Game Theory as a Tool for Market Design

By Alvin E. Roth

Harvard University, Department of Economics and Graduate School of Business
Administration

Introduction:

Markets evolve, but they are also designed. Entrepreneurs and managers, legislators and
regulators, lawyers and judges, all get involved, at least indirectly, in market design.
Recently game theorists have also started to take a direct role in design. This is a natural
development, because game theory is the part of economics that deals with the "rules of
the game" that define market operations.

What makes the practical design of markets different from studying them conceptually is
that practical design carries with it a responsibility for detail. For this reason, the simple
models and analytical methods that characterize most game-theoretic analysis in the
economic literature need to be supplemented with tools to deal with complexity. This
complexity comes in two forms: complexity of the strategic environment itself, and
complexity of participants' behavior.

Design practice is in its early stages, and we are learning by doing, teaching ourselves as
we go. In my limited experience, some of which I will briefly describe below, a natural
set of tools with which to attack complex design problems consists of analytical game
theory, experimental economics, and computation. Laboratory experiments will help
clarify how people behave, both when faced with new markets, and as they gain
experience. Computational methods will help us explore strategic environments that may
be too complex to solve analytically.

Market design:

Market design concerns the creation of a venue for buyers and sellers, and a format for
transactions. (A market as a "pure venue" can be seen in perhaps its clearest form in
internet auctions, where some of the questions that arise about the location of a market
are almost purely conceptual.) Game theorists have taken the lead in designing a number
of different kinds of markets. Perhaps the three best known of these are auction markets
for radio spectrum licenses, spot markets for electric power, and labor market
clearinghouses (references: Chao and Wilson 1999, Cramton 1997, McAffee and
McMillan 1996, McMillan 1994, 1995, Milgrom 1998, 1999;, Roth and Peranson 1997,
1999, Wilson, 1998, 1999). The primary design problem has been different in each of
these.

In the case of radio spectrum auctions in the United States, the federal government used to give away licenses, but was ordered by Congress to sell them, both to raise revenue and to promote efficient use. The chief design question was how to design an auction procedure to promote price discovery of complicated, inter-related packages of spectrum rights in different regions. Because licenses may be more valuable in combination than separately, it was necessary to allow bidders to bid on bundles of their own choosing, in a way that allowed them to change packages in response to price changes during the auction. For this reason, multi-round auctions were adopted, and much of the design focus was on "activity rules" intended to promote efficient price discovery, by preventing bidders from concealing their interest in certain licenses and then bidding "at the last minute."

The spot market for electric power also came into being as a result of legislation, in this case to separate power transmission from power generation. Some of the chief design questions for electricity exchanges have to do with the physics of power transmission, including the negative externalities in the form of line congestion that each transaction brings to other users of the network. Thus electricity markets, in addition to mediating transactions between buyers and sellers, have to balance the transmission network geographically, and ensure the provision of adequate reserve capacity to increment and decrement the flows in response to network requirements (see Chao and Wilson, 1999).

My own experience in market design has been with entry-level professional labor markets. Since 1998, the vast majority of jobs for new physicians in the U.S. (about 20,000 per year) are filled by a clearinghouse whose design I directed (Roth and Peranson 1997, 1999). It updated and replaced an older clearinghouse, which in turn replaced a chaotic, decentralized market, which had suffered from severe coordination failures.

It turns out that these kinds of coordination failures are quite common in professional, entry-level labor markets. (And the newly designed clearinghouse has in fact been adopted by entry level labor markets in a number of other professions, since its adoption by American physicians.) The demand for the design of market clearinghouses occurs in markets in which strategic behavior has led to very inefficient matching of firms and workers. The principal design task is to create a centralized market clearinghouse that will compete effectively for participants with the inefficient alternative of decentralized bilateral contracting.

To fully understand the problems that clearinghouses are intended to solve, as well as the strategic pressures that sometimes cause them to fail, it is helpful to understand the history of these various markets, and the failures they experienced that made the clearinghouse form of organization an attractive alternative. That is beyond the scope of the present paper. But some of the history of the American medical market is given in Roth (1984), some comparisons to other medical markets are found in Roth (1990, 1991), and a variety of market failures in the entry level labor markets of other professions are discussed in Roth and Xing (1994, 1997).

We can see how design problems build on and extend traditional game theoretic models by looking first at a simple model of matching, and then considering how it must be extended to deal with the practical design problems involved in the medical labor market. Models of this kind have been the basis of a large theoretical literature (dating from Gale and Shapley, 1962) and a growing empirical literature. (Roth and Sotomayor 1990 survey much of this literature.)

A simple model of matching:

There are two disjoint sets of players, Firms = $\{f_1,..., f_n\}$, and Workers = $\{w_1,..., w_p\}$. Associated with each firm fi is the number of positions they have to offer, qi. Agents on each side of the market have (transitive and strict) preferences over agents on the other side, with whom they could be matched. These are denoted by ordered lists: e.g.

$\qquad$ $P(f_i) = w_3, w_2, ... , w_k, f_i ...$

is read as saying that firm f_i's first choice is w_3, it prefers w_3 to w_2 (denoted $w_3 P(f_i) w_2$), and so on, and that w_k is the least desirable worker it is willing to employ (after that it would rather leave a position vacant than fill it with a less desirable worker). Similarly, each worker w_j has preferences of the form $P(w_j) = f_2, f_4, ... w_j ...$

An outcome of the game is a *matching* x that matches firms and workers. Formally, x: F∪W → F∪W, such that x(f) = w iff x(w) = f, and for all f and w, |x(f)| is less than or equal to q_f, and either x(w) is in F or x(w) = w.

A matching x will be said to be *blocked by an individual* k if k prefers being unmatched to being matched with x(k), i.e. if kP(k) x(k). A matching x is *blocked by a pair of agents* (f,w) if they each prefer each other to their matches at x, i.e. if

$\qquad$ w P(f) w' for some w' in x(f) or w P(f) f if |x(f)| < q_f
and
$\qquad$ f P(w) x(w).

(The first line says that f prefers w to some w' it is matched with, or f prefers w to an empty position if it has one; the second line says that w prefers f to its match at x.)

A matching x is *stable* if it isn't blocked by any individual or pair of agents. That is, a matching is stable if it assigns no agent to an unacceptable match, and if no pair of agents not matched to each other would mutually prefer to be matched to one another.

Some empirical evidence:

Unstable matchings are hard to sustain, regardless of whether they are produced by decentralized negotiations or by centralized clearinghouses. If, for example, you are a worker holding an offer from your third choice firm, you have only to make two phone calls to find out if you are part of a blocking pair, i.e. to find out if one of the firms you would prefer to work for would prefer to employ you. As Table 1 shows, centralized market

clearinghouses that produce stable matchings have succeeded much more reliably than those that do not. (Table 1 is drawn both from unpublished notes and from Roth 1984, 1991, Mongell and Roth 1991, Roth and Xing 1994, and Bodin and Pankin, 1999.)

Table 1: <u>Stable and Unstable (Centralized) Clearinghouses</u>

Market	Stable?	Still in use (halted unraveling)?
American medical and other health professions labor markets:		
NRMP	yes	yes(new design adopted 1997)
Medical Specialties	yes	yes (about 30 markets)
Dental Residencies	yes	yes (6 markets)
Osteopaths (before '94)	no	no
Osteopaths (since '94)	yes	yes (2 markets)
Pharmacists	yes	yes
Clinical psychologists	yes (first used in '99)	yes
Regional medical markets in the U.K.		
Edinburgh ('69)	yes	yes
Cardiff	yes	yes
Cambridge	no	yes
London Hospital	no	yes
Birmingham	no	no
Edinburgh ('67)	no	no
Newcastle	no	no
Sheffield	no	no
Canadian Doctors (CaRMS)	yes	yes
Canadian Lawyers ("articling positions")		
Ontario	yes	yes
Alberta	yes	yes
British Columbia	yes	no (abandoned in '96)
Sororities	yes (at equilibrium)	yes
Reform rabbis	yes (first used in '97-98)	yes

As the Table makes clear, stability is an important feature of a centralized labor market clearinghouse. A lot more can be said about this than is evident from the Table, but I'll limit myself here to mentioning that the regional medical markets in the U.K. provide a nice natural experiment that lets us learn more about how unstable mechanisms fail, as well as how stable ones succeed. But even the best natural experiment does not permit a close examination of individual behavior (particularly when strategic behavior of various kinds may violate market rules and hence be conducted secretly). And even the best natural experiment allows for other possible interpretations—e.g. there are other differences between Edinburgh and Newcastle than the way they organized their medical markets. It is therefore very helpful to be able to also examine these phenomena under controlled conditions in the laboratory.

A laboratory experiment:

Kagel and Roth (1999) report such an experiment, in which a small, decentralized laboratory market was established. It experienced costly unraveling over time in response to strategic behavior by the participants. A centralized clearinghouse was then made available. The experimental treatment was that in one condition of the experiment the clearinghouse produced stable matches (in terms of participants' stated preferences—see below), using essentially the algorithm employed in Edinburgh and Cardiff, and in the other condition, the unstable algorithm found in Birmingham and Newcastle was used. The results of the laboratory market essentially followed those observed in the field markets, and thus add an element of confidence to the interpretation of the natural experiment observed in the U.K. The control available in the laboratory, which allows us to look at markets that differ *only* in their matching mechanism, confirms that the observed changes can be accounted for by the difference between the mechanisms. The experiment also reveals some of the complexities of participants' behavior during the transition from decentralized to centralized markets that helps explain the dynamics of how centralized clearinghouses succeed or fail.

Stable matching mechanisms:

The reason that stable mechanisms have been invented in so many markets is that stable matches can be produced by mimicking the "folk idea" of how a decentralized market operates. One version of this is stated below so as to make this clear. That is, although the algorithm described below is for a centralized clearinghouse that takes as inputs rank order lists from both firms and workers, the description is "as if" the participants were making offers and considering them, in order to make clear the analogy to a decentralized market. (Of course in a real decentralized market, we observe many varieties of strategic behavior that don't occur in the algorithm. In centralized clearinghouses, some of this strategic behavior takes place instead in the decision of what rank order lists to state. And both centralized and decentralized markets continue to promote strategic behavior at the interviewing stage. See Roth, 1991, and Roth and Xing 1994, 1997.) The algorithm described below is roughly in the form proposed by Gale and Shapley (1962).

Deferred Acceptance Algorithm, with Firms making offers:

Step 1 a. Each firm f (with q_f positions) offers them to its q_f 1st choice workers (i.e. to the top q_f workers on its rank order list).

b. Each worker rejects any unacceptable offers (i.e. offers from firms not on its rank order list) and, if more than one acceptable offer is received, "holds" the most preferred (i.e. the highest ranked).
.
.
.

Step k a. Any firm that had one or more offers rejected at step k-1 makes new offers to its most preferred acceptable workers (up to the number of rejections received).

b. Each worker holds her most preferred acceptable offer to date, and rejects the rest.

STOP: when no further offers are made, and match each worker to the firm (if any) whose offer she is holding.

It is easy to see that the resulting matching is stable, because no firm or worker is matched to an unacceptable counterpart, and if firm f, say, prefers to be matched to worker w, then firm f must have proposed to w and been rejected, so w doesn't prefer to be matched to f.

It is also easy to see that there is another, similar deferred acceptance algorithm, in which the roles of the firms and workers are reversed, i.e. in which the workers propose (or "apply" for positions) and each firm f holds its q_f most preferred applicants and rejects the rest. It can be shown that all the workers like the stable matching that results from the worker proposing algorithm at least as well as the matching that results from the firm-proposing algorithm, and that the firms similarly all do better when the firms propose.

However neither of these conclusions will hold when we extend the algorithm to deal with some of the complications found in the medical labor market. The NRMP has several kinds of complications (or "match variations"), two of which I will describe here: both cause pairs of positions to be linked as complements in the preferences of some applicants.

Complexities of the medical market:

The first of these complications involves applicants who need two positions. Of the approximately 20,000 applicants each year, about ten percent are attempting to match to a second year position, such that, if they succeed, they will also require a prerequisite first year position. These applicants submit two kinds of rank order lists. The first is their primary list,

which indicates their preferences over positions in the usual way. For each position on their primary list that requires a prerequisite first year position, they also submit a supplementary list ranking these positions.

A second, somewhat similar complication is that about five percent of the applicants are members of married couples who are seeking two residency positions in the same city. These couples submit a single rank order list, on which are ranked *pairs* of positions, one for each spouse. Thus a couple may indicate, for example, that its first choice is a particular pair of positions in Boston, and its second choice is another particular pair of positions in New York.

These complications matter for two related reasons:

1.   They may change the properties of the market; and
2.   The clearinghouse algorithm must be designed to accommodate them.

Regarding the first point, when I was asked in 1995 to redesign the clearinghouse for American physicians, I had to come to grips with the fact that the only results in the existing theory of matching that applied directly to that market were the counterexamples (see Roth and Sotomayor, 1990). The theorems were all of the form that certain things *always* happened in simple markets (or never happened), while the counterexamples all showed that when markets have complexities such as those described above, less regular behavior could *sometimes* occur. But the existing theory didn't give much guidance about *magnitudes*: how often different kinds of irregular behavior might occur, and just how irregular it might be. However questions about magnitudes quickly come to the fore in making design choices.

To fix ideas, below are three theorems about simple markets whose conclusions do not carry over to markets with the NRMP match variations described above.

In a simple matching market:

1. the set of stable matchings is always nonempty

2. the set of stable matchings always contains a "firm-optimal" stable matching that is produced by the firm-proposing deferred acceptance algorithm, and an "applicant optimal" stable matching that is produced by the applicant-proposing version of the algorithm.

3. When the applicant proposing algorithm is used (but not when the firm proposing algorithm is used) no applicant can possibly improve his match by submitting an ROL that is different from his true preferences.

Because these theorems about simple markets can have counterexamples in the complex medical market, we relied on computational explorations to see how close an approximation the simple theory provides for the complex market. We relied on

computation in three places, each of which will be briefly described below (for details see Roth and Peranson, 1999).

Computational experiments to assist in the algorithm design:

The process by which the applicant-proposing algorithm was designed is roughly as follows. First, a conceptual design was formulated and circulated for comment (Roth, 1996). (It was a multi-pass algorithm designed to deal with complications and resolve instabilities one at a time, and was based in part on the deferred acceptance algorithms described in Roth and Vande Vate 1990.) In order for this to be coded into a working algorithm, a number of choices had to be made, chiefly concerning the sequence in which operations would be conducted. Most of these decisions can be shown to have no effect on the outcome of simple matches, but could potentially effect the outcome when the NRMP match variations are present. Consequently, we (Elliott Peranson and I) performed computational experiments before making sequencing choices.

The results of these computational experiments on sequencing gave us our first indication that the set of stable matchings was going to turn out to be quite small. Different variations of the algorithm could indeed produce different stable matchings (in ways that would not be the case in a simple match), but it turns out that this mostly effected fewer than 2 applicants a year. That is, the differences in outcomes due to sequencing decisions in the algorithm were on the order of one in ten thousand applicants. Furthermore, the differences due to sequencing appeared to be unsystematic, and in no case did the algorithm fail to converge to a stable matching. Consequently we sequenced the algorithm in ways that cut down the number of times it cycled, and speeded convergence to a stable matching.

Computational explorations of the data to assess the effect of the new algorithm:

Once the design of the algorithm was completed, we turned to computational explorations of the rank order lists submitted in previous years. For the moment, I will describe the analysis as if these rank order lists represent the true preferences of the medical students and residency programs, and then explain in the next section why this is not an unreasonable interpretation, for the purposes to which we put it.

In a simple market, the size of the set of stable outcomes, as measured by how many firms and workers receive different matches at different stable outcomes, can be assessed by comparing the results of the firm-proposing and the worker-proposing deferred acceptance algorithms. When this was done on the data of previous markets, we found that only about 20 applicants a year would be affected by this change, i.e. only one in a thousand.

If this were a simple market, the small number of applicants whose matching changes when we switch from hospitals proposing to applicants proposing would imply that there was also little room for strategic behavior when it comes time to state rank order lists. But this doesn't follow automatically in the complex market. And the direct

computational experiment, in which each applicant's strategy choices are varied while holding all others constant, would be infeasible to conduct, given the large number of applicants in each year's data. However, for studying the strategic options available to applicants in this market, it proved possible to design efficient computational experiments that begin by truncating all applicants' rank order lists, and proceed by iteratively reducing the set of applicants whose lists are truncated, to determine upper bounds on the number of applicants who could potentially profit from manipulating their rank order lists. The number of applicants who could even potentially profit from such manipulations when the applicant-proposing algorithm is employed varied from 0 to 9 in the years examined.

Theoretical computation on a simple model, to study the effect of market size:

The computational explorations of the data from the complex market suggest that the set of stable matchings is surprisingly small, and that the opportunities for strategic manipulation by applicants are correspondingly small. (The same can be said for residency programs, although I have not discussed that here; see Roth and Peranson, 1999). But there might be other explanations for these computational results.

One hypothesis could be that the reason we find such small potential for strategic manipulation is that our data have been collected after such manipulation has already taken place. That is, perhaps there are substantial opportunities for strategic manipulation, but these have been exhausted by the time we look at the ROLs submitted to the match, because the participants have already behaved strategically in an optimal way. We turn to computational experiments to see if this is plausible.

The computational experiments involve randomly generated simple markets with n workers and n firms, each of whom seeks to fill a single position. The set of stable matchings will be small if the preferences are highly correlated; therefore we concentrate on uncorrelated preferences. In this case, if agents on one side of the market have (randomly generated) preferences over all the agents on the other side of the market, the size of the set of stable matchings (which equals the core of the market) grows large as the market grows large: over 90 percent of firms and workers get different assignments at different stable matches by the time n=1,000.

However an important feature of the real market is that no one can participate in 1,000 interviews, and therefore no worker of firm submits a rank order list over that many firms or workers. And when we fix some maximum length k for submitted preference lists, we find that the core shrinks rapidly as the number of firms and workers grow. Simple markets of comparable size (in which we know we are looking at the "true," unmanipulated preferences) exhibit the same size sets of stable matchings as we see in the NRMP, and in smaller specialty medical markets. Consequently the size of the set of stable matchings is a function of the market size and length of preference lists, and the estimates on artificial markets correspond closely to the small size observed in the data from the medical markets. This small core is a property of the market structure (and not of manipulated preference lists). In fact the small size of the set of stable matchings

implies that there is virtually no opportunity for applicants to profitably manipulate their preference lists, regardless of what stable matching mechanism is employed.

Some concluding comments:

If game theory is going to become as important a part of applied economics as it is a part of economic theory, we have some work ahead of us. For one thing, we'll have to develop tools to study complex environments.

Laboratory experiments will help inform us about how people behave, not only in environments too complex to analyze analytically, but also in simple environments (in which economists' customary assumptions about behavior may not always be such good approximations as we would hope). One fact that leaps out of both experimental and field data is that behavior in markets evolves, as people learn about the environment by interacting with it. And in new markets, what people learn initially is about the behavior of other inexperienced players, so that an analysis that focuses only on equilibrium behavior may miss the importance of a market's initial history. For market designers, the early history is of critical importance, since markets that fail early will not have the opportunity for players to learn about their possibly desirable equilibria.

Computational methods will help us analyze games that may be too complex to solve analytically. When game theory is used primarily as a conceptual tool, it is a great virtue to concentrate on very simple games (and even these can present formidable analytical difficulties.) When game theory is used to study empirical environments, simplicity of the models is a more nuanced virtue. But when game theory is used for designing working markets, there is no choice but to deal with the complexities that the market requires. As we have seen, computation can play different roles, from explorations of alternative design choices, to data exploration, to theoretical computation (i.e. from using computational experiments to test alternative designs, to directly exploring complex market data, to exploring related simple models in ways that nevertheless elude simple analytical solution).

In conclusion, it appears to me that the design of markets—and of other kinds of economic environments—presents both the biggest challenge and the biggest opportunity facing applied game theory at the brink of the twenty first century.

Bibliography:

Bodin, Lawrence and (Rabbi) Aaron Panken [1999], "High Tech for a Higher Authority: The Placement of Graduating Rabbis From Hebrew Union College-Jewish Institute of Religion," mimeo.

Chao, Hung-po and Robert Wilson, "Design of Wholesale Electricity Markets," 1999, in preparation.

Cramton, Peter "The FCC Auctions: An Early Assessment," Journal of Economics and Management Strategy 6 (3), Fall 1997, 431-497

Gale, David and Lloyd Shapley [1962], "College Admissions and the Stability of Marriage," American Mathematical Monthly, 69, pp9-15.

Kagel, John H. and A.E. Roth, "The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment," mimeo, 1999.

McAfee, R. Preston, and John McMillan, "Analyzing the Airwaves Auction," Journal of Economic Perspectives 10, Winter 1996, 159-176

McMillan, John "Why Auction the Spectrum?" Telecommunications Policy 19, April 1995, 191-199

McMillan, John "Selling Spectrum Rights," Journal of Economic Perspectives 8, Summer 1994, 145-162

Milgrom, Paul "Game Theory and the Spectrum Auctions," European Economic Review 42, 1998, 771-778.

Milgrom, Paul "Auction Theory for Privatization," in preparation.

Mongell, Susan J. and Alvin E. Roth [1991] "Sorority Rush as a Two- Sided Matching Mechanism," American Economic Review,   81, June, 441-464.

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

Roth, Alvin E. "New Physicians: A Natural Experiment in Market Organization," Science, 250, 1990, 1524-1528.

Roth, Alvin E, "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.," American Economic Review, 81, June 1991, 415-440.

Roth, Alvin E. [1996], "Interim Report #1: Evaluation of the current NRMP algorithm, and preliminary design of an applicant-processing algorithm," consultant's report, and http://www.economics.harvard.edu/~aroth/interim1.html

Roth, Alvin E. and Elliott Peranson, "The effects of the change in the NRMP matching algorithm," Journal of the American Medical Association, 278, 9, September 3, 1997, 729-732.

Roth, A.E. and E. Peranson, "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," American Economic Review, forthcoming.

Roth, Alvin E. and Marilda Sotomayor Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis, Econometric Society Monograph Series, Cambridge University Press, 1990.

Roth, Alvin E. and John H. Vande Vate [1990], "Random Paths to Stability in Two-Sided Matching," Econometrica, 58, 1990, 1475-1480.

Roth, A.E. and X. Xing "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," American Economic Review, 84, September, 1994, 992-1044.

 Roth, A.E. and X. Xing "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists," Journal of Political Economy, 105, April 1997, 284-329.

Wilson, Robert   "Design Principles," chapter 11 in H. Chao and H. Huntington (eds.), Design of Restructured Power Markets, 1998. Norwell MA: Kluwer Academic Press.

Wilson, Robert   "Activity Rules for an Iterative Double Auction," chapter 10 in K. Chatterjee and W. Samuelson (eds.), Business Applications of Game Theory, 1999. Norwell MA: Kluwer Academic Press.