from taking such a course.

Similarly, Goodwin (pers. comm., 31 October 1984) adds:

Virtually all candidates accept the posts allocated to them. Indeed, we achieve a match of the candidates' and consultants' first choices in around 70 percent of cases which is obviously satisfactory to all concerned. The remainder nearly always accept the post to which they have been allocated. Of course, we have no legal right to prevent someone from declining a post after it has been allocated to him but this would be considered pretty bad form and the candidates know it.

Thus, in markets of this size and composition, it may be that participation in the matching scheme may be somewhat less voluntary than in larger markets or markets in which many of the agents are not so intimately connected with one another.<sup>37</sup>

Another hypothesis, in view of the high reported percentage of (1, 1) matches, is that the agents manage to adapt to the system by coordinating among themselves before the formal match or by modifying the rank-orderings they submit. In such a case, the outcome of the match might even be stable. (To test such a hypothesis, it would be necessary to have better information about the submitted rank-orderings than I presently have. However, see Susan Mongell and Roth [1991] for an account of an unstable match-

<sup>36</sup>They are thus not completely "impersonal" markets, particularly since the most desirable positions are in the associated teaching hospital, and the most desirable house officers are the top students in the class. Since total class sizes are around 100, the key players on both sides of the market have often had the opportunity to get to know one another in the course of the students' medical education.

<sup>37</sup>B. T. Colvin, the present Postgraduate Medical Sub-Dean at The London Hospital Medical College, writes that "Much of the goodwill in the system relies on the Postgraduate Sub-Dean's personal knowledge of the candidates, consultants and posts, and his ability to impose on both parties the moral obligation to comply with the allocations..." (pers. comm., 10 August 1989). To the extent that agents in this market are obliged or feel obliged to comply, it would of course be unsurprising that unstable matchings and matching procedures can persist: the National Football League draft and the process by which graduates of the U.S. Naval Academy obtain their first assignments (see Roth and Marilda Sotomayor, 1990) are good examples of such markets.

ing procedure for which such data were available; the data gathered for that procedure supported an equilibrium misrepresentation hypothesis of this kind.)

In either case, these matching schemes and the environment in which they operate appear to make prematch coordination both more difficult and less rewarding than do the priority matching schemes discussed in Section III.

## VI. Interregional Instabilities and the Current State of Affairs

Students who fail to obtain two positions in the regional markets can participate in a secondary market, called the Safety Net, organized by the Councils for Postgraduate Medical Education,<sup>38</sup> which distributes information on unfilled posts and unplaced students. Consultants with unfilled positions may hire senior house officers (instead of preregistration house officers) after a certain date. In 1983, about 25 graduates of U.K. medical schools failed to obtain any preregistration position, in part because some consultants apparently hired senior house officers before the preregistration market had cleared. The incentives for consultants to do so were apparently heightened by the fact that some students would break arrangements made in this market if a more desirable position became available at the last minute (see Henry Yellowlees, 1983; E. D. Acheson, 1984).

In response to these events, a working group was established to study the current market, and their 1987 report (Department of Health and Social Security, 1987) describes the current state of affairs as follows:

...matching arrangements are tending to take place increasingly early in the

clinical years. In some cases, students are matched to posts even before they have entered the final clinical year.... The earlier a student is matched to a post, the greater the chance that for some reason he or she may decide not to fulfil the arrangements but to take some other post instead. Where medical schools run particularly early matching schemes it appears that some Districts not involved in the schemes then advertise their pre-registration posts at an even earlier date in order to try to pre-empt the matching scheme. This has led to some posts being advertised as much as 18 months before their start date. [p. 11]

The report also notes that medical schools should help in:

...ensuring that their unmatched students are made aware that it is unethical, while holding an offer for one post, to accept an offer for another before negotiating release from the prior bargain.... [p. 11]

What seems to be occurring is that, as the centralized matching schemes in individual regions of the National Health Service have unravelled backward in time and been abandoned, the situation in most regions today has come to resemble that of the late 1960's.<sup>39</sup> The instabilities that underlie the contemporary problems appear to involve not only instabilities in a given region, but also instabilities between regions.<sup>40</sup>

<sup>39</sup>In another contemporary report, J. H. Gillard and T. H. S. Dent (1988 p. 344) note that "Both matching schemes and free markets with an official start date were reported to be vulnerable to pre-empting. This criticism was conveyed most vigorously from Nottingham and Southampton [both with "free market" decentralized systems], and caused considerable anxiety. Students feared that their colleagues were making private arrangements with consultants, while consultants were keen to avoid missing the more attractive students. If a formal match operates later, students might waste an application if the posts have been covertly allocated by student and consultant agreeing to place each other first in their order of preference."

<sup>40</sup>Instabilities involving positions in other regions can affect even regions with stable regional matching

<sup>38</sup>The Council for Postgraduate Medical Education in England and Wales has recently been replaced by the Standing Committee on Postgraduate Medical Education (SCOPME). As of 1 April 1989, the responsibility for the Safety Net has passed to the Department of Health (no longer DHSS, social security having been separated from health in mid-1988).

In view of the experience in the regional markets and in the American market, perhaps the most promising plan to remedy the unraveling of appointment dates and the problems that accompany it would be to replace the patchwork of regional markets with a national market organized around a set of stable procedures (such as those in use in Edinburgh).<sup>41</sup>

## VII. The Implications for Other Markets

The environment in which the various U.K. markets operate differs in important ways from that of the American market in size, in the number of jobs sought by each student, and in the centralization of authority (at least within a given region). Still, the experience of the various centralized matching schemes considered here allows two very different hypotheses about matching markets which might have been formed on the basis of the evidence of the American market to be rejected.

The first of these hypotheses, which might be called the “pure-transaction-cost hypothesis,” is that, because a centralized matching mechanism reduces the high transaction costs found in the chaotic decentralized markets, *any* centralized procedure would obtain high rates of participation. The rapid failure of the priority matching schemes in the regions where they were tried is clear evidence against this hypothesis.

The second of these hypotheses, which might be called the “pure-efficiency hypothesis” (or perhaps the “core hypothesis”) is that, in order to achieve high rates of participation, a scheme must produce matchings that do not allow coalitions of *any* size to profit by defecting from the scheme. The

higher-order instabilities that can occur in the Edinburgh (1969) and Cardiff schemes (and which Proposition 2 suggests may be endemic), not to mention the pairwise instabilities that can arise in the linear programming schemes, suggest that this is not the case either.

A more accurate description seems to lie somewhere in between. Centralized matching schemes like the priority matching schemes, which make it easy for pairwise coalitions to defect profitably, experience widespread defection even when the system is endowed with enough authority to require at least pro forma participation. However, orderly participation in a centralized matching scheme appears to be much more likely when the opportunities it presents for defection are rare and difficult to find or when they primarily involve large coalitions.

Another observation that may have considerable generality is that markets of this kind have a propensity to unravel backwards in time, with dates of appointment becoming earlier and earlier. This phenomenon seems to occur in very different kinds of markets. (Indeed, the membership drives of American sororities are called “rush” precisely because they have unraveled in this way [see Mongell and Roth, 1991]. Among the markets presently experiencing this sort of unraveling are some of those for newly graduated lawyers, about which I hope to have more to say in the future.) This suggests that the kind of models considered here may be useful tools with which to explore a wide variety of entry-level labor markets and related matching processes. Centralized markets have been a productive place to begin this kind of study, because they make it relatively straightforward to determine the “rules of the game” in the kind of detail demanded by game-theoretic analysis. However, with the experience gained from studying such markets, it should be possible to study decentralized markets in related ways.<sup>42</sup>

---

procedures. Sir James Fraser (pers. comm., 10 July 1987), writing of the Edinburgh PRAMS, notes that “... one of the principles behind the Scheme is that a student is committed to accept the Unit to which he is allotted. Rarely, this formal agreement is broken, more commonly by applicants from outside Edinburgh...”

<sup>41</sup>Of course, there may be formidable and perhaps intractable problems of coordination, having to do with different schedules and jurisdictions, involved in setting up such a national market.

<sup>42</sup>A theoretical question related to the study of decentralized markets has recently been resolved in a paper by Roth and John Vande Vate (1990), which

In closing, a few words are probably in order about the methodology this paper shares with the related empirical studies of matching in Roth (1984a) and Mongell and Roth (1991). The chief analytical tools (stability and strategic equilibrium) are a mix drawn from what has traditionally been called cooperative and noncooperative game theory. Their use together here should help make clear why this is not always a useful distinction, since these two approaches to game theory are complements, rather than substitutes. While the rules of the game must be identified in great detail in order to speak about equilibrium, the stability of a matching may be discussed somewhat independently of the specific rules of the market. One of the phenomena that emerges from these studies is that, when the outcomes are unstable, agents have incentives to change the rules of the game, as when they decide to introduce a centralized matching procedure or to defect from one. In principle, all such decisions could be modeled in terms of some larger, all-encompassing game, and much of the contemporary theoretical literature in game theory takes this point of view. However, this will seldom be an option in modeling complex real systems, whose rules need to be determined by observation (and determining rules by observation will often be subject to uncertainty, since as has been seen, formal rules sometimes turn out to be less than fully binding, while informal unwritten rules may impose real constraints). If game theory is to play as important a role in empirical economics as it already plays in economic theory, increased attention must be paid to modeling complex games whose rules can only be observed imperfectly.

#### APPENDIX

Proposition 10 below is an impossibility result that strengthens the conclusions of

---

shows that a wide class of random processes converge with probability 1 to a stable matching. Some preliminary results on strategic behavior in such markets are contained in Roth and Vande Vate (1991).

Proposition 4 and shows that no priority matching scheme always produces stable matchings. An arbitrary priority matching scheme is given by a priority function  $f: \mathbf{N} \cup \{x\} \times \mathbf{N} \cup \{x\} \rightarrow \mathbf{N} \cup \{x\}$ , where  $\mathbf{N}$  is the set of positive integers and  $x$  indicates that a match is unacceptable. Thus, if  $a_{ij}$  is student  $j$ 's rank in consultant  $i$ 's ordering (i.e.,  $a_{ij} = k$  means student  $j$  is consultant  $i$ 's  $k$ th choice) and  $b_{ji}$  is consultant  $i$ 's rank in student  $j$ 's ordering,  $f(a_{ij}, x) = f(x, b_{ji}) = x$ , and  $f(a_{ij}, b_{ji}) \in \mathbf{N}$ , where  $f(a_{ij}, b_{ji})$  is the priority of the match between  $i$  and  $j$ . The matching  $\mu_f$  is created by performing all first-priority matches [ $f(a_{ij}, b_{ji}) = 1$ ], followed by all second-priority matches among those agents not yet matched, and so forth. Note that the function  $f$  restricted to acceptable matches must be one-to-one (i.e.,  $f: \mathbf{N} \times \mathbf{N} \rightarrow \mathbf{N}$  is one-to-one) because, for example, (1,2) matches cannot have the same priority as (2,1) matches, since a given consultant might be part of a (1,2) match with  $s_1$  and a (2,1) match with  $s_2$ . Without loss of generality (by taking a monotone transformation of  $f$  if necessary), one may also assume that the function  $f$  is onto  $\mathbf{N}$  [i.e., that there are values of  $a_{ij}$  and  $b_{ji}$  for which  $f(a_{ij}, b_{ji})$  equals 1, 2, etc., with no gaps]. The impossibility result can now be stated.

**PROPOSITION 10:** *No priority matching function  $f$  always produces stable matches  $\mu_f$ .*

#### PROOF:

It will be sufficient to prove the proposition for the symmetric case of one-to-one matching. First observe that if (1,1) matches are not first-priority (i.e., if  $f(1,1) \neq 1$ ) then  $f$  does not always produce stable matchings. To see this, suppose  $(i,j)$  matches are first-priority, and suppose  $j > 1$ . Let  $a_{21} = i$  and  $b_{12} = j$ , so  $C_2$  and  $s_1$  are matched at the matching  $\mu_f$ ; but let  $a_{11} = 1 = b_{11}$ , so  $C_1$  and  $s_1$  are a (1,1) match. Then  $\mu_f$  is unstable with respect to  $C_1$  and  $s_1$ , which proves the first observation.

Next observe that if  $(i,j) \neq (1,1)$  is second-priority and  $i > 1$  then  $\mu_f$  may be unstable, even if (1,1) matches are first-priority. To see this, let (1,1) matches be first-priority and let  $C_1, C_2, s_1, s_2$ , and  $s_3$  be

such that  $a_{11} = 1$ ,  $a_{12} = i > 1$ ,  $b_{11} = 2$ ,  $b_{12} = 1$ ,  $a_{23} = 1$ ,  $b_{21} = j$ , and  $b_{32} = 1$ . Then  $C_2$  and  $s_3$  are matched in the (1,1) phase, and  $C_1$ ,  $s_1$ , and  $s_2$  are not (if  $j \neq 1$ , let  $b_{22} = 1$ ). Therefore,  $C_1$  and  $s_2$  are matched in the  $(i, j)$  phase, and the resulting matching is unstable with respect to  $C_1$  and  $s_1$ , which proves the second observation.

A symmetric argument shows that unstable matchings may result unless  $j = 1$  also. However, since  $(i, j) \neq (1, 1)$ , this completes the proof of the proposition.

A formal model of matching sufficiently general for all the cases covered here is the following. There are two sets of agents,  $C = \{C_1, \dots, C_n\}$  and  $S = \{s_1, \dots, s_m\}$ , namely consultants and students. Each consultant  $C$  seeks to fill no more than some number  $q_C$  of positions, and each student  $s$  seeks to obtain no more than  $q_s = 2$  positions. A *matching* is a function from  $C \cup S$  into the set of all subsets of  $C \cup S$  such that, for every  $C$  in  $C$  and  $s$  in  $S$ :

- 1)  $\mu(C)$  is contained in  $S$  and  $\mu(s)$  is contained in  $C$  (either set may be empty);
- 2)  $|\mu(C)| \leq q_C$  for all  $C$  in  $C$ ;
- 3)  $|\mu(s)| \leq q_s = 2$  for all  $s$  in  $S$ ;
- 4)  $s$  is in  $\mu(C)$  if and only if  $C$  is in  $\mu(s)$ .

Each agent  $A$  has preferences over sets of agents on the other side of the market that determine a choice function  $\text{Ch}_A$  as in Section II. Each agent's preferences over matchings are sensitive only to his own part of the matching, so that if  $\mu$  and  $\nu$  are matchings such that  $\mu(A) = \nu(A)$ , then agent  $A$  is indifferent between  $\mu$  and  $\nu$ .

An agent  $A$  has *responsive* preferences if he has a rank-ordering of the individual agents on the other side of the market such that, for any subset  $T$  of such agents,  $\text{Ch}_A(T)$  is the set of the  $q_A$  highest-ranked acceptable individuals in  $T$  (or all the acceptable individuals in  $T$  if there are fewer than  $q_A$ ). Note that this is a slightly weaker definition than that given in Section I, since this definition deals only with the agents' choice sets and ignores preferences between subsets  $T$  that would not arise as choices from a common set. Note also that responsive prefer-

ences are substitutable, as defined in Section II.

The following example proves the second part of Proposition 2, namely that stable matchings need not be in the core, nor even be Pareto optimal, in many-to-many matching with responsive preferences.

*Example 3.* There are four consultants, each of whom seeks two house officers, and four students, each of whom seeks two positions. The (responsive) preferences of the agents are given by the following rank-orderings of acceptable subsets:

- $s_1$ :  $\{C_1, C_2\}$ ,  $\{C_1, C_3\}$ ,  $\{C_1, C_4\}$ ,  $\{C_2, C_3\}$ ,  $\{C_2, C_4\}$ ,  $\{C_3, C_4\}$ ,  $\{C_1\}$ ,  $\{C_2\}$ ,  $\{C_3\}$ ,  $\{C_4\}$   
 $s_2$ :  $\{C_2, C_1\}$ ,  $\{C_2, C_4\}$ ,  $\{C_2, C_3\}$ ,  $\{C_1, C_4\}$ ,  $\{C_1, C_3\}$ ,  $\{C_4, C_3\}$ ,  $\{C_2\}$ ,  $\{C_1\}$ ,  $\{C_4\}$ ,  $\{C_3\}$   
 $s_3$ :  $\{C_3, C_4\}$ ,  $\{C_3, C_1\}$ ,  $\{C_3, C_2\}$ ,  $\{C_4, C_1\}$ ,  $\{C_4, C_2\}$ ,  $\{C_1, C_2\}$ ,  $\{C_3\}$ ,  $\{C_4\}$ ,  $\{C_1\}$ ,  $\{C_2\}$   
 $s_4$ :  $\{C_4, C_3\}$ ,  $\{C_4, C_2\}$ ,  $\{C_4, C_1\}$ ,  $\{C_3, C_2\}$ ,  $\{C_3, C_1\}$ ,  $\{C_2, C_1\}$ ,  $\{C_4\}$ ,  $\{C_3\}$ ,  $\{C_2\}$ ,  $\{C_1\}$   
 $C_1$ :  $\{s_4, s_3\}$ ,  $\{s_4, s_2\}$ ,  $\{s_4, s_1\}$ ,  $\{s_3, s_2\}$ ,  $\{s_3, s_1\}$ ,  $\{s_2, s_1\}$ ,  $\{s_4\}$ ,  $\{s_3\}$ ,  $\{s_2\}$ ,  $\{s_1\}$   
 $C_2$ :  $\{s_3, s_4\}$ ,  $\{s_3, s_1\}$ ,  $\{s_3, s_2\}$ ,  $\{s_4, s_1\}$ ,  $\{s_4, s_2\}$ ,  $\{s_1, s_2\}$ ,  $\{s_3\}$ ,  $\{s_4\}$ ,  $\{s_1\}$ ,  $\{s_2\}$   
 $C_3$ :  $\{s_2, s_1\}$ ,  $\{s_2, s_4\}$ ,  $\{s_2, s_3\}$ ,  $\{s_1, s_4\}$ ,  $\{s_1, s_3\}$ ,  $\{s_4, s_3\}$ ,  $\{s_2\}$ ,  $\{s_1\}$ ,  $\{s_4\}$ ,  $\{s_3\}$   
 $C_4$ :  $\{s_1, s_2\}$ ,  $\{s_1, s_3\}$ ,  $\{s_1, s_4\}$ ,  $\{s_2, s_3\}$ ,  $\{s_2, s_4\}$ ,  $\{s_3, s_4\}$ ,  $\{s_1\}$ ,  $\{s_2\}$ ,  $\{s_3\}$ ,  $\{s_4\}$ .

Each agent's preferences over individuals can be read from the last four (singleton) entries in his preference list  $P$  over sets. The matching  $\mu$  that matches each agent to his fourth-choice set of agents [emphasized in boldface in the preferences; i.e.,  $\mu(s_1) = \{C_2, C_3\}$ ,  $\mu(C_1) = \{s_3, s_2\}$ , etc.] is stable. To see this, note that for each  $s_i$ , all improvements on  $\mu(s_i)$  involve consultant  $C_i$ , but  $C_i$  is not interested in dealing with  $s_i$  alone, since  $s_i$  is  $C_i$ 's last-choice individual. However,  $\mu$  is not in the core and is not Pareto optimal, since it is dominated by the matching  $\mu'$  that gives each agent his third-choice set of agents.<sup>43</sup>

<sup>43</sup>In each case, the responsive preferences have been chosen so that an agent prefers to be matched with the set consisting of his first- and fourth-choice

To adapt the example to the requirement that students must obtain one medical and one surgical position, suppose  $C_1$  and  $C_3$  offer surgical positions, while  $C_2$  and  $C_4$  offer medical positions. Then, the student preferences given above can be modified by simply deleting from the preference lists the unacceptable sets  $\{C_2, C_4\}$  and  $\{C_1, C_3\}$ , which leaves the conclusion of the example unchanged. Similarly, by deleting the set  $\{s_1, s_2\}$  from the preferences of  $C_4$  (who might be a urological surgeon from Edinburgh), we obtain the kind of substitutable preferences that result when a gender quota is invoked.

Proposition 3 (and therefore the first part of Proposition 2) is proved next.

**PROPOSITION 3:** *In the many-to-many matching model in which all agents have substitutable preferences, the set of stable matchings is nonempty.*

The proof will use the following algorithm.

Step 1: Each  $C$  makes offers to every  $s$  in the set  $Q(C, 1) = \text{Ch}_C(S)$ .

...

Step  $k$ : (i) Each  $s$  rejects any offers received so far that are not in the set  $\text{Ch}_s(O(s, k-1))$ , where  $O(s, k-1)$  is the set of offers  $s$  has received in steps  $1, \dots, k-1$  and not yet rejected by the end of step  $k-1$ . (ii) Each  $C$  who has received at least one rejection in part (i) of step  $k$  and who now has (nonrejected) offers outstanding to the students in the set  $N(C, k)$ , and has been rejected in steps  $1, \dots, k$  by the students in the set  $R(C, k)$ , makes (or renews) offers to his most preferred set of students  $Q(C, k)$  in the collection of sets  $\{Q \subset S \setminus R(C, k) \text{ such that}$

$Q \text{ contains } N(C, k)\}$ . (Subtraction of sets is denoted by  $\setminus$ .) That is,  $C$ 's outstanding offers at the end of step  $k$  include all those issued at previous steps and not yet rejected and none of those that have already been rejected. (iii) The algorithm stops at any step  $k = T$  at which no rejections are issued, and the resulting matching is  $\mu$  such that  $\mu(C) = N(C, T)$  for each  $C$  in  $C$ .

The proof, which involves showing that the matching  $\mu$  resulting from the algorithm is stable (as defined in Section II) when preferences are substitutable, makes use of the following lemmas. The first says that the requirement that consultants keep open all offers that have not been rejected does not constrain them: at each step, they issue offers to their choice set from among those that have not yet rejected them. (The lemmas, like the algorithm, are stated as if preferences are strict so that there is always a unique choice set. This assumption is for notational convenience only; if it were relaxed, it would be necessary to denote a choice set, meaning one among the collection of most-preferred sets, instead of the choice set as below.)

**LEMMA 1:** *For each step  $k = 1, \dots, T$  and each  $C$  in  $C$ ,*

$$Q(C, k) = \text{Ch}_C(S \setminus R(C, k)).$$

**PROOF:**

For  $k = 1$ ,  $R(C, 1) = \emptyset$ , so  $Q(C, 1) = \text{Ch}(S \setminus R(C, 1))$ . Suppose now, inductively, that  $Q(C, j) = \text{Ch}_C(S \setminus R(C, j-1))$  for  $j = 1, \dots, k-1$ . Then, if  $s$  is in  $N(C, k)$ ,  $s$  is in  $Q(C, k-1) = \text{Ch}_C(S \setminus R(C, k-1))$ ; but  $\{S \setminus R(C, k-1)\}$  contains  $\{S \setminus R(C, k)\}$ , and  $s$  is an element of  $\{S \setminus R(C, k)\}$ , so substitutability implies that  $s$  is an element of  $\text{Ch}_C(S \setminus R(C, k))$  [i.e., that  $\text{Ch}_C(S \setminus R(C, k))$  contains  $N(C, k)$ ]. Therefore,  $Q(C, k) = \text{Ch}_C(S \setminus R(C, k))$ , which completes the proof by induction.

The next lemma says that, no matter what the sequence of offers a student receives, he

---

individuals rather than his second and third choices. This is not a consequence of responsiveness; it is easy to check that reversing this preference would still leave the preferences responsive.



never regrets having rejected earlier offers when he discovers what the later offers are.

**LEMMA 2:** *For each  $s$  in  $S$ , if the algorithm stops at step  $T$ ,*

$$\begin{aligned}\mu(s) &\equiv O(s, T-1) \\ &\equiv \text{Ch}_s(O(s, T-1)) \\ &= \text{Ch}_s\left(\bigcup_{k=2}^T O(s, k-1)\right).\end{aligned}$$

**PROOF:**

Suppose that  $C$  is an element of  $\text{Ch}_s(\bigcup_k O(s, k-1))$ , and let  $k$  be such that  $C$  is an element of  $O(s, k-1)$ . Then by substitutability,  $C$  is an element of  $\text{Ch}_s(O(s, k-1))$ . Therefore,  $s$  does not reject  $C$  at any step of the algorithm, so  $C$  is an element of  $O(s, T-1)$ , and  $\text{Ch}_s(\bigcup_k O(s, k-1))$  is a subset of  $\mu(s)$ . However,  $\mu(s) = \text{Ch}_s(O(s, T-1))$  cannot be strictly preferred to its subset  $\text{Ch}_s(\bigcup_k O(s, k-1))$ , and so when preferences are strict the two sets must be equal (and when preferences are not strict the choice sets can always be selected so that the two sets are equal).

The proof of Proposition 3 is now simple.

**PROOF OF PROPOSITION 3:**

The lemmas imply that  $\mu(s) = \text{Ch}_s(\mu(s))$  for all students  $s$ , and  $\mu(C) = \text{Ch}_C(\mu(C))$  for all consultants  $C$ . Now consider a student  $s$  and consultant  $C$ , not matched to one another at  $\mu$ , such that  $s$  is an element of  $\text{Ch}_C(\mu(C) \cup s)$ . Lemma 1 implies that, during the course of the algorithm,  $C$  made an offer to  $s$  and was rejected. Then, Lemma 2 implies that  $\text{Ch}_s(\mu(s) \cup C) = \mu(s)$ . Therefore,  $\mu$  is stable.

The above proof demonstrates that the set of stable matchings is nonempty in many-to-two matching with substitutable preferences. Proposition 6 states that preferences constrained to obey constraints 1 or 2 described in Section IV are substitutable,

while those constrained to obey 1 and 3 are not.

**PROOF OF PROPOSITION 6:**

First consider an agent  $A$  with quota  $q_A$  ( $= 2$  if  $A$  is a student) whose preferences are responsive except insofar as they satisfy either constraint 1 (if  $A$  is a student) or 2 (in which case  $A$  must be a consultant surgeon). That is,  $A$  has a rank-ordering over the agents in the opposite set, which can be represented as a set of the form  $B = B_1 \cup B_2$ , and  $A$  finds it unacceptable to be matched to any subset  $S$  of  $B$  that contains more than one member of  $B_2$  (and, if  $A$  is a student, more than one member of  $B_1$ ). Thus, for any subset  $T$  of  $B$ ,  $\text{Ch}_A(T)$  is a set consisting of the  $q_A$  highest-rank acceptable members of  $T$  (or all of them if there are fewer than  $q_A$ ), providing that these contain no more than one member of  $B_2$  (or  $B_1$ ), and otherwise it is a set consisting of the highest-ranked member of  $B_2 \cap T$  and the  $q_A - 1$  highest-ranked acceptable members of  $B_1 \cap T$  (or all of them if there are fewer than  $q_A - 1$ ). Now it is easy to check that, for any  $t$  in  $B_1$  or  $B_2$  and  $t' \neq t$  in  $B_1$  or  $B_2$ , if  $t$  is in  $\text{Ch}_A(T)$  then it is also in  $\text{Ch}_A(T - t')$ . This completes the proof of the proposition for the case of preferences satisfying constraints 1 or 2.

Now consider a student  $s$  whose preferences satisfy constraints 1 and 3 but are otherwise responsive to a rank-ordering over individual consultants, and let this rank ordering be  $C_1, C_2, C_3, C_4$ , where  $C_1$  and  $C_3$  are teaching-hospital positions in medicine and surgery, respectively, and where  $C_2$  and  $C_4$  are nonteaching hospital positions in medicine and surgery, respectively. Then, it is possible to have  $\text{Ch}_s(\{C_1, C_2, C_3, C_4\}) = \{C_1, C_4\}$  but  $\text{Ch}_s(\{C_2, C_3, C_4\}) = \{C_2, C_3\}$ , which does not contain  $C_4$ , so these preferences are not substitutable. This completes the proof of Proposition 6.

**PROOF OF PROPOSITION 7:**

Recall that each consultant in the deferred-acceptance algorithm as modified at Edinburgh submits a rank-ordering of individual students. Following a rejection in the

algorithm, a consultant extends an offer to the highest-ranked (eligible) student who has not yet received an offer from him, where only male students may be eligible if the consultant is a surgeon who has exercised his option to limit the number of female house officers. Each student has submitted separate lists of medical and surgical positions and rejects all but the best offer in each category.

It is immediate from the above description of the algorithm and from Proposition 6 that this conforms to the version of the algorithm used here to prove Proposition 3, if the consultants have responsive preferences that are lexicographic in their lower-ranked (more preferred) candidates (e.g., a consultant  $C$  whose rank-ordering is  $s_1, s_2, s_3, s_4, \dots$  must prefer  $\{s_1, s_4\}$  to  $\{s_2, s_3\}$  in order for the algorithm to correspond to that used to prove Proposition 3). Thus, if the consultants have such preferences, the matching is stable. However, suppose they have responsive preferences that are not lexicographic in this way (e.g., suppose the above consultant  $C$  prefers  $\{s_2, s_3\}$  to  $\{s_1, s_4\}$ ). If there were a student  $s$  such that the resulting matching was unstable with respect to  $(s, C)$ , then the matching would be unstable with respect to *any* preferences that are responsive to  $C$ 's rank-ordering of individual students. Since this cannot be, the resulting matching is stable.

The following example shows that the set of stable matchings may be empty when student preferences obey both constraints 1 and 3 from Section IV:

**Example 4.** Consider a market that includes four students applying to consultants  $MT_1$ ,  $ST_1$ ,  $ST_2$ ,  $SNT$ ,  $MNT$ , where  $M$  stands for a medical position,  $S$  for a surgical position,  $T$  for a teaching-hospital position and  $NT$  for a nonteaching-hospital position. For simplicity, each consultant seeks only one student. Each student seeks one medical and one surgical position and may not hold more than one teaching-hospital position. The preferences over individual students and positions are given by the following rank

orderings:

$s_1$ : $SNT, ST_2, MT_1$	$MT_1$ : $s_1, s_2$
$s_2$ : $MT_1, ST_1$	$ST_1$ : $s_2, s_3$
$s_3$ : $ST_1, SNT$	$ST_2$ : $s_1, s_4$
$s_4$ : $ST_2$	$SNT$ : $s_3, s_1, s_2$
	$MNT$ : $s_1, s_2$ .

To see that every matching  $\mu$  is unstable in this example, first suppose that  $\mu$  is stable and that  $\mu(SNT) \neq s_1$ . Stability implies  $\mu(SNT) = s_3$ , which implies  $\mu(ST_1) = s_2$ , which implies  $\mu(MT_1) = s_1$ , which implies that  $\mu$  is unstable with respect to  $(s_1, ST_2)$ , a contradiction. Thus, if  $\mu$  is stable,  $\mu(SNT) = s_1$ . However, this implies  $\mu(ST_1) = s_3$ , which implies  $\mu(MT_1) = s_2$ , which implies that  $\mu$  is unstable with respect to  $(s_1, MT_1)$ , which is the contradiction needed to prove that no matching is stable.

## REFERENCES

- Acheson, E. D., "Preregistration House-Officers" (letter), *The Lancet*, 14 April 1984 (Vol. I, No. 8381), 852.
- Alexander-Williams, J. and Stephenson, Ivor G., "Appointment of Preregistration House Officers," *British Medical Journal*, 9 June 1973 (Vol. 2, No. 5866), 605–6.
- Blair, Charles, "The Lattice Structure of the Set of Stable Matchings with Multiple Partners," *Mathematics of Operations Research*, November 1988, 13, 619–28.
- Clayden, A. D. and Parkhouse, James, "Allocation of Preregistration Posts," *British Journal of Medical Education*, March 1971, 5, 5–12.
- Doig, A. and Munday, G., "A Coordinated Scheme for the Allocation of Preregistration House-Officer Posts," *The Lancet*, 21 June 1969 (Vol. I, No. 7608), 1250–2.
- Gale, David and Shapley, Lloyd, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, January 1962, 69, 9–15.
- Gillard, J. H. and Dent, T. H. S., "The Allocation of House Officer Posts: A UK Sur-

- vey," *Medical Education*, July 1988, 22, 342-4.
- Kelso, Alexander J., Jr. and Crawford, Vincent P.**, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, November 1982, 50, 1483-1504.
- Leishman, A. G. and Ryan, R. P.**, "Appointment of Provisionally Registered House-Officers by Computer Match," *The Lancet*, 29 August 1970 (Vol. I, No. 7670), 459-61.
- McVitie, D. G. and Wilson, L. B.**, (1970a) "Stable Marriage Assignments for Unequal Sets," *BIT*, 1970, 10 (3), 295-309.
- \_\_\_\_\_, and \_\_\_\_\_, (1970b) "The Application of the Stable Marriage Assignment to University Admissions," *Operational Research Quarterly*, December 1970, 21, 425-33.
- \_\_\_\_\_, and \_\_\_\_\_, "The Stable Marriage Problem," *Communications of the ACM*, July 1971, 14, 486-92.
- Mongell, Susan and Roth, Alvin E.**, "Sorority Rush as a Two-Sided Matching Mechanism," *American Economic Review*, June 1991, 81, 441-64.
- Roth, Alvin E.**, "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, November 1982, 7, 617-28.
- \_\_\_\_\_, (1984a) "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, December 1984, 92, 991-1016.
- \_\_\_\_\_, (1984b) "Misrepresentation and Stability in the Marriage Problem," *Journal of Economic Theory*, December 1984, 34, 383-7.
- \_\_\_\_\_, (1985a) "The College Admissions Problem Is Not Equivalent to the Marriage Problem," *Journal of Economic Theory*, August 1985, 36, 277-88.
- \_\_\_\_\_, (1985b) "Common and Conflicting Interests in Two-Sided Matching Markets," *European Economic Review* (special issue: Market Competition, Conflict, and Collusion), February 1985, 27, 75-96.
- \_\_\_\_\_, "Two Sided Matching with Incomplete Information about Others' Preferences," *Games and Economic Behavior*, June 1989, 1, 191-209.
- \_\_\_\_\_, "New Physicians in the U.S. and the U.K.: A Natural Experiment in the Organization of Labor Markets," *Science*, 14 December 1990, 250, 1524-8.
- \_\_\_\_\_, and **Sotomayor, Marilda**, *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monograph Series, Cambridge: Cambridge University Press, 1990.
- \_\_\_\_\_, and **Vande Vate, John H.**, "Random Paths to Stability in Two-Sided Matching," *Econometrica*, November 1990, 58, 1475-80.
- \_\_\_\_\_, and \_\_\_\_\_, "Incentives in Two-Sided Matching with Random Stable Mechanisms," *Economic Theory*, January 1991, 1, 31-44.
- Shah, A. R. and Farrow, S. C.**, "Pre-registration House Appointments: A Computer Aided Allocation Scheme," *Medical Education*, November 1976, 10, 474-9.
- Townsend, H. R. A.**, "Pre-registration House Appointments: A Computer-Aided Allocation Scheme" (letter), *Medical Education*, March 1977, 11, 160-1.
- \_\_\_\_\_, *PRAMS.80 Manual*, Cirencester, U.K.: Special Medical Micro Software Ltd., 1981.
- Yellowlees, Henry**, "Difficulties in Obtaining Preregistration Posts" (letter), *The Lancet*, 26 November 1983 (Vol. II, No. 8361), 1254.
- Department of Health and Social Security (U.K.)**, *Report of the Pre-registration House Officers Working Group*, unpublished memorandum, 1987.