The Evolution of Core Stability in Decentralized Matching Markets

Heinrich H. Nax, Bary S. R. Pradelski & H. Peyton Young*

January 10, 2013

Abstract

Decentralized matching platforms on the internet allow large numbers of agents to interact anonymously at virtually no cost. Very little information is available to market participants and trade takes place at many different prices simultaneously. We propose a decentralized, completely uncoupled learning process in such environments that leads to stable and efficient outcomes. Agents on each side of the market make demands of potential partners and are matched if their demands are mutually profitable. Matched agents occasionally experiment with higher demands, while single agents lower their demands in the hope of attracting partners. This simple and intuitive learning process implements core allocations even though agents have no knowledge of other agents' strategies, payoffs, or the structure of the game, and there is no central authority with such knowledge either.

JEL classifications: C71, C73, C78, D83

Keywords: assignment games, cooperative games, core, evolutionary game theory, learning, matching markets

*We thank Itai Arieli for suggesting a number of improvements to an earlier draft, and are grateful to participants at the 23rd International Conference on Game Theory at Stony Brook University and the Paris Game Theory Seminar. This research was supported by the United States Air Force Office of Scientific Research Grant FA9550-09-1-0538. Bary Pradelski also acknowledges the support of the Oxford-Man Institute of Quantitative Finance.

1. *Introduction*

Electronic technology has created new forms of markets that involve large numbers of agents who interact in real time at virtually no cost. Interactions are driven by repeated online participation over extended periods of time without public announcements of bids, offers, or realized prices. Even after many encounters, agents may learn little or nothing about the preferences and past actions of other market participants. In this paper we propose a dynamic model that incorporates these features and explore its convergence and welfare properties. We see this as a first step towards developing a better understanding of how such markets operate, and how they might be more effectively designed.

We shall be particularly interested in bilateral markets where agents on each side of the market submit prices at which they are willing to be matched. Examples include online platforms for matching buyers and sellers of goods, for matching workers and firms, for matching hotels with clients, and for matching men and women.[1] Matching markets have traditionally been analyzed using game-theoretic methods (Gale & Shapley [1962], Shapley & Shubik [1972], Roth & Sotomayor [1990]). In much of this literature, however, it is assumed that agents submit *preference menus* to a central authority, which then employs a suitably designed algorithm to match them. The model we propose is different in character: agents make bids that are conditional on the characteristics of those with whom they wish to be matched, and a profitable (not necessarily optimal) set of matches is realized at each point in time. There is no presumption that agents or a central authority know anything about others' preferences, or that they can deduce such information from prior rounds. Instead, the agents, through trial-and-error, look for profitable matches and adjust their demands dependent on whether being matched or being single.

Rules of this type have a long history in the psychology literature (Thorndike [1898], Hoppe [1931], Estes [1950], Bush & Mosteller [1955], Herrnstein [1961]). To the best of our knowledge, however, such a framework has not previously been used in the study of matching markets in cooperative games.[2] The approach seems especially well-suited to modeling behavior in large decentralized matching markets, where agents have little information about the overall game and about the identity of the other market participants. We show that a class of learning rules with simple adjustment dynamics of this type implements the core with probability one after finite time. The main contribution of the paper is to show that this can be achieved even though agents have no knowledge of other agents' strategies or preferences, and there is no central authority with such knowledge either.

The paper is structured as follows. The next section discusses the related literature on matching and core implementation. Section 3 formally introduces assignment games and the concepts of bilateral stability and the core. Section 4 describes the process of adjustment and search by individual agents. In section 5 we prove that this process converges to the core. Section 6 concludes.

---

[1] Examples include www.priceline.com's Name-Your-Own-Price®, the first reverse auction on the internet allowing buyers to post bids for products such as hotel rooms or airplane tickets. www.HireMeNow.com's Name-Your-Own-Wage™ uses a similar mechanism for temporary employment.

[2] For a review of matching mechanisms in the literature see Sandholm [2008].

2. *Related literature*

There is a sizeable literature on matching algorithms that grows out of the seminal paper by Gale & Shapley [1962]. In this approach agents submit preferences for being matched with agents on the other side of the market, and a central clearing algorithm matches them in a way that yields a core outcome (provided that the reports are truthful). For subsequent literature, see Crawford & Knoer [1981], Kelso & Crawford [1982], Demange & Gale [1985], Demange, Gale & Sotomayor [1986], Shimer [2007], [2008], Elliott [2010], [2011].[3] These algorithms have been successfully applied in situations where agents engage in a formal application process, such as students seeking admission to universities, doctors applying for hospital residencies, or transplant patients looking for organ donors.[4]

In the present paper, by contrast, we consider situations where the market is fluid and decentralized. Agents are matched and rematched over time, and the information they submit takes the form of prices rather than preferences. Examples include markets matching buyers with sellers or firms with workers. These constitute a special class of cooperative games with transferable utility (Shapley & Shubik [1972]). We shall show that even when agents have minimal amounts of information and use very simple price adjustment rules, the market evolves towards core outcomes.

There is a clearing mechanism, "the Matchmaker", whose function is to match agents with mutually profitable bids and offers. Neither the players nor the Matchmaker have enough information to optimize the value of the matches. This limited role is what distinguishes our Matchmaker from a central authority governing a traditional matching environment as in, for example, the National Resident Matching Program (Roth & Peranson [1999]). We shall show that simple adjustment rules by the agents lead to efficient and stable outcomes without any centralized information about which matches are best.

This result fits into a growing literature showing how cooperative game solutions can be understood as outcomes of a dynamic learning process (Agastya [1997], [1999]; Arnold & Schwalbe [2002]; Rozen [2010a], [2010b]; Newton [2010], [2012]; Sawa [2011]). To illustrate the differences between these approaches and ours, we shall briefly outline Newton's model here; the others are similar in spirit.[5] In each period a player is activated at random and demands a share of the surplus from some targeted coalition of players. He chooses a demand that amounts to a best reply to the expected demands of the others in the coalition, where his expectations are based on a random sample of the other players' past demands. In fact he chooses a best reply with probability close to one, but with small probability he may make some other demand. This noisy best response process leads to a Markov chain whose ergodic distribution can be characterized using the theory of large deviations. Newton shows that, subject to various regularity conditions, this process converges to a core allocation provided the game has a nonempty interior core.

---

[3] Shimer [2007], [2008] and Elliott [2010], [2011] explore empirical and network elements of matching.

[4] See Roth [1984], Roth & Peranson [1999] for discussions of the US medical resident market, and Roth, Sönmez & Ünver [2005] for the kidney exchange market.

[5] Newton's [2012] model nests the models of Agastya [1997], [1999] and Rozen [2010a], [2010b] as special cases.

The main difference between existing learning models and ours is the amount of information available to market participants.[6] The approach we take here requires considerably less information on the part of the agents: players know nothing about the other players' current or past behavior, or their payoffs. Thus, they have no basis on which to best respond to the other players' strategies; they simply experiment to see whether they might be able to do better. Adaptive rules of this type are said to be *completely uncoupled* (Foster & Young [2006]).[7] In recent years it has been shown that there are families of such rules that lead to equilibrium behavior in generic non-cooperative games (Karandikar, Mookherjee, Ray & Vega-Redondo [1998], Foster & Young [2006], Germano & Lugosi [2007], Marden, Young, Arslan & Shamma [2009], Young [2009], Pradelski & Young [2012]). Here we shall demonstrate that a very simple rule of this form leads to stability and optimality in two-sided matching markets.

3. Matching markets with transferable utility

In this section we shall introduce the conceptual framework for analyzing matching markets with transferable utility; in the next section we introduce the learning process itself. The population $N = \{F \cup W\}$ consists of firms $F = \{f_1, ..., f_m\}$ and workers $W = \{w_1, ..., w_n\}$.[8] They interact by submitting bids and offers to "the Matchmaker", whose function is to propose matches between firms and workers whose bids and offers are mutually profitable.

3.1 Static components

Willingness to pay. Each firm i has a *willingness to pay*, $p^+_{ij} \geq 0$, for being matched to worker j.

Willingness to accept. Each worker j has a *willingness to accept*, $q^-_{ij} \geq 0$, for being matched with firm i.

We assume that these numbers are specific to the agents and are not known to the other market participants or to the Matchmaker.

Match value. Assume that utility is linear and separable in money. The *value* of a match $(i, j) \in F \times W$ is the potential surplus

$$\alpha_{ij} = (p^+_{ij} - q^-_{ij})_+. \tag{1}$$

It will be convenient to assume that all values p^+_{ij}, q^-_{ij} and α_{ij} can be expressed as multiples of some minimal unit of currency δ, e.g., "dollars".

[6] Moreover, the core of an assignment game typically has an empty interior, so that the aforementioned results cannot be applied directly to the present set-up.

[7] This definition is a strengthening of *uncoupled* rules introduced by Hart & Mas-Colell [2003].

[8] More generally, the two sides of the market could represent buyers and sellers, or men and women in a (monetized) marriage market.

3.2 Dynamic components

Let $t = 0, 1, 2, ...$ be the time periods.

Assignment. For all agents $(i, j) \in F \times W$, let $a_{ij}^t \in \{0, 1\}$.

$$\text{If } (i,j) \text{ is} \begin{cases} \textit{matched} & \text{then} \quad a_{ij}^t = 1, \\ \textit{unmatched} & \text{then} \quad a_{ij}^t = 0. \end{cases} \tag{2}$$

If for a given agent $i \in N$ there exists j such that $a_{ij}^t = 1$ we shall refer to that agent as *matched*; otherwise i is *single*.

Bids. In period t each agent submits conditional bids to the Matchmaker. Agent $i \in F$ submits a vector of bids $b_i^t = (p_{i1}^t, ..., p_{in}^t)$, where p_{ij}^t is the amount i would currently pay if matched with $j \in W$. Similarly, agent $j \in W$ submits a vector of bids $b_j^t = (q_{1j}^t, ..., q_{mj}^t)$, where q_{ij}^t is the amount j would currently accept if matched with $i \in F$. We assume that the bids are multiples of δ and that they are *individually rational*, that is,

$$\text{for all } (i,j), \qquad p_{ij}^t \leq p_{ij}^+ \qquad \text{and} \qquad q_{ij}^t \geq q_{ij}^-. \tag{3}$$

Tie-breaking. A firm (worker) prefers to be matched at p_{ij}^+ (q_{ij}^-) rather than being single.

We shall assume that each bid by a given agent would yield the same payoff if realized at the bid price. In other words, an agent evaluates the expected benefits from each possible match and equalizes them across potential partners.[9] This means that each agent's bid vector has the following simple structure:

Demands. There exist *demands* $d_i^t, d_j^t \geq 0$, such that

$$\begin{aligned} &\text{for every } i, \quad d_i^t = p_{ij}^+ - p_{ij}^t \quad \text{for all } j, \\ &\text{for every } j, \quad d_j^t = q_{ij}^t - q_{ij}^- \quad \text{for all } i. \end{aligned} \tag{4}$$

For brevity let $\mathbf{d}^t = (d_1^t, d_2^t, ..., d_{m+n}^t)$ represent the players' demands in period t. We may also interpret them as *aspiration levels*. Note that a demand vector allows an agent to differentiate his *bids* according to his subjective valuation of different possible matches. In practice, an agent may lump the options into broad "quality" categories that he treats as equivalent. For example, an agent looking for a hotel room might bid the same amount for all one-star hotels, another amount for all two-star hotels, and so forth.

Profitability. A pair of bids (p_{ij}^t, q_{ij}^t) is *profitable* if $p_{ij}^t > q_{ij}^t$, or if $p_{ij}^t = q_{ij}^t$ and both i and j are single. Given (1) and (4), it is equivalent to say that a pair of demands (d_i^t, d_j^t) is *profitable* if $d_i^t + d_j^t < \alpha_{ij}$, and/or $d_i^t + d_j^t = \alpha_{ij}$ and both i and j are single.

Matchmaker. At each moment in time, at most one player is "active". The Matchmaker observes

[9] A given agent has no knowledge of the other agents' characteristics or the likelihood that he will be matched with them, hence he has no basis on which to choose his bids strategically.

- the current bids and which agent is currently active,
- who is currently matched with whom and which bids are profitable.

The Matchmaker then proposes a profitable match, if one exists, to the active agent, who accepts with positive probability. (Details about the Matchmaker and about how players are activated are specified in the next section.)

Prices. When i is matched with j given bids $p_{ij}^t \geq q_{ij}^t$, the resulting *price*, P_{ij}^t, is the average of the players' bids (subject to "rounding"). Namely, there is an integer k such that

$$\begin{array}{lll} \text{if } p_{ij}^t + q_{ij}^t = 2k\delta & \text{then} & P_{ij}^t = k\delta, \\ \text{if } p_{ij}^t + q_{ij}^t = (2k+1)\delta & \text{then} & \begin{cases} P_{ij}^t = k\delta & \text{with probability } 0.5, \\ P_{ij}^t = (k+1)\delta & \text{with probability } 0.5. \end{cases} \end{array} \tag{5}$$

This implies that when a pair is matched we have

$$d_i^t + d_j^t = \alpha_{ij}. \tag{6}$$

Note that when a new match forms that is profitable (as defined earlier), neither of the agents is worse off, and if one agent was previously matched both agents are better off in expectation.

3.3 Assignment games

We are now in a position to define matching markets and assignment games more formally.

Matching market. The *matching market* is described by $[F, W, \boldsymbol{\alpha}, \mathbf{A}]$:

- $F = \{f_1, ..., f_m\}$ is a set of m firms (or men or sellers),
- $W = \{w_1, ..., w_n\}$ is a set of n workers (or women or buyers),
- $\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \alpha_{ij} & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$ is the matrix of match values,

$\boldsymbol{\alpha}$ specifies the total surplus generated by every possible match.

- $\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ is a particular assignment matrix with

$$a_{ij} = 0 \text{ or } 1 \ \sum_{j \in W} a_{ij} \leq 1, \text{ and } \sum_{i \in F} a_{ij} \leq 1 \text{ for all } (i, j) \in F \times W.$$

The set of all possible assignments is denoted by $\mathcal{A}$.

Cooperative assignment game. Given $[F, W, \boldsymbol{\alpha}]$, the *cooperative assignment game* $G(v, N)$ is defined as follows. Let $N = F \cup W$ and define $v : S \subseteq N \to \mathbb{R}$ such that

- $v(i) = v(\emptyset) = 0$ for all singletons $i \in N$,
- $v(S) = \alpha_{ij}$ for all $S = (i, j)$ such that $i \in F$ and $j \in W$,
- $v(S) = \max\{v(i_1, j_1) + ... + v(i_k, j_k)\}$ for every $S \subseteq N$,

where the maximum is taken over all sets $\{(i_1, j_1), ..., (i_k, j_k)\}$ consisting of disjoint pairs that can be formed by matching firms and workers in S. The number $v(N)$ specifies the value of an optimal assignment.

Optimality.
An assignment $\mathbf{A}$ is *optimal* if $\sum_{(i,j)\in F\times W} \alpha_{ij} a_{ij} = v(N)$.

Outcome. An *outcome* $[\mathbf{A}^t, \mathbf{d}^t]$ is an assignment $\mathbf{A} \in \mathcal{A}$ together with a demand vector $\mathbf{d}^t \in \mathbf{R}_+^{m+n}$.

Pairwise stability.
$[\mathbf{A}^t, \mathbf{d}^t]$ is *pairwise stable* if for all (i, j) with $a_{ij} = 1$, $d_i^t + d_j^t = \alpha_{ij}$, and $d_{i'}^t + d_j^t \geq \alpha_{i'j}$ for every alternative firm i' and $d_i^t + d_{j'}^t \geq \alpha_{ij'}$ for every alternative worker j'.

The Core.
The *core* of an assignment game, $G(v, N)$, consists of the set $\mathbf{C}$ of all outcomes, $[\mathbf{A}, \mathbf{d}]$, such that $\mathbf{A}$ is an optimal assignment and $\mathbf{d}$ is pairwise stable.

Shapley & Shubik [1972] show that the core of any assignment game is always non-empty and coincides with the set of pairwise stable allocations that are supported by optimal assignments. Subsequent literature has investigated the structure of the assignment game core, which turns out to be very rich.[10]

Payoffs. Given $[\mathbf{A}, \mathbf{d}]$ the *payoffs* to the players are

$$\tilde{d}_i = \begin{cases} d_i & \text{if } i \text{ is matched,} \\ 0 & \text{if } i \text{ is single.} \end{cases} \tag{7}$$

In our framework, $[\mathbf{A}, \mathbf{d}]$ is in the core if all $a_{ij} = 0$ or 1, all $d_i \geq 0$ and the following conditions hold for the payoffs $\tilde{d}_i$:[11]

(i) $\forall i \in F$, $\sum_{j\in W} a_{ij} \leq 1$ and $\forall j \in W$, $\sum_{i\in F} a_{ij} \leq 1$,

(ii) $\forall i, j \in F \times W$, $\tilde{d}_i + \tilde{d}_j \geq \alpha_{ij}$,

(iii) $\forall i \in F$, $\sum_{j\in W} a_{ij} < 1 \Rightarrow \tilde{d}_i = 0$ and $\forall j \in W$, $\sum_{i\in F} a_{ij} < 1 \Rightarrow \tilde{d}_j = 0$.

(iv) $\forall i, j \in F \times W$, $a_{ij} = 1 \Rightarrow \tilde{d}_i + \tilde{d}_j = \alpha_{ij}$.

[10] See, for example, Roth & Sotomayor [1992], Balinski & Gale [1987], Sotomayor [2003].

[11] These are the feasibility and complementary slackness conditions for the associated linear program and its dual.

4. Evolving play

A fixed population of agents, $N = F \cup W$, repeatedly plays the assignment game $G(v, N)$ by submitting nonnegative demands to the Matchmaker and by adjusting them dynamically as the game evolves. (Recall that demands translate directly into bids and vice versa.) Agents become activated spontaneously according to independent Poisson arrival processes. For simplicity we shall assume that the arrival rates are the same for all agents, but our results also hold when the rates differ across agents (for example, single agents might become active at a faster rate than matched agents). The distinct times at which one agent becomes active will be called *periods*.

States. At the end of period t, the *state* Z^t consists of a pair $[\mathbf{A}^t, \mathbf{d}^t]$ where the assignment is $\mathbf{A}^t = \{a^t_{ij}\}_{F\times W}$, the demand vector is $\mathbf{d}^t = \{d^t_1, ..., d^t_{m+n}\}$, and by (9)

$$\text{if } \ a^t_{ij} = 1 \ \text{ then } \ d^t_i + d^t_j = \alpha_{ij}.$$

Denote the set of all states by Ω.

4.1. Behavioral dynamics

The essential steps and features of the learning process are as follows. At the start of period $t + 1$:

1. A unique agent becomes active.

2a. If a profitable match exists, the Matchmaker proposes a randomly drawn profitable match to the active agent. This match is accepted (and formed) with probability p.

2b. If no profitable match exists, the Matchmaker rejects the demand.

3a. If a new match (i, j) is formed, the price is the average of the two bids (subject to rounding). The demands of i and j next period are their realized payoffs this period.

3b. If no new match is formed, the active agent, if he was previously matched, keeps his previous demand and stays with his previous partner. If he was previously single, he remains single and lowers his demand by δ.

We shall now describe the process in more detail, distinguishing the cases where the active agent is currently *matched* or *single*. Let Z^t be the state at the end of period t (and the beginning of period $t + 1$), and let i be the unique *active* agent.

I. The active agent is currently matched

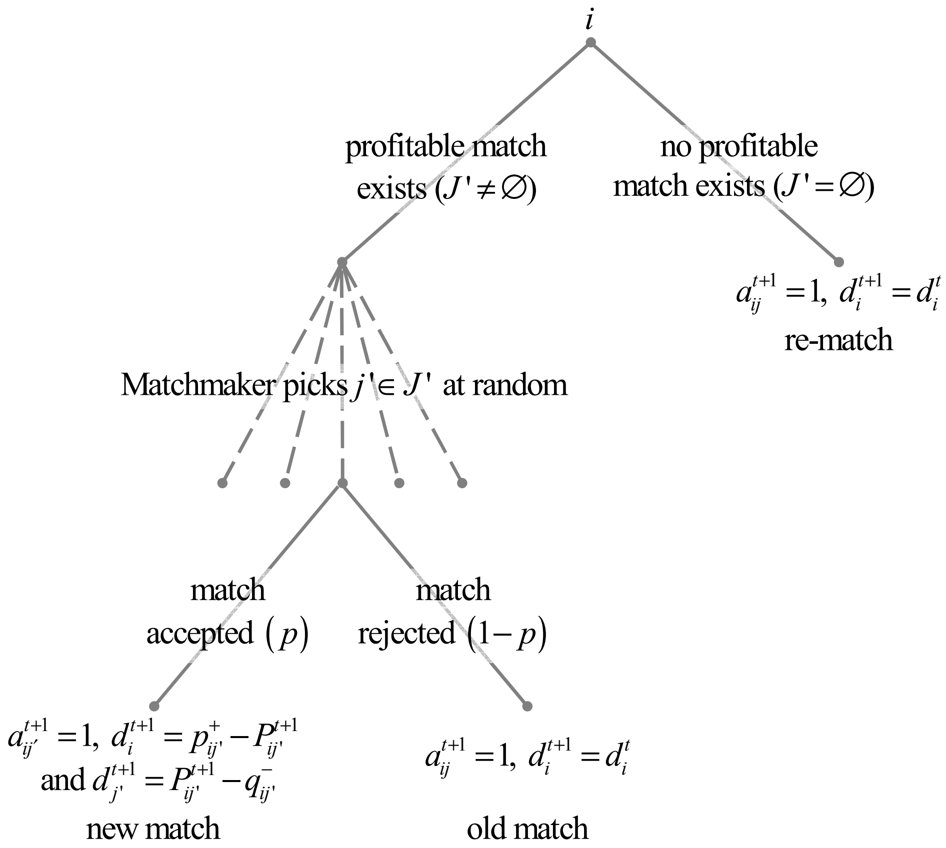
Let J' be the set of players with whom i can be profitably matched, that is,

$$J' = \{j' : p^t_{ij'} > q^t_{ij'}\}. \tag{8}$$

These are the only matches in which i will be better off in expectation. If $J' \neq \emptyset$, some agent $j' \in J'$ is drawn uniformly at random by the Matchmaker, and j' accepts the proposed match with probability $p \in (0, 1]$.[12] As a result, i's former partner is now single (and so is j''s former partner if j' was matched in period t). The price governing the new match, $P^{t+1}_{ij'}$, is the average (subject to rounding) of $p^t_{ij'}$ and $q^t_{ij'}$.

At the end of period $t + 1$, the demands (and the corresponding bids) of the newly matched pair are adjusted to equal their newly realized payoffs, while all other demands and matches remain fixed. If $J' = \emptyset$ or the proposed match is not accepted (which occurs with probability $1 - p$), i remains matched with his previous partner and keeps his previous demand. See Figure 1 for an illustration.

Figure 1: Transition diagram for active, matched agent (period $t + 1$).



[12]Instead of a uniform random draw from the profitable matches, priority could be given to those involving single agents; indeed any distribution with full support on the profitable matches can be used.

II. The active agent is currently single

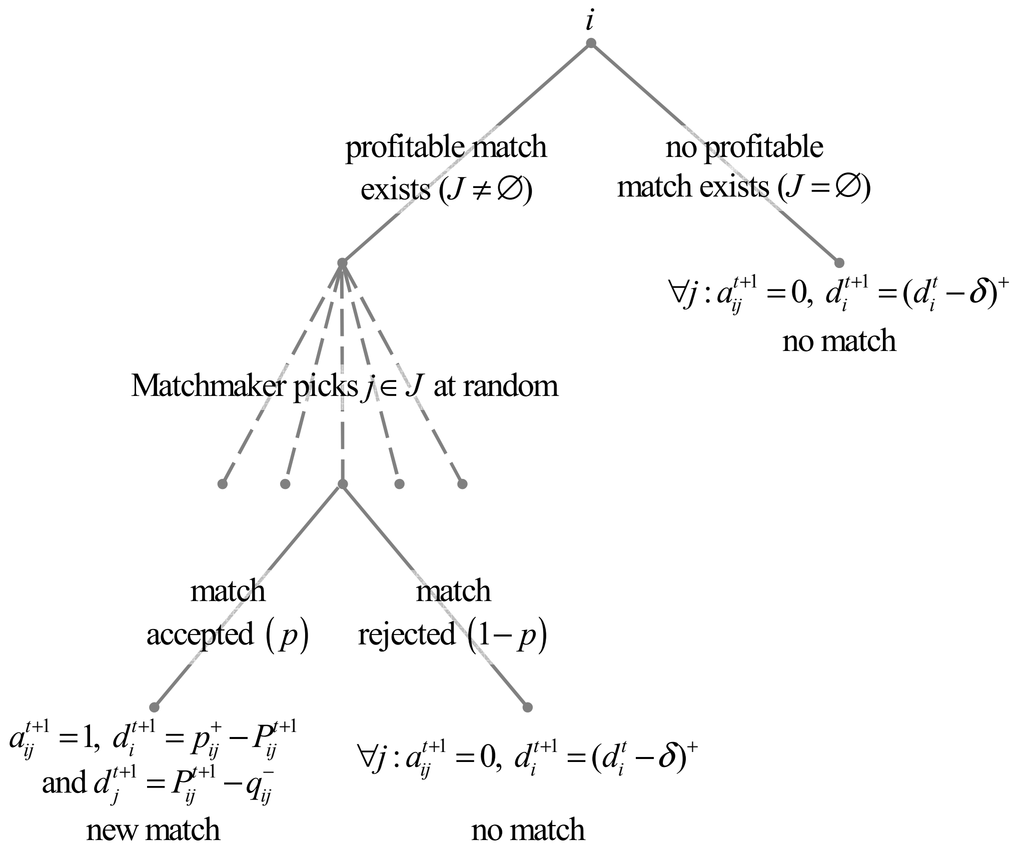
Let J be the set of players with whom i can be profitably matched, that is,

$$J = \{j : j \text{ single}, p_{ij}^t \geq q_{ij}^t\} \cup \{j : j \text{ matched and } p_{ij}^t > q_{ij}^t\}. \quad (9)$$

If $J \neq \emptyset$, some agent $j \in J$ is drawn uniformly at random by the Matchmaker and j accepts the proposed match with probability $p \in (0, 1]$. If j was matched in period t his former partner is now single. The price governing the new match, P_{ij}^{t+1}, is the average (subject to rounding) of p_{ij}^t and q_{ij}^t.

At the end of period $t + 1$, the demands (and the corresponding bids) of the newly matched pair are adjusted to equal their newly realized payoffs, while all other demands and matches remain fixed. If $J = \emptyset$ or the proposed match is not accepted (which occurs with probability $1 - p$), i remains single and reduces his demand (and the corresponding bid vector) to $d_i^{t+1} = d_i^t - \delta$. A special case arises when $d_i^t = 0$, in which case we assume that $d_i^{t+1} = 0$. See Figure 2 for an illustration.

Figure 2: Transition diagram for active, single agent (period $t + 1$).

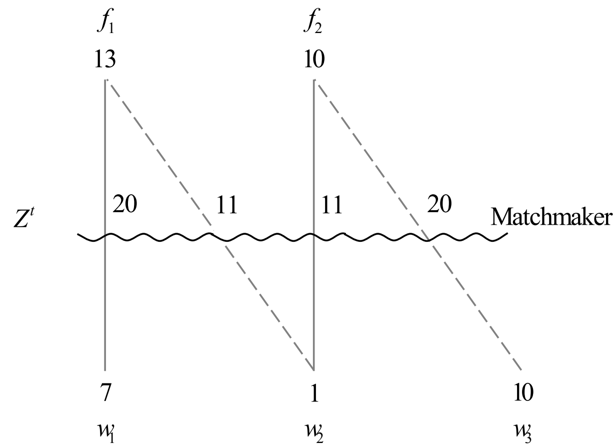


4.2. *Example*

Let $N = F \cup W = \{f_1, f_2\} \cup \{w_1, w_2, w_3\}$, $\alpha_{11} = \alpha_{23} = 2$, $\alpha_{12} = \alpha_{22} = 1$ and $\alpha_{ij} = 0$ for all other pairs (i, j). Let $\delta = 1$.

period t: *Current state*

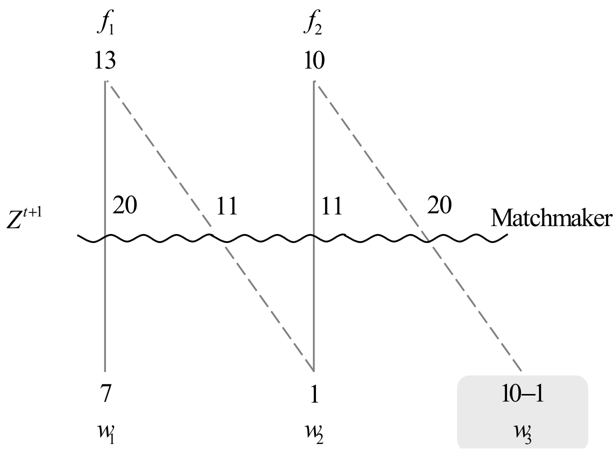
Suppose that, in some period t, (f_1, w_1) and (f_2, w_2) are matched and w_3 is single. In the illustrations below, the current demand of each agent is shown next to the name of that agent, and the values α_{ij} are shown next to the edges. Solid edges indicate matched pairs, and dashed edges indicate unmatched pairs. (Edges with value zero are not shown.) The wavy line indicates that no player can see the bids or the status of the players on the other side of the market.



Note that the demands satisfy $d_i^t + d_j^t \geq \alpha_{ij}$ for all i and j, but the assignment is not optimal (firm 2 should match with worker 3).

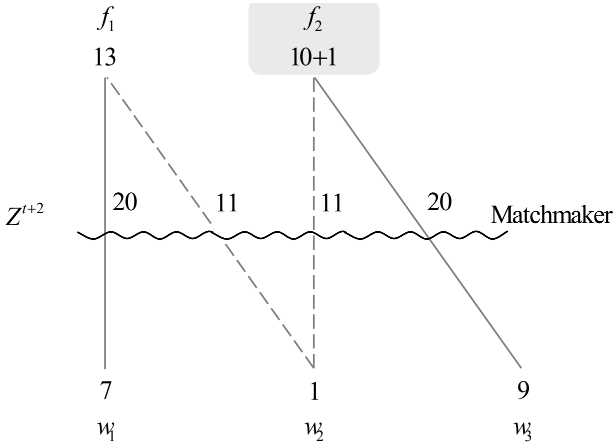
period $t + 1$: *Activation of single agent w_3*

w_3's current demand is too high in the sense that he has no profitable matches. Hence he remains single and reduces his demand by 1.



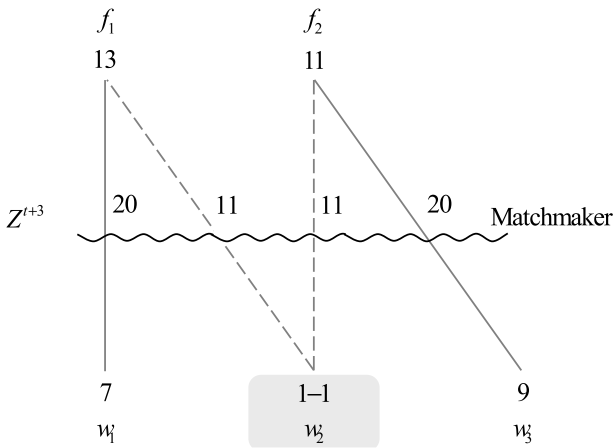
period $t + 2$**:** *Activation of matched agent* f_2

f_2's only profitable match is with w_3. With probability p the match forms. With probability 0.5 the price is set such that f_2 receives one unit more (11) and w_3 receives 9, while with probability 0.5 f_2 receives 10 and w_3 receives 10. (Thus in expectation the active agent f_2 gets a higher payoff than before.)



period $t + 3$**:** *Activation of single agent* w_2

w_2's current demand is too high in the sense that he has no profitable matches. Hence he remains single and reduces his demand by 1.



The resulting state is in the core.[13]

[13]Note that the states Z^{t+2} and Z^{t+3} are both in the core, but Z^{t+3} is absorbing whereas Z^{t+2} is not.

5. *Core stability*

Recall that a state Z^t is defined by an assignment $\mathbf{A}^t$ and demands $\mathbf{d}^t$ that jointly determine the payoffs. Recall that an outcome $[\mathbf{A}^t, \mathbf{d}^t]$ is in the core, $\mathbf{C}$, if conditions (i)-(iv) are satisfied.

Theorem 1. *Given an assignment game $G(v, N)$, from any initial state $[\mathbf{A}^0, \mathbf{d}^0]$, the process is absorbed into the core in finite time with probability 1.*

Throughout the proof we shall omit the time superscript since the process is time-independent. We shall proceed by establishing the following two claims.

Claim 1. There is a positive probability path to demands $\mathbf{d}$ such that $d_i + d_j \geq \alpha_{ij}$ for all i, j and such that, for every i, either there exists a j such that $d_i + d_j = \alpha_{ij}$ or else $d_i = 0$.

Any demands satisfying Claim 1 will be called *good*.

Claim 2. Starting at any state with good demands, there is a positive probability path to a pair $(\mathbf{A}, \mathbf{d})$ where $\mathbf{d}$ is good, $\mathbf{A}$ is optimal, and all singles demand zero.[14]

Proof of Claim 1.

Case 1. Suppose the demands $\mathbf{d}$ are such that $d_i + d_j < \alpha_{ij}$ for some i, j.

Case 1a. i and j are not matched with each other.

With positive probability, either i or j is activated and i and j become matched. The new demands are set equal to the new payoffs. Thus the sum of the demands is equal to the match's value α_{ij}.

Case 1b. i and j are matched with each other.

In this case, $d_i + d_j = \alpha_{ij}$ because whenever two players are matched the entire surplus is allocated.

Therefore, there is a positive probability path along which $\mathbf{d}$ increases monotonically until $d_i + d_j \geq \alpha_{ij}$ for all i, j.

Case 2. Suppose the demands $\mathbf{d}$ are such that $d_i + d_j \geq \alpha_{ij}$ for all i, j.

We can suppose that there exists a single agent i with $d_i > 0$ and $d_i + d_j > \alpha_{ij}$ for all j, else we are done. With positive probability, i is activated. Since no profitable match exists, he lowers his demand by δ. In this manner, a suitable path can be constructed along which $\mathbf{d}$ decreases monotonically until the demands are good. Note that at the end of such a path, the assignment does not need to be optimal and not every agent

[14] Note that this claim describes an absorbing state in the core. It may well be that the core is reached while a single demands more than zero. The latter state, however, is transient and will converge to the corresponding absorbing state.

with a positive demand needs to be matched. (See the period-t example in the preceding section.)

□

Proof of Claim 2.

Suppose that the state $(\mathbf{A}, \mathbf{d})$ satisfies Claim 1 and that some single exists who demands a positive amount. (If no such single exists and $\mathbf{d}$ is good, the assignment is optimal and we have reached a core state.) We show that there exists $m > 0$ and a probability $q > 0$ such that within m periods the probability is at least q that one of the following holds:

The demands are good, the number of single agents demanding a positive amount decreases, and the sum of the demands remains constant. (10)

The demands are good, the sum of the demands decreases by $\delta > 0$, and the number of single agents demanding a positive amount does not increase. (11)

In general, say an edge is *tight* if $d_i + d_j = \alpha_{ij}$ and *loose* if $d_i + d_j = \alpha_{ij} - \delta$. Define a *maximal alternating path* P to be a maximal-length path that starts at a single firm with positive demand, and that alternates between unmatched tight edges and matched tight edges. Note that, for every single demanding a positive amount, at least one maximal alternating path exists. Figure 3 (left panel) illustrates a maximal alternating path starting at f_1. Unmatched tight edges are indicated by dashed lines, matched tight edges by solid lines and loose edges by dotted lines.

Without loss of generality, let f_1 be a single firm demanding a positive amount.

Case 1. Starting at f_1, there exists a maximal alternating path P of odd length.

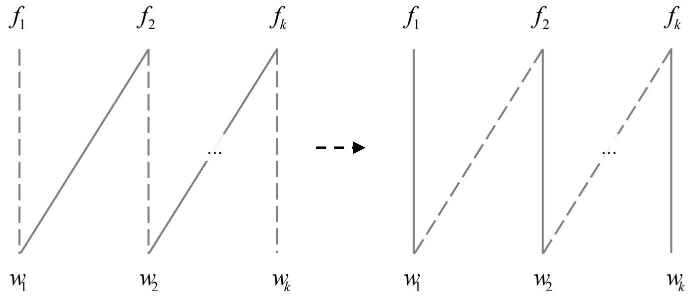
Case 1a. All firms on the path demand a positive amount.

We shall demonstrate a sequence of adjustments leading to a state as in (10).

Let $P = (f_1, w_1, f_2, w_2, ..., w_{k-1}, f_k, w_k)$. Note that, since the path is maximal and of odd length, w_k must be single. With positive probability, f_1 is activated. Since no profitable match exists, he lowers his demand by δ. With positive probability, f_1 is activated again next period, he snags w_1 and with probability 0.5 he receives the residual δ. At this point the demands are unchanged but f_2 is now single. With positive probability, f_2 is activated. Since no profitable match exists, he lowers his demand by δ. With positive probability, f_2 is activated again next period, he snags w_2 and with probability 0.5 he receives the residual δ. Within a finite number of periods a state is reached where all players on P are matched and the demands are as before.

In summary, the number of matched agents has increased by two and the number of single agents demanding a positive amount has decreased by at least one. The demands did not change, hence they are still good. (See Figure 3 for an illustration.)

Figure 3: Transition diagram for Case 1a.



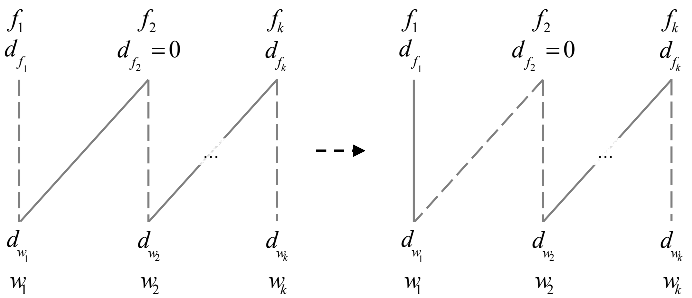
Case 1b. At least one firm on the path demands zero.

We shall demonstrate a sequence of adjustments leading to a state as in (10).

Let $P = (f_1, w_1, f_2, w_2, ..., w_{k-1}, f_k, w_k)$. There exists a firm $f_i \in P$ with current demand zero (f_2 in the illustration), hence no further reduction by f_i can occur. (If multiple firms on P have demand zero, let f_i be the first such firm on the path.) Apply the same sequence of transitions as in Case 1a up to firm f_i. At the end of this sequence the demands are as before. Once f_{i-1} snags w_{i-1}, f_i becomes single and still demands zero.

In summary, the number of single agents demanding a positive amount has decreased by one because f_1 is no longer single and the new single agent f_i demands zero. The demands did not change, hence they are still good. (See Figure 4 for an illustration.)

Figure 4: Transition diagram for Case 1b.



Case 2. Starting at f_1, *all* maximal alternating paths are of even length.

Case 2a. All firms on the paths demand a positive amount.

We shall demonstrate a sequence of adjustments leading to a state as in (11).

With positive probability f_1 is activated. Since no profitable match exists, he lowers his demand by δ. Hence, all previously tight edges starting at f_1 are now loose.

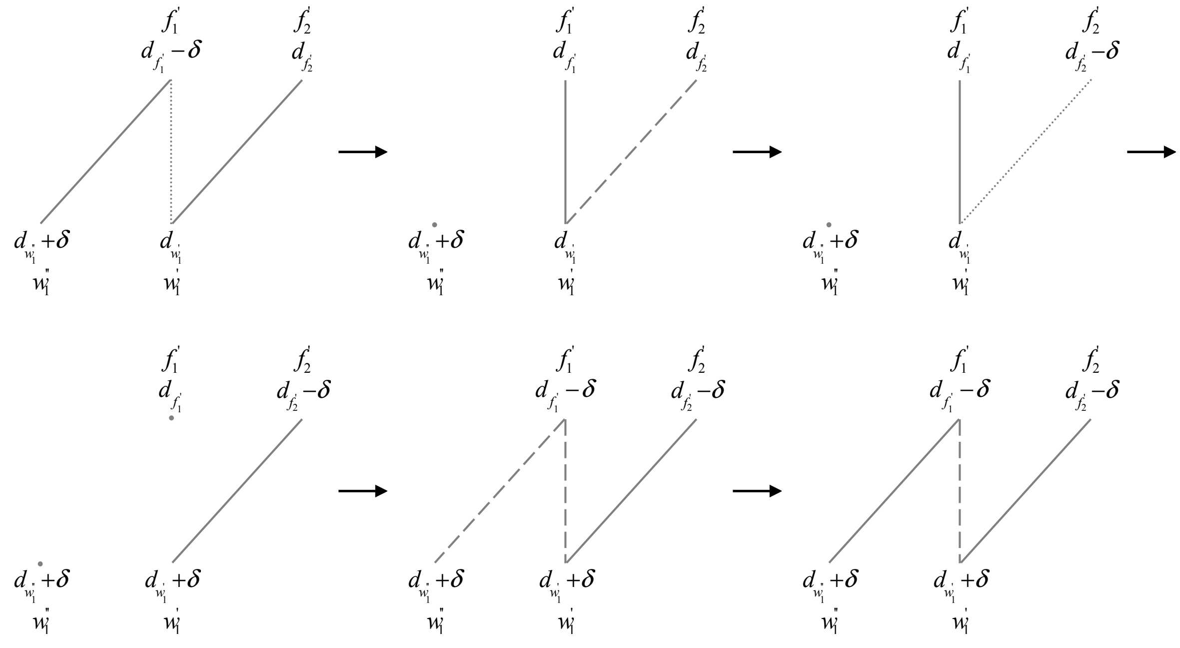
We shall describe a sequence of transitions under which a given loose edge is eliminated (by making it tight again), the matching does not change and the sum of demands remains fixed. Consider a loose edge between a firm, say f'_1, and a worker, say w'_1. Since all maximal alternating paths starting at f_1 are of even length, the worker has to be matched to a firm, say f_2'. With positive probability w_1' is activated, snags f_1', and with probability 0.5 f_1' receives the residual δ. (Such a transition occurs with strictly positive probability whether or not f_1' is matched because demands are strictly below the match value of (w_1', f_1').) Note that f_2' and possibly f_1''s previous partner, say w_1'', are now single. With positive probability f_2' is activated. Since no profitable match exists, he lowers his demand by δ. (This occurs because all firms on the maximal alternating paths starting at f_1 demand at least δ.) With positive probability, f_2' is activated again, snags w_1', and with probability 0.5 w_1' receives the residual δ. Finally, with positive probability f_1' is activated. Since no profitable match exists, he lowers his demand by δ. If previously matched, f_1' is activated again in the next period and matches with w_1''. At the end of this sequence the matching is the same as at the beginning. Moreover, w_1''s demand went up by δ while f_2''s demand went down by δ and all other demands stayed the same. The originally loose edge between f_1' and w_1' is now tight.

We iterate this construction until all loose edges at f_1' have been eliminated. However, given f_2''s reduction by δ there may be new loose edges connecting f_2' to workers. In this case we repeat the preceding construction for f_2' until all of the loose edges at f_2' have been eliminated. If any firms still exist with loose edges we repeat the construction again. This iteration eventually terminates given the following observation. Any worker on a maximal alternating path who previously increased his demand cannot still be connected to a firm by a loose edge. Similarly, any firm that previously reduced its demand cannot now be matched to a worker with a loose edge because such a worker increased his demand. Therefore the preceding construction involves any given firm (or worker) at most once. It follows that, in a finite number of periods, all firms on maximal alternating paths starting at f_1 have reduced their demand by δ and all workers have increased their demand by δ.

In summary, the number of demand reductions outnumbers the number of demand increases by one, hence the sum of the demands has decreased. The number of single agents demanding a positive amount has not increased. Moreover the demands are still good. (See Figure 5 for an illustration.)

Note that the δ-reductions may lead to new tight edges, resulting in new maximal alternating paths of odd or even lengths.

Figure 5: Transition diagram for Case 2a.



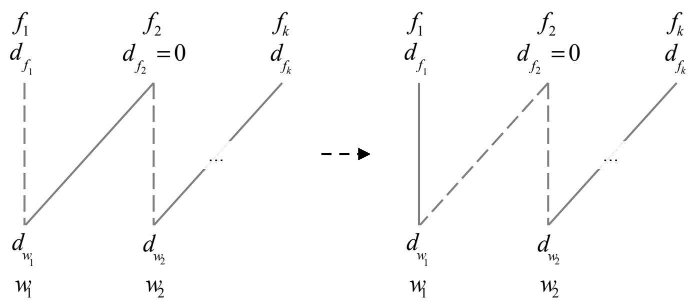
Case 2b. At least one firm on a path demands zero.

We shall demonstrate a sequence of adjustments leading to a state as in (10).

Let $P = (f_1, w_1, f_2, w_2, ..., w_{k-1}, f_k)$. There exists a firm $f_i \in P$ with current demand zero (f_2 in the illustration), hence no further reduction by f_i can occur. (If multiple firms on P have demand zero, let f_i be the first such firm on the path.) With positive probability f_1 is activated. Since no profitable match exists, he lowers his demand by δ. With positive probability, f_1 is activated again next period, he snags w_1 and with probability 0.5 he receives the residual δ. Now f_2 is single. With positive probability f_2 is activated, lowers, snags w_2, and so forth. This sequence continues until f_i is reached, who is now single with demand zero.

In summary, the number of single agents demanding a positive amount has decreased. The demands did not change, hence they are still good. (See Figure 6 for an illustration.)

Figure 6: Transition diagram for Case 2b.



Let us summarize the argument so far. Starting in a state $[\mathbf{A}, \mathbf{d}]$ with good demands $\mathbf{d}$, we successively eliminate the odd paths (if any exist) and then we eliminate the even paths (if any exist), while maintaining good demands. This process must come to an end because at each iteration either the sum of demands goes down by δ and the number of single agents with positive demand stays fixed, or the sum of demands stays fixed and the number of single agents with positive demand goes down. Finally, single agents (with demand zero) successively match at demand zero until all agents on the smaller side of the market are matched. The resulting state must be in the core. Moreover any core state reached by such a sequence of transitions is absorbing because singles cannot reduce further and no new matches can be formed. We have therefore shown that the process $[\mathbf{A}^t, \mathbf{d}^t]$ is absorbed into the core in finite time with probability 1. □

6. Conclusion

In this paper we have shown that agents in large decentralized matching markets can learn to play stable and efficient outcomes through a trial-and-error learning process. We assume that the agents have no information about the distribution of others' preferences, their past actions and payoffs, or about the value of different matches. Nevertheless the learning process leads to the core with probability one. The proof uses integer programming arguments (Kuhn [1955], Balinski [1965]), but the Matchmaker does not "solve" an integer programming problem. Rather, a path into the core is discovered in finite time by a random sequence of adjustments by the agents.

A crucial feature of our model is that the Matchmaker has no knowledge of match values, hence standard matching procedures cannot be used. In fact, the role of the Matchmaker can be eliminated entirely, and the process can be interpreted as a purely evolutionary process with no third party at all. As before, let agents be activated by independent Poisson clocks. Suppose that an active agent randomly encounters one agent from the other side of the market drawn from a distribution with full support. The two players enter a new match with positive probability if their match is potentially profitable, which they can see from their current bids and offers. If the two players are already matched with each other, they remain so. If both are single, they agree to be matched if their bid and offer cross. If at least one agent is matched (but with someone else), they agree to be matched if their bid and offer *strictly* cross. This is essentially the same process as the one described above, and the same proof shows that it leads to the core in finite time with probability one.

References

M. Agastya, "Adaptive Play in Multiplayer Bargaining Situations", *Review of Economic Studies 64*, 411-26, 1997.

M. Agastya, "Perturbed Adaptive Dynamics in Coalition Form Games", *Journal of Economic Theory 89*, 207-233, 1999.

T. Arnold & U. Schwalbe, "Dynamic coalition formation and the core", *Journal of Economic Behavior and Organization 49*, 363-380, 2002.

M. L. Balinski, "Integer Programming: Methods, Uses, Computations", *Management Science 12*, 253-313, 1965.

M. L. Balinski & D. Gale, "On the Core of the Assignment game", in *Functional Analysis, Optimization and Mathematical Economics*, L. J. Leifman (ed.), Oxford University Press, 274-289, 1987.

R. Bush & F. Mosteller, *Stochastic Models of Learning,* Wiley, 1955.

V. P. Crawford & E. M. Knoer, "Job Matching with Heterogeneous Firms and Workers", *Econometrica 49*, 437-540, 1981.

G. Demange & D. Gale, "The strategy of two-sided matching markets", *Econometrica 53,* 873-988, 1985.

G. Demange, D. Gale & M. Sotomayor, "Multi-item auctions", *Journal of Political Economics 94,* 863-872, 1986.

M. L. Elliott, "Inefficiencies in networked markets", *working paper,* Stanford University, 2010.

M. L. Elliott, "Search with multilateral bargaining", *working paper,* Stanford University, 2011.

W. Estes, "Towards a statistical theory of learning", *Psychological Review 57*, 94-107, 1950.

D. Foster & H. P. Young, "Regret testing: Learning to play Nash equilibrium without knowing you have an opponent", *Theoretical Economics 1*, 341-367, 2006.

D. Gale & L. S. Shapley, "College admissions and the stability of marriage", *American Mathematical Monthly 69*, 9-15, 1962.

F. Germano & G. Lugosi, "Global Nash convergence of Foster and Young's regret testing", *Games and Economic Behavior 60*, 135-154, 2007.

S. Hart & A. Mas-Colell, "Uncoupled Dynamics Do Not Lead to Nash Equilibrium", *American Economic Review 93*, 1830-1836, 2003.

R. J. Herrnstein, "Relative and absolute strength of response as a function of frequency of reinforcement", *Journal of Experimental Analysis of Behavior 4*, 267-272, 1961.

F. Hoppe, "Erfolg und Mißerfolg", *Psychologische Forschung 14,* 1-62, 1931.

R. Karandikar, D. Mookherjee, D. Ray & F. Vega-Redondo, “Evolving Aspirations and Cooperation”, *Journal of Economic Theory 80*, 292-331, 1998.

A. S. Kelso & V. P. Crawford, “Job Matching, Coalition Formation, and Gross Substitutes”, *Econometrica 50,* 1483-1504, 1982.

H. W. Kuhn, “The Hungarian Method for the assignment problem”, *Naval Research Logistic Quarterly 2,* 83-97, 1955.

J. R. Marden, H. P. Young, G. Arslan, J. Shamma, “Payoff-based dynamics for multiplayer weakly acyclic games”, *SIAM Journal on Control and Optimization 48,* special issue on “Control and Optimization in Cooperative Networks”, 373-396, 2009.

J. Newton, “Non-cooperative convergence to the core in Nash demand games without random errors or convexity assumptions”, Ph.D. thesis, University of Cambridge, 2010.

J. Newton, “Recontracting and stochastic stability in cooperative games”, *Journal of Economic Theory 147(1)*, 364-381, 2012.

B. S. R. Pradelski & H. P. Young, “Learning Efficient Nash Equilibria in Distributed Systems”, *Games and Economic Behavior 75*, 882-897, 2012.

A. E. Roth, “The Evolution of the Labor Markets for Medical Interns and Residents: A Case Study in Game Theory”, *Journal of Political Economy 92,* 991-1016, 1984.

A. E. Roth & E. Peranson, “The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design”, *The American Economic Review 89,* 756-757, 1999.

A. E. Roth, T. Sönmez & U. Ünver, “Pairwise kidney exchange”, *Journal of Economic Theory 125*, 151-188, 2005.

A. E. Roth & M. Sotomayor, *Two-Sided Matching: A Study in Game Theoretic Modeling and Analysis*, Cambridge University Press, 1990.

A. E. Roth & M. Sotomayor, “Two-sided matching”, in *Handbook of Game Theory with Economic Applications*, Volume 1, R. Aumann & S. Hart (eds.), 485-541, 1992.

K. Rozen, “Conflict Leads to Cooperation in Nash Bargaining”, *mimeo,* Yale University, 2010a.

K. Rozen, “Conflict Leads to Cooperation in Nash Bargaining: Supplemental Result on Evolutionary Dynamics”, *web appendix,* Yale University, 2010b.

T. Sandholm, “Computing in Mechanism Design”, *New Palgrave Dictionary of Economics*, 2008.

R. Sawa, “Coalitional stochastic stability in games, networks and markets”, *working paper*, University of Wisconsin-Madison, 2011.

L. S. Shapley & M. Shubik, “The Assignment Game I: The Core”, *International Journal of Game Theory 1*, 111-130, 1972.

R. Shimer, “Mismatch”, *The American Economic Review 97*, 1074-1101, 2007.

R. Shimer, “The Probability of Finding a Job”, *The American Economic Review 98* (Papers and Proceedings), 268-273, 2008.

M. Sotomayor, “Some further remark on the core structure of the assignment game”, *Mathematical Social Sciences 46*, 261-265, 2003.

E. Thorndike, “Animal Intelligence: An Experimental Study of the Associative Processes in Animals”, *Psychological Review 8*, 1898.

H. P. Young, “Learning by trial and error”, *Games and Economic Behavior 65*, 626-643, 2009.