Heinrich H. Naxy, Bary S. R. Pradelski

Evolutionary dynamics and equitable core selection in assignment games

Article (Accepted version) (Refereed)

Original citation:
Nax, Heinrich H. and Pradelski, Bary S. R. (2015) *Evolutionary dynamics and equitable core selection in assignment games.* International Journal of Game Theory, 44 (4). pp. 903-932. ISSN 0020-7276

DOI: 10.1007/s00182-014-0459-1

Evolutionary dynamics and equitable core selection in assignment games*

Heinrich H. Nax† & Bary S. R. Pradelski‡

June 1, 2012
This version: November 10, 2014

Abstract

We study evolutionary dynamics in assignment games where many agents interact anonymously at virtually no cost. The process is decentralized, very little information is available and trade takes place at many different prices simultaneously. We propose a completely uncoupled learning process that selects a subset of the core of the game with a natural equity interpretation. This happens 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. In our model, agents randomly encounter other agents, make bids and offers for potential partnerships and match if the partnerships are profitable. Equity is favored by our dynamics because it is more stable, not because of any ex ante fairness criterion.

JEL classifications: C71, C73, C78, D83

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

*Foremost, we thank Peyton Young for his guidance. He worked with us throughout large parts of this project and provided invaluable guidance. Further, we thank Itai Arieli, Peter Biró, Gabrielle Demange, Sergiu Hart, Gabriel Kreindler, Jonathan Newton, Tom Norman, Tamás Solymosi, Zaifu Yang and anonymous referees for suggesting a number of improvements to earlier versions. We are also grateful for comments by participants at the 23rd International Conference on Game Theory at Stony Brook, the Paris Game Theory Seminar, the AFOSR MUIR 2013 meeting at MIT, the 18th CTN Workshop at the University of Warwick, the Economics Department theory group at the University of York, the Theory Workshop at the Center for the Study of Rationality at the Hebrew University, and the Game Theory Seminar at the Technion. The research was supported by the United States Air Force Office of Scientific Research Grant FA9550-09-1-0538 and the Office of Naval Research Grant N00014-09-1-0751.

†Department of Social Sciences-SOMS, ETH Zürich, Clausiusstrasse 50, 8092 Zürich, Switzerland. Nax acknowledges support by the European Commission through the ERC Advanced Investigator Grant 'Momentum' (Grant No. 324247).

‡University of Oxford, Eagle House, Walton Well Road, Oxford OX2 6ED, UK. Pradelski acknowledges support of the Oxford-Man Institute of Quantitative Finance.

1. *Introduction*

Many matching markets are decentralized and agents interact repeatedly with very little knowledge about the market as a whole. Examples include online markets for bringing together buyers and sellers of goods, matching workers with firms, matching hotels with clients, and matching men and women. In such markets matchings are repeatedly broken, reshuffled, and restored. Even after many encounters, however, agents may still have little information about the preferences of others, and they must experiment extensively before the market stabilizes.

In this paper we propose a simple adaptive process that reflects the participants' limited information about the market. Agents have aspiration levels that they adjust from time to time based on their experienced payoffs. Matched agents occasionally experiment with higher bids in the hope of extracting more from another match, while single agents occasionally lower their bids in the hope of attracting a partner. There is no presumption that market participants or a central authority know anything about the distribution of others' preferences or that they can deduce such information from prior rounds of play. Instead they follow a process of trial and error in which they adjust their bids and offers in the hope of increasing their payoffs. Such aspiration adjustment rules are rooted in the psychology and learning literature.[1] A key feature of the rule we propose is that an agent's behavior does not require any information about other agents' actions or payoffs: the rule is completely uncoupled.[2] It is therefore particularly well-suited to environments such as decentralized online markets where players interact anonymously and trades take place at many different prices. We shall show that this simple adaptive process leads to equitable solutions inside the core of the associated *assignment game* (Shapley and Shubik 1972). In particular, core stability and equity are achieved even though agents have no knowledge of the other agents' strategies or preferences, and there is no ex ante preference for equity.

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 relevant solution concepts. Section 4 describes the process of adjustment and search by individual agents. In sections 5 and 6 we show that the stochastically stable states of the process lie inside the core. Section 7 concludes with several open problems.

[1] There is an extensive literature in psychology and experimental game theory on trial and error and aspiration adjustment. See in particular the learning models of Thorndike (1898), Hoppe (1931), Estes (1950), Bush and Mosteller (1955), Herrnstein (1961), and aspiration adjustment and directional learning dynamics of Heckhausen (1955), Sauermann and Selten (1962), Selten and Stoecker (1986), Selten (1998).

[2] This idea was introduced by Foster and Young (2006) and is a refinement of the concept of uncoupled learning due to Hart and Mas-Colell (2003, 2006). Recent work has shown that there exist completely uncoupled rules that lead to pure Nash equilibrium in generic noncooperative games with pure Nash equilibria (Germano and Lugosi 2007, Marden et al. 2009, Young 2009, Pradelski and Young 2012).

2. *Related literature*

Our results fit into a growing literature showing how cooperative game solutions can be implemented via noncooperative dynamic learning processes (Agastya 1997, 1999, Arnold and Schwalbe 2002, Newton 2010, 2012, Sawa 2011, Rozen 2013). A particularly interesting class of cooperative games are assignment games, in which every candidate matched pair has a cooperative 'value'. Shapley and Shubik (1972) showed that the core of such a game is always nonempty.[3] Subsequently various authors have explored refinements of the assignment game core, including the kernel (Rochford 1984) and the nucleolus (Huberman 1980, Solymosi and Raghavan 1994, Nunez 2004, Llerena et al. 2012). To the best of our knowledge, however, there has been no prior work showing how a core refinement is selected via a decentralized learning process, which is the subject of the present paper.

This paper establishes convergence to the core of the assignment game for a class of natural dynamics and selection of a core refinement under payoff perturbations. We are not aware of prior work comparable with our selection result. There are, however, several recent papers that also address the issue of core convergence for a variety of related processes (Chen et al. 2011, Biró et al. 2012, Klaus and Payot 2013, Bayati et al. 2014). These processes are different from ours, in particular they are not aspiration-adjustment learning processes, and they do not provide a selection mechanism for a core refinement as we do here. The closest relative to our paper is the concurrent paper by Chen et al. (2011), which demonstrates a decentralized process where, similarly as in our process, pairs of players from the two market sides randomly meet in search of higher payoffs. This process also leads almost surely to solutions in the core. Chen et al. (2011) and our paper are independent and parallel work. They provide a constructive proof based on their process which is similar to ours for the proof of the convergence theorem. Thus, theirs as well as our algorithm (proof of Theorem 1) can be used to find core outcomes. Biró et al. (2012) generalizes Chen et al. (2011) to transferable-utility roommate problems. In contrast to Chen et al. (2011) and our proof, Biró et al. (2012) use a target argument which cannot be implemented to obtain a core outcome. Biró et al. (2012)'s proof technique is subsequently used in Klaus and Payot (2013) to prove the result of Chen et al. (2011) for continuous payoff space in the assignment game. A particularity in this case is the fact that the assignment may continue to change as payoffs approximate a core outcome. Finally, Bayati et al. (2014) study the rate of convergence of a related bargaining process for the roommate problem in which players know their best alternatives at each stage. The main difference of this process to ours is that agents best reply (i.e. they have a lot of information about their best alternatives). Moreover, the order of activation is fixed, not random, and matches are only formed once a stable outcome is found.[4] An important feature of our learning process is that it is explicitly formulated in terms of random bids of workers and random offers of firms (as in the original set-up by Shapley and Shubik 1972), which allows a completely uncoupled set-up of the dynamic.

[3]Important subsequent papers include Crawford and Knoer (1981), Kelso and Crawford (1982), Demange and Gale (1985), and Demange et al. (1986).

[4]In a recent paper Pradelski (2014) discusses the differences to our set-up in more detail. He then investigates the convergence rate properties of a process closely related to ours.

There is also a related literature on the marriage problem (Gale and Shapley 1962). In this setting the players have ordinal preferences for being matched with members of the other population, and the core consists of matchings such that no pair would prefer each other to their current partners.[5] Typically, many matchings turn out to be stable. Roth and Vande Vate (1990) demonstrate a random blocking pair dynamic that leads almost surely to the core in such games. Chung (2000), Diamantoudi et al. (2004) and Inarra et al. (2008, 2013) establish similar results for nontransferable-utility roommate problems, while Klaus and Klijn (2007) and Kojima and Ünver (2008) treat the case of many-to-one and many-to-many nontransferable-utility matchings. Another branch of the literature considers stochastic updating procedures that place high probability on core solutions, that is, the stochastically stable set is contained in the core of the game (Jackson and Watts 2002, Klaus et al. 2010, Newton and Sawa 2013).

The key difference between marriage problems and assignment games is that the former are framed in terms of nontransferable (usually ordinal) utility, whereas in the latter each potential match has a transferable 'value'. The core of the assignment game consists of outcomes such that the matching is optimal and the allocation is pairwise stable. Generically, the optimal matching is unique and the allocations supporting it infinite. On the face of it one might suppose that the known results for marriage games would carry over easily to assignment games but this is not the case. The difficulty is that in marriage games (and roommate games) a payoff-improving deviation is determined by the players' current matches and their preferences, whereas in an assignment game it is determined by their matches, the value created by these matches, and by how they currently split the value of the matches. Thus the core of the assignment game tends to be significantly more constrained and paths to the core are harder to find than in the marriage game.

The contribution of the present paper is to demonstrate a simple completely uncoupled adjustment process that has strong selection properties for assignment games. Using a proof technique introduced by Newton and Sawa (2013) (the one-period deviation principle), we show that the stochastically stable solutions of our process lie in a subset of the core of the assignment game. These solutions have a natural equity interpretation: namely, every pair of matched agents splits the difference between the highest and lowest payoffs they could get without violating the core constraints.

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.

[5]See Roth and Sotomayor (1992) for a textbook on two-sided matching.

The population $N = F \cup W$ consists of firms $F = \{f_1, ..., f_m\}$ and workers $W = \{w_1, ..., w_n\}$.[6] They interact by making bids and offers to potential partners whom they randomly encounter. We assume matches form only if these bids are mutually profitable for both agents.

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

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, for example, 'dollars'.

We shall introduce time at this stage to consistently develop our notation. Let $t = 0, 1, 2, ...$ be the time periods.

Assignment. For all pairs of agents $(i, j) \in F \times W$, let $a^t_{ij} \in \{0, 1\}$.

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

If for a given agent $i \in N$ there exists j such that $a^t_{ij} = 1$ we shall refer to that agent as *matched*; otherwise i is *single*. An assignment $\mathbf{A}^t = (a^t_{ij})_{i \in F, j \in W}$ is such that if $a^t_{ij} = 1$ for some (i, j), then $a^t_{ik} = 0$ for all $k \neq j$ and $a^t_{lj} = 0$ for all $l \neq i$.

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

- $F = \{f_1, ..., f_m\}$ is the set of m firms (or men or sellers),
- $W = \{w_1, ..., w_n\}$ is the 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.
- $\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ is the assignment matrix with 0/1 values and row/column sums at most one.

[6] The two sides of the market could also, for example, represent buyers and sellers, or men and women in a (monetized) marriage market.

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

Note, that the game at hand is a cooperative game:

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 $S \subseteq F \times W$,

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.

3.2 Dynamic components

Aspiration level. At the end of any period t, a player has an *aspiration level*, d_i^t, which determines the minimal payoff at which he is willing to be matched. Let $\mathbf{d}^t = \{d_i^t\}_{i \in F \cup W}$.

Bids. In any period t, one pair of players is drawn at random and they make bids for each other. We assume that the two players' bids are such that the resulting payoff to each player is at least equal to his aspiration level, and with positive probability is exactly equal to his aspiration level.

Formally, firm $i \in F$ encounters $j \in W$ and submits a random bid $b_i^t = p_{ij}^t$, where p_{ij}^t is the maximal amount i is currently willing to pay if matched with j. Similarly, worker $j \in W$ submits $b_j^t = q_{ij}^t$, where q_{ij}^t is the minimal amount j is currently willing to accept if matched with i. A bid is separable into two components; the current (deterministic) aspiration level and a random variable that represents an exogenous shock to the agent's aspiration level. Specifically let P_{ij}^t, Q_{ij}^t be independent random variables that take values in $\delta \cdot \mathbb{N}_0$ where 0 has positive probability.[7] We thus have, for all i, j,

$$p_{ij}^t = (p_{ij}^+ - d_i^{t-1}) - P_{ij}^t \qquad \text{and} \qquad q_{ij}^t = (q_{ij}^- + d_j^{t-1}) + Q_{ij}^t. \tag{3}$$

Consider, for example, worker j's bid for firm i. The amount q_{ij}^- is the minimum that j would ever accept to be matched with i, while d_j^{t-1} is his previous aspiration level over and above the minimum. Thus Q_{ij}^t is j's attempt to get even more in the current period. Note that if the random variable is zero, the agent bids exactly according to his current aspiration level.

Prices. When i is matched with j they trade at a unique *price*, π_{ij}^t.

[7]Note that $\mathbb{P}[P_{ij}^t = 0] > 0$ and $\mathbb{P}[Q_{ij}^t = 0] > 0$ are reasonable assumptions, since we can adjust p_{ij}^+ and q_{ij}^- in order for it to hold. This would alter the underlying game but then allow us to proceed as suggested.

Payoffs. Given $[\mathbf{A^t}, \mathbf{d^t}]$ the *payoff* to firm i / worker j is

$$\phi_i^t = \begin{cases} p_{ij}^+ - \pi_{ij}^t & \text{if } i \text{ is matched to } j, \\ 0 & \text{if } i \text{ is single.} \end{cases} \qquad \phi_j^t = \begin{cases} \pi_{ij}^t - q_{ij}^- & \text{if } j \text{ is matched to } i, \\ 0 & \text{if } j \text{ is single.} \end{cases} \tag{4}$$

Note that players' payoffs can be deduced from the aspiration levels and the assignment matrix.

Profitability. A pair of bids (p_{ij}^t, q_{ij}^t) is *profitable* if both players, in expectation, receive a higher payoff if the match is formed.

Note that, if two players' bids are at their aspiration levels and $p_{ij}^t = q_{ij}^t$, then they are profitable only if both players are currently single. Also note that a pair of players (i, j) with $\alpha_{ij} = 0$ will never match.

Re-match. At each moment in time, a pair (i, j) that randomly encounters each other matches if their bids are profitable. The resulting price, π_{ij}^t, is set anywhere between q_{ij}^t and p_{ij}^t. (Details about how players are activated are specified in the next section.)

To summarize, when a new match forms that is profitable, both agents receive a higher payoff in expectation due to the full support of the resulting price.[8]

States. The *state* at the end of period t is given by $Z^t = [\mathbf{A}^t, \mathbf{d}^t]$ where $\mathbf{A}^t \in \mathcal{A}$ is an assignment and $\mathbf{d}^t$ is the aspiration level vector. Denote the set of all states by Ω.

3.3 Solution concepts

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

Pairwise stability. An aspiration level vector $\mathbf{d}^t$ is *pairwise stable* if, $\forall i, j$ and $a_{ij}^t = 1$,

$$p_{ij}^+ - d_i^t = q_{ij}^- + d_j^t, \tag{5}$$

and $p_{i'j}^+ - d_{i'}^t \leq q_{i'j}^- + d_j^t$ for every alternative firm $i' \in F$ with $i' \neq i$ and $q_{ij'}^- + d_{j'}^t \geq p_{ij'}^+ - d_i^t$ for every alternative worker $j' \in W$ with $j' \neq j$.

Core (Shapley and Shubik 1972). The *core* of any assignment game is always non-empty and consists of the set $\mathbf{C} \subseteq \Omega$ of all states Z such that $\mathbf{A}$ is an optimal assignment and $\mathbf{d}$ is pairwise stable.

Subsequent literature has investigated the structure of the core of the assignment game, which turns out to be very rich.[9] In order to investigate the constraints of pairwise stability in more detail the concept of 'payoff excess' will be useful:

Excess. Given state Z^t, the *excess* for a player i who is matched with j is

$$e_i^t = \phi_i^t - \max_{k\neq j}(\alpha_{ik} - \phi_k^t)_+. \tag{6}$$

[8] In this sense any alternative match that may block a current assignment because it is profitable (as defined earlier) is a strict blocking pair.

[9] See, for example, Roth and Sotomayor (1992), Balinski and Gale (1987), Sotomayor (2003).

The excess for player i describes the gap to his next-best alternative, that is, the smallest amount he would have to give up in order to profitably match with some other player $k \neq j$. If a player has negative excess, pairwise stability is violated. In a core allocation, therefore, all players have nonnegative excess. For the analysis of absorbing core states, note that the excess in payoff can be equivalently expressed in terms of the excess in aspiration level. This is the case since in absorbing core states aspiration levels are directly deducible from payoffs.

Minimal excess. Given state Z^t, the *minimal excess* is

$$e^t_{\min}(Z^t) = \min_{i:\, i \text{ matched}} e^t_i. \tag{7}$$

Based on the minimal excess of a state, we can define the kernel (Davis and Maschler 1965). For assignment games, the kernel coincides with the solution concept proposed by Rochford (1984), which generalizes a pairwise equal split solution à la Nash (1950).

Kernel (Davis and Maschler 1965, Rochford 1984). The *kernel* $\mathbf{K}$ of an assignment game is the set of states such that the matching is optimal and, for all matched pairs (i, j),

$$e^t_i =_\delta e^t_j, \tag{8}$$

where $=_\delta$ means "equality up to δ". (This is necessary given that we operate on the discrete grid.)

Given Z^t, extend the definition of excess to *coalitional excess* for coalition $S \subseteq N$; $e^t(S) = \sum_{i \in S} \phi^t_i - v(S)$. Now let $\mathbf{E}(\phi^t) \in \mathbb{R}^{2^{m+n}}$ be the vector of coalitional excesses for all $S \subseteq N$, ordered from smallest to largest. Say $\mathbf{E}(\phi)$ is lexicographically larger than $\mathbf{E}(\phi')$ for some k, if $\mathbf{E}_i(\phi) = \mathbf{E}_i(\phi')$ for all $i < k$ and $\mathbf{E}_k(\phi) < \mathbf{E}_k(\phi')$.[10]

Nucleolus (Schmeidler 1969). The *nucleolus* $\mathbf{N}$ of the assignment game is the unique solution that minimizes the lexicographic measure. (See also Huberman 1980, Solymosi and Raghavan 1994.)

For an analysis of the welfare properties and of the links between the kernel and the nucleolus of the assignment game see Nunez (2004) and Llerena et al. (2012).

Least core (Maschler et al. 1979). The *least core* $\mathbf{L}$ of an assignment game is the set of states Z such that the matching is optimal and the minimum excess is maximized, that is,

$$e_{\min}(Z) = \max_{Z' \in \mathbf{C}} e_{\min}(Z'). \tag{9}$$

Note that our definition of excess (equations (6) and (7)) applies to essential coalitions only (that is, for the case of the assignment game, to two-player coalitions involving exactly one agent from each market side). Hence, the least core generalizes the nucleolus of the assignment game in the following sense. Starting with the nucleolus, select any

[10]Note that the excess for coalitions, $e^t(S)$, is usually defined with a reversed sign. In order to make it consistent with definition (6) we chose to reverse the sign.

player with minimum excess (according to equation (6)): the least core contains all outcomes with a minimum excess that is not smaller.[11]

The following inclusions are known for the assignment game:[12]

$$\mathbf{N} \in (\mathbf{K} \cap \mathbf{L}), \ \ \mathbf{K} \subseteq \mathbf{C}, \ \ \mathbf{L} \subseteq \mathbf{C}. \tag{10}$$

4. Evolving play

A fixed population of agents, $N = F \cup W$, plays the assignment game $G(v, N)$. Repeatedly, a randomly activated agent encounters another agent, they make bids for each other and match if profitable. The distinct times at which one agent becomes active will be called *periods*. Agents are activated by independent Poisson clocks.[13] 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 if their match is profitable, which they can see from their current bids, offers and their own payoffs. If the two players are already matched with each other, they remain so.

4.1. Behavioral dynamics

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

1. The activated agent i makes a random encounter j.

2a. If the encounter is profitable given their current bids and assignment, the pair matches.

2b. If the match is not profitable, both agents return to their previous matches (or remain single).

3a. If a new match (i, j) forms, the price is set anywhere between bid and offer. The aspiration levels of i and j are set to equal their realized payoffs.

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

[11] See Shapley and Shubik (1963, 1966) for the underlying idea of the least core, the *strong ϵ-core*. See Maschler et al. (1979), Driessen (1999), Llerena and Nunez (2011) for geometric interpretations of these concepts.

[12] $\mathbf{N} \in \mathbf{K}$ is shown by Schmeidler (1969) for general cooperative games. Similarly $\mathbf{N} \in \mathbf{L}$ is shown by Maschler et al. (1979). Driessen (1998) shows for the assignment game that $\mathbf{K} \subseteq \mathbf{C}$. $\mathbf{L} \subseteq \mathbf{C}$ follows directly from the definitions.

[13] The Poisson clocks' arrival rates may depend on the agents' themselves or on their position in the game. Single agents, for example, may be activated faster than matched agents.

Our rules have antecedents in the psychology literature (Thorndike 1898, Hoppe 1931, Estes 1950, Bush and 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. The approach seems especially well-suited to modeling behavior in large decentralized assignment markets, where agents have little information about the overall game and about the identity of the other market participants. Following aspiration adjustment theory (Sauermann and Selten 1962, Selten 1998) and related bargaining experiments on directional and reinforcement learning (e.g., Tietz and Weber 1972, Roth and Erev 1995), we shall assume a simple directional learning model: matched agents occasionally experiment with higher offers if on the sell-side (or lower bids if on the buy-side), while single agents, in the hope of attracting partners, lower their offers if on the sell-side (or increase their bids if on the buy-side).

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 $i \in F$ be the unique *active* agent, for ease of exposition assumed to be a firm.

I. The active agent is currently matched and meets j

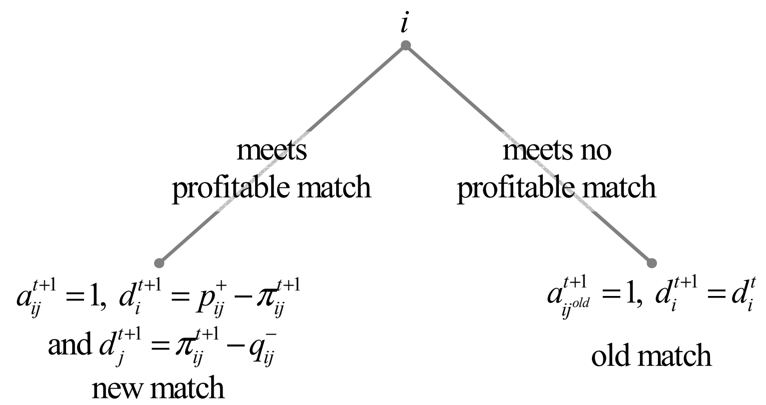
If i, j are profitable (given their current aspiration levels) they match. As a result, i's former partner is now single (and so is j's former partner if j was matched in period t). The price governing the new match, π_{ij}^{t+1}, is randomly set between p_{ij}^{t+1} and q_{ij}^{t+1}.

At the end of period $t+1$, the aspiration levels of the newly matched pair (i, j) are adjusted according to their newly realized payoffs:

$$d_i^{t+1} = p_{ij}^{+} - \pi_{ij}^{t+1} \quad \text{and} \quad d_j^{t+1} = \pi_{ij}^{t+1} - q_{ij}^{-}. \tag{11}$$

All other aspiration levels and matches remain fixed. If i, j are not profitable, i remains matched with his previous partner and keeps his previous aspiration level. See figure 1 for an illustration.

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



II. The active agent is currently single and meets j

If i, j are profitable (given their current aspiration levels) they match. As a result, j's former partner is now single if j was matched in period t. The price governing the new match, π_{ij}^{t+1}, is randomly set between p_{ij}^{t+1} and q_{ij}^{t+1}.

At the end of period $t+1$, the aspiration levels of the newly matched pair (i, j) are adjusted to equal their newly realized payoffs:

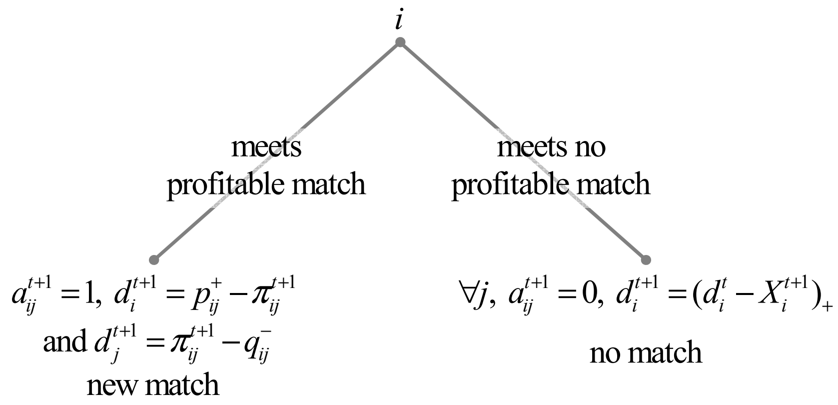
$$d_i^{t+1} = p_{ij}^{+} - \pi_{ij}^{t+1} \quad \text{and} \quad d_j^{t+1} = \pi_{ij}^{t+1} - q_{ij}^{-}. \tag{12}$$

All other aspiration levels and matches remain as before. If i, j are not profitable, i remains single and, with positive probability, reduces his aspiration level,

$$d_i^{t+1} = (d_i^t - X_i^{t+1})_+, \tag{13}$$

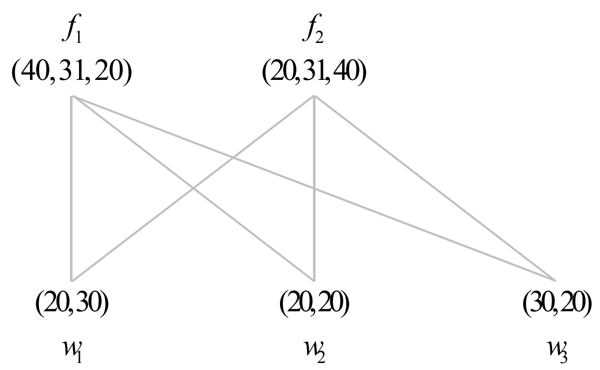
where X_i^{t+1} is an independent random variable taking values in $\delta \cdot \mathbb{N}_0$, and δ occurs with positive probability. See figure 2 for an illustration.

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



4.2. Example

Let $N = F \cup W = \{f_1, f_2\} \cup \{w_1, w_2, w_3\}$, $p_{1j}^{+} = (40, 31, 20)$ and $p_{2j}^{+} = (20, 31, 40)$ for $j = 1, 2, 3$, and $q_{i1}^{-} = (20, 30)$, $q_{i2}^{-} = (20, 20)$ and $q_{i3}^{-} = (30, 20)$ for $i = 1, 2$.



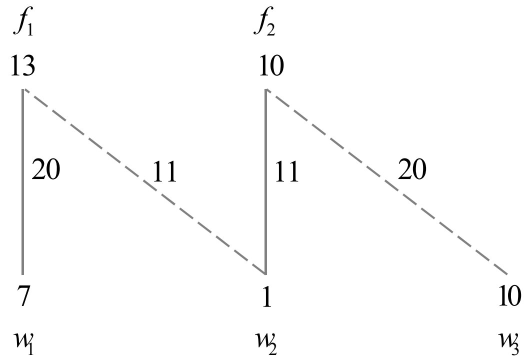
Then one can compute the match values: $\alpha_{11} = \alpha_{23} = 20$, $\alpha_{12} = \alpha_{22} = 11$, and $\alpha_{ij} = 0$ for all other pairs (i, j). Let $\delta = 1$.

period t: *Current state*

Suppose that, at the end of some period t, (f_1, w_1) and (f_2, w_2) are matched and w_3 is single.

The current aspiration level is shown next to the name of that agent, and the values α_{ij} are shown next to the edges (if positive). Bids will be shown to the right of the aspiration level. Solid edges indicate matched pairs, and dashed edges indicate unmatched pairs. (Edges with value zero are not shown.) Note that no player can see the bids or the status of the players on the other side of the market.

Note that some matches can never occur. For example f_1 is never willing to pay more than 20 for w_3, but w_3 would only accept a price above 30 from f_1.

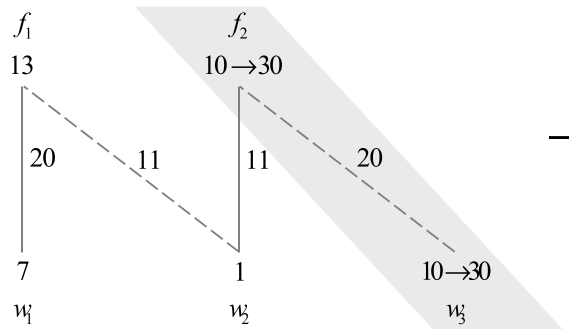


Note that the aspiration levels satisfy $d_i^t + d_j^t \geq \alpha_{ij}$ for all i and j, but the assignment is not optimal (firm 2 should match with worker 3).

period $t + 1$: *Activation of single agent w_3 and encounter of f_2*

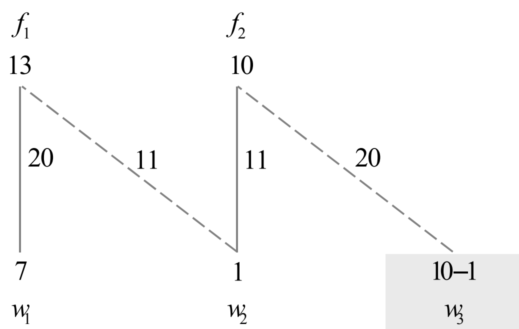
w_3's current aspiration level is too high in order to be profitable with f_2. Hence, independent of the specific bid he makes, he remains single and, with positive probability, reduces his aspiration level by 1.

w_3 encounters f_2



$\longrightarrow$

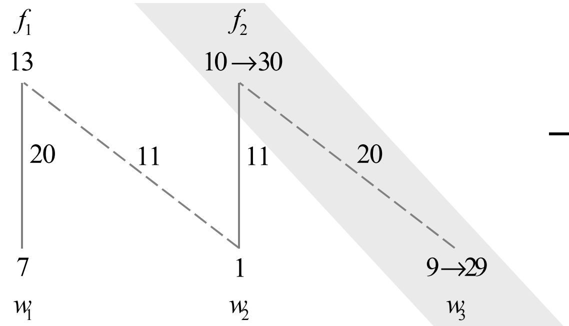
w_3 reduces aspiration level



period $t+2$: *Activation of matched agent* f_2 *and encounter of* w_3

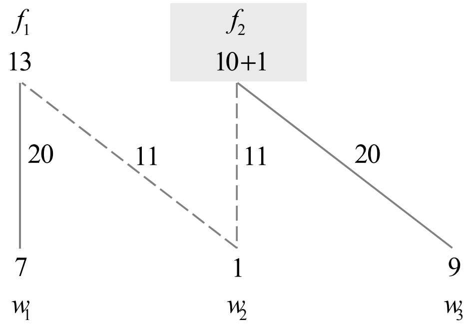
f_2 and w_3 are profitable. With positive probability f_2 bids 30 for w_3 and w_3 bids 29 for f_2 (hence the match is profitable), and the match forms. The price is set at random to either 29 such that f_2 raises his aspiration level by one unit (11) and w_3 keeps his aspiration level (9), or to 30 such that f_2 keeps his aspiration level (10) and w_3 raises his aspiration level by one unit (10). (Thus in expectation the agents get a higher payoff than before.)

f_2 encounters w_3



$\longrightarrow$

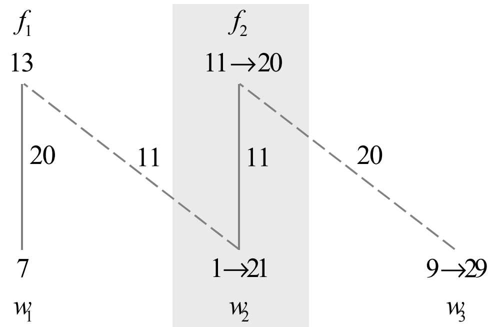
Successful match; f_2 increases aspiration level



period $t+3$: *Activation of single agent* w_2 *and encounter of* f_2

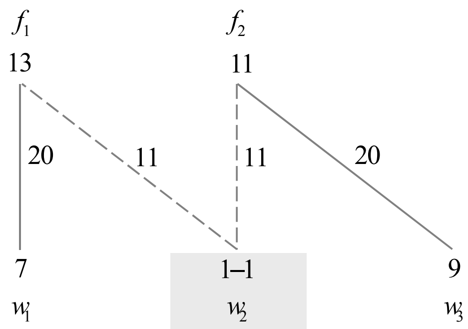
w_2's current aspiration level is too high in the sense that he has no profitable matches, and thus in particular is not profitable with f_2. Hence he remains single and, with positive probability, reduces his aspiration level by 1.

w_2 encounters f_2

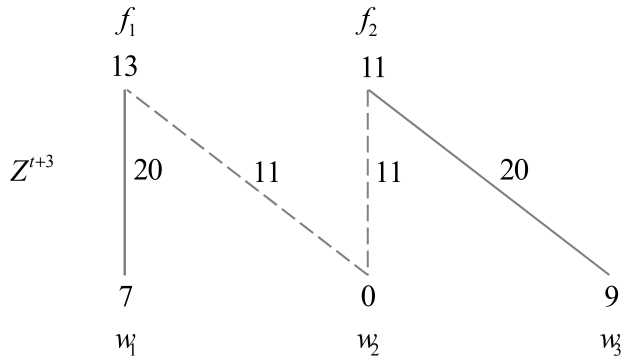


$\longrightarrow$

w_2 reduces aspiration level



The resulting state is in the core:[14]



5. Core stability – absorbing states of the unperturbed process

Recall that a state Z^t is defined by an assignment $\mathbf{A}^t$ and aspiration levels $\mathbf{d}^t$ that jointly determine the payoffs. $\mathbf{C}$ is the set of core states; let $\mathbf{C}_0$ be the set of core states such that singles' aspiration levels are zero.

Theorem 1. *Given an assignment game $G(v, N)$, from any initial state $Z^0 = [\mathbf{A}^0, \mathbf{d}^0] \in \Omega$, the process is absorbed into the core in finite time with probability 1. The set of absorbing states consists of $\mathbf{C}_0$. Further, starting from $\mathbf{d}^0 = 0$ any absorbing state is attainable.*

Throughout the proof we shall omit the time superscript since the process is time-homogeneous. The general idea of the proof is to show a particular path leading into the core which has positive probability. The proof uses integer programming arguments (Kuhn 1955, Balinski 1965) but no single authority 'solves' an integer programming problem. It will simplify the argument to restrict our attention to a particular class of paths with the property that the realizations of the random variables P_{ij}^t, Q_{ij}^t are always 0 and the realizations of X_i^t are always δ. P_{ij}^t, Q_{ij}^t determine the gaps between the bids and the aspiration levels, and X_i^t determines the reduction of the aspiration level by a single agent. One obtains from equation (3) for the bids:

$$\text{for all } i, j, \qquad p_{ij}^t = p_{ij}^+ - d_i^{t-1} \qquad \text{and} \qquad q_{ij}^t = q_{ij}^- + d_j^{t-1} \tag{14}$$

Recall that any two agents encounter each other in any period with positive probability. It shall be understood in the proof that the relevant agents in any period encounter each other. Jointly with equation (3), we can then say that a pair of aspiration levels (d_i^t, d_j^t) is *profitable* if

$$\text{either} \quad d_i^t + d_j^t < \alpha_{ij} \qquad \text{or} \qquad d_i^t + d_j^t = \alpha_{ij} \text{ and both } i \text{ and } j \text{ are single.} \tag{15}$$

Restricting attention to this particular class of paths will permit a more transparent analysis of the transitions, which we can describe solely in terms of the aspiration levels.

[14]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.

We shall proceed by establishing the following two claims.

Claim 1. There is a positive probability path to aspiration levels $\mathbf{d}$ 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 aspiration levels satisfying Claim 1 will be called *good*. Note that, even if aspiration levels are good, the assignment does not need to be optimal and not every agent with a positive aspiration level needs to be matched. (See the period-t example in the preceding section.)

Claim 2. Starting at any state with good aspiration levels, 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' aspiration levels are zero.[15]

Proof of Claim 1.

Case 1. Suppose the aspiration levels $\mathbf{d}$ are such that $d_i + d_j < \alpha_{ij}$ for some i, j. Note that this implies that i and j are not matched with each other since otherwise the entire surplus is allocated and $d_i + d_j = \alpha_{ij}$. With positive probability, either i or j is activated and i and j become matched. The new aspiration levels are set equal to the new payoffs. Thus the sum of the aspiration levels is equal to the match value α_{ij}. 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 aspiration levels $\mathbf{d}$ 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 aspiration level by δ. In this manner, a suitable path can be constructed, along which $\mathbf{d}$ decreases monotonically until the aspiration levels are good. Note that at the end of such a path, the assignment does not need to be optimal and not every agent with a positive aspiration level needs to be matched. (See the period-t example in the preceding section.) $\square$

Proof of Claim 2.

Suppose that the state $(\mathbf{A}, \mathbf{d})$ satisfies Claim 1 ($\mathbf{d}$ is good) and that some single exists whose aspiration level is positive. (If no such single exists, the assignment is optimal and we have reached a core state.) Starting at any such state, we show that, within a bounded number of periods and with positive probability (bounded below), one of the following holds:

> The aspiration levels are good, the number of single agents with positive aspiration level decreases, and the sum of the aspiration levels remains constant. (16)

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

The aspiration levels are good, the sum of the aspiration levels decreases by $\delta > 0,$ and the number of single agents with a positive aspiration level does not increase. (17)

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 path that starts at a single player with positive aspiration level, and that alternates between unmatched tight edges and matched tight edges such that it cannot be extended (hence maximal). Note that, for every single with a positive aspiration level, 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 with positive aspiration level.

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

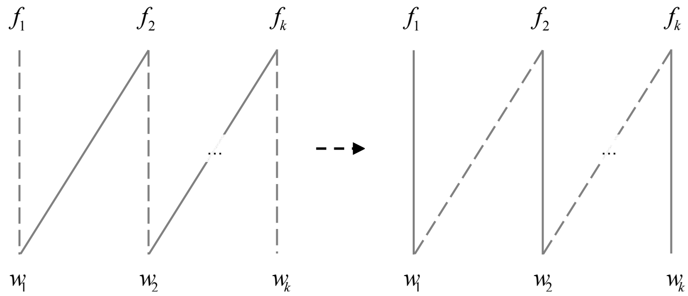
Case 1a. All firms on the path have a positive aspiration level.

We shall demonstrate a sequence of adjustments leading to a state as in (16).

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 aspiration level by δ. With positive probability, f_1 is activated again next period, matches with w_1 and receives the residual δ. At this point the aspiration levels are unchanged but f_2 is now single. With positive probability, f_2 is activated. Since no profitable match exists, he lowers his aspiration level by δ. With positive probability, f_2 is activated again next period, matches with w_2 and receives the residual δ. Within a finite number of periods a state is reached where all players on P are matched and the aspiration levels are as before. (Note that f_k is matched with w_k without a previous reduction by f_k since w_k is single and thus their bids are profitable.)

In summary, the number of matched agents has increased by two and the number of single agents with positive aspiration level has decreased by at least one. The aspiration levels did not change, hence they are still good. (See figure 3 for an illustration.)

Figure 3: Transition diagram for Case 1a.



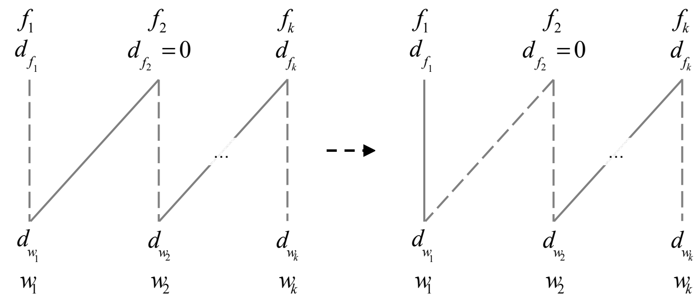
Case 1b. At least one firm on the path has aspiration level zero.

We shall demonstrate a sequence of adjustments leading to a state as in (16).

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 aspiration level zero (f_2 in the illustration), hence no further reduction by f_i can occur. (If multiple firms on P have aspiration level 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 aspiration levels are as before. Once f_{i-1} matches with w_{i-1}, f_i becomes single and his aspiration level is still zero.

In summary, the number of single agents with a positive aspiration level has decreased by one because f_1 is no longer single and the new single agent f_i has aspiration level zero. The aspiration levels 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 all maximal alternating paths starting at f_1 have a positive aspiration level.

We shall demonstrate a sequence of adjustments leading to a state as in (17).

Since aspiration levels will have changed by the end of the sequence of transitions, it does not suffice to only consider players along one maximal alternating path (for otherwise the aspiration levels will no longer be good). Instead, we need to consider *the union* of alternating paths (which are sets of edges) starting at f_1. Note that this union is a connected graph, say G_{f_1}. (Players may be part of multiple maximal alternating paths starting at f_1.) Now, fix a subgraph $T_{f_1} \subseteq G_{f_1}$ that includes all firms and workers, is connected, has no cycles and contains all of the matched edges. T_{f_1} is a *spanning tree* of G_{f_1} and thus always exists. Note that T_{f_1} only has alternating paths (between unmatched tight and matched tight edges) and all maximal alternating paths on this tree which start at f_1 contain an even number of edges. Thus there is one more firm on this graph than there are workers. (Think of f_1 as that extra firm.)

We shall describe a sequence of transitions along this tree such that, at its end, all firms on the graph have reduced their aspiration level by δ and all workers have increased their aspiration level by δ. We shall label firms and workers such that, for $i \geq 2$, f_i is at distance $2i-2$ from f_1 on the tree T_{f_1}, and, for $i \geq 1$, w_i is at distance $2i-1$ from f_i. Note that this implies that, for two agents with label i and j and $i < j$, the former agent is closer to f_1 than the latter. Let k be the maximal i such that f_i is on T_{f_1}.

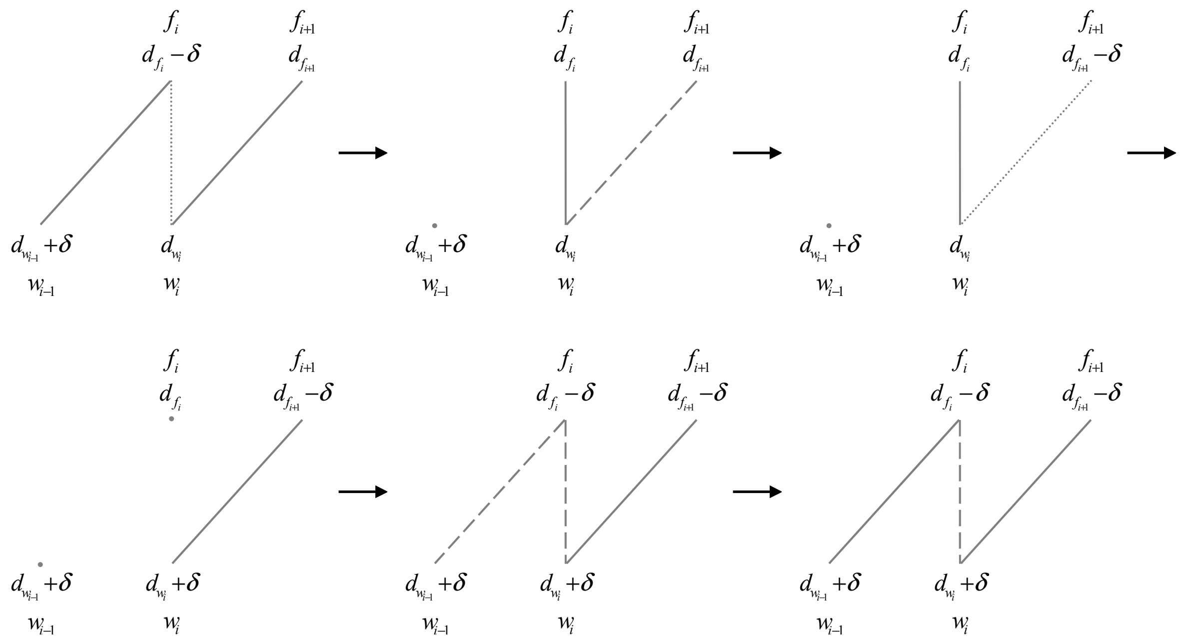
With positive probability f_1 is activated. Since no profitable match exists, f_1 lowers his aspiration level by δ. Hence, all previously tight edges starting at f_1 are now loose.

We shall describe a sequence of transitions which will tighten a loose edge along the tree T_{f_1} by making another edge on the tree loose. This new loose edge is further away from f_1 on the tree. At the end of this sequence the matching will not have changed and the sum of aspiration levels will have remained fixed.

Consider one such loose edge starting at $f_i \in T_{f_1}$. The loose edge from f_i is shared with a worker, say w_i. Since all maximal alternating paths starting at f_1 are of even length, the worker has to be matched to a firm, say f_{i+1}. With positive probability, w_i is activated, matches with f_i, and f_i receives the residual δ. (Such a transition occurs with strictly positive probability, whether or not f_i is matched, because the sum of aspiration levels is strictly below the match value of (w_i, f_i).) Note that f_{i+1} and, for $i \neq 1$, f_i's previous partner, w_{i-1}, are now single. With positive probability, f_{i+1} is activated and meets w_i again who is now matched with f_i. Since the aspiration levels of both players have not changed, their match is not profitable, and thus f_{i+1} lowers his aspiration level by δ. (Note that this reduction is possible because all firms on any maximal alternating path starting at f_1 have aspiration level at least δ.) With positive probability, f_{i+1} is activated again, matches with w_i, and w_i receives the residual δ. Finally, with positive probability, f_i is activated. Since no profitable match exists, he lowers his aspiration level by δ. For $i \neq 1$, f_i was previously matched, then f_i is activated with positive probability again in the next period and matches with the single w_{i-1} (there is no additional surplus to be split).

At the end of the sequence described above the assignment is the same as at the beginning. Moreover, w_i's aspiration level went up by δ while f_{i+1}'s aspiration level went down by δ and all other aspiration levels stayed the same. The originally loose edge between f_i and w_i is now tight. If f_{i+1} is not the last firm on its maximal alternating paths on T_{f_1}, there will be a new loose edge on T_{f_1} between f_{i+1} and workers with label $i+1$ (at distance $2i+1$). Note that there may be other loose edges not on T_{f_1}. These we shall not consider in our construction since they disappear at its end. (See figure 5 for an illustration.)

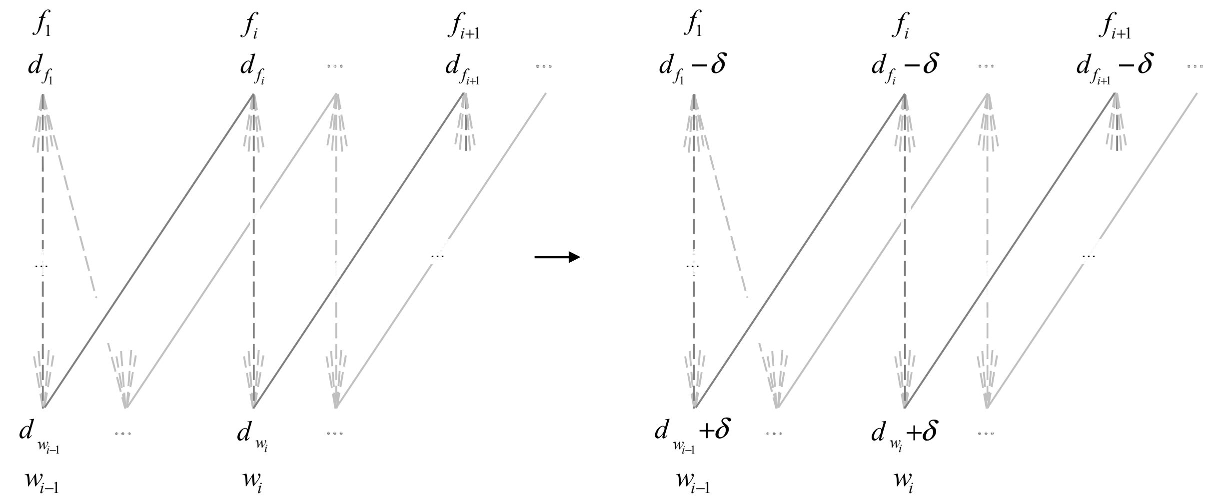
Figure 5: Detailed transitions for Case 2a.



We repeat the latter construction on T_{f_1} for $f_i = f_1$ until all loose edges at f_1 have been eliminated (*iteration 1*). This may create new loose edges on T_{f_1} between firms at distance 2 from f_1 (on T_{f_1}) and workers at distance 3 from f_1. Next we repeat the construction for all firms at distance 2 from f_1 on T_{f_1} (*iteration 2*). Note that the distance of loose edges to f_1 is increasing after each such iteration (recall that we consider the fixed spanning tree T_{f_1}). We thus keep iterating the construction for firms at distance $2i$ from f_1 for $i = 2, 3, \ldots, k-1$. The final iteration of our construction on a maximal alternating path does not lead to new loose edges on T_{f_1} (for otherwise the last firm on the unique path from f_1 on T_{f_1} would previously have had a tight edge, contradicting the assumption that the alternating path is maximal.) Hence, in a finite number of periods, all firms on T_{f_1}, and thus on G_{f_1}, have reduced their aspiration level by δ, and all workers have increased their aspiration level by δ. Given that G_{f_1} contained *all* maximal alternating paths starting at f_1, and therefore all tight paths starting at f_1, there are no more loose edges connected to any firm or worker on G_{f_1}. This is the case for two reasons. First, all edges on G_{f_1} were tight to begin with and, after each firm reduced his aspiration level by δ and each worker increased his aspiration level by δ, they are again tight at the end of all iterations. Second, there cannot be loose edges with firms or workers outside G_{f_1} since they were not tight before, and thus one δ-reduction cannot make them loose.

In summary, aspiration level reductions outnumber aspiration level increases by one (namely by the δ-reduction by firm f_1), hence the sum of the aspiration levels has decreased. The number of single agents with a positive aspiration level has not increased. Moreover the aspiration levels are still good. (See figure 6 for an illustration.)

Figure 6: Transition diagram for Case 2a.



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

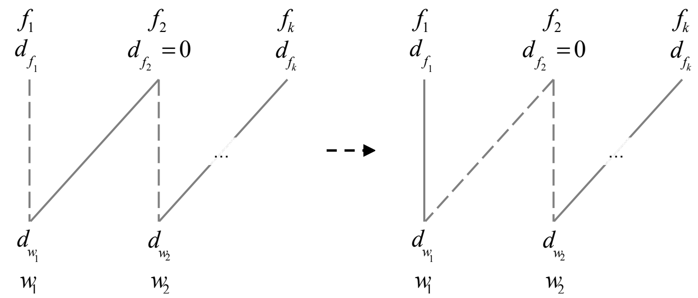
Case 2b. At least one firm on a maximal alternating path starting at f_1 has aspiration level zero.

We shall demonstrate a sequence of adjustments leading to a state as in (16).

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

In summary, the number of single agents with a positive aspiration level has decreased. The aspiration levels did not change, hence they are still good. (See figure 7 for an illustration.)

Figure 7: Transition diagram for Case 2b.



Let us summarize the argument. Starting in a state $[\mathbf{A}, \mathbf{d}]$ with good aspiration levels $\mathbf{d}$, we successively (if any exist) eliminate the odd paths starting at firms/workers followed by the even paths starting at firms/workers, while maintaining good aspiration levels. This process must come to an end because, at each iteration, either the sum of aspiration levels decreases by δ and the number of single agents with positive aspiration levels stays fixed, or the sum of aspiration levels stays fixed and the number of single agents with positive aspiration levels decreases. The resulting state must be in the core and is absorbing because single agents cannot reduce their aspiration level further and no new matches can be formed. Since an aspiration level constitutes a lower bound on a player's bids we can conclude that the process Z^t is absorbed into the core in finite time with probability 1. Finally note that, starting from $\mathbf{d}^0 = 0$, we can trivially reach any state in $\mathbf{C}_0$. $\square$

6. *Core selection*

In this section, we investigate the effects of random perturbations to the adjustment process. Suppose that players occasionally experience shocks when in a match and that larger shocks are less likely than smaller shocks. The effect of such a shock is that a player receives more or less payoff than anticipated given the current price he agreed to with his partner. We shall formalize these perturbations and investigate the resulting selection of stochastically stable states as the probability of shocks becomes vanishingly small (Foster and Young 1990, Kandori et al. 1993, Young 1993). It turns out that the set of stochastically stable states is contained in the *least core*; moreover there are natural conditions under which it coincides with the least core.

Given a player i who is matched in period t, suppose his unperturbed payoff ϕ_i^t is subject to a shock. Denote the new payoff by $\hat{\phi}_i^t$ and define:

$$\hat{\phi}_i^t = \begin{cases} \phi_i^t + \delta \cdot R_i^t \text{ with probability } 0.5, \\ \phi_i^t - \delta \cdot R_i^t \text{ with probability } 0.5, \end{cases} \tag{18}$$

where R_i^t is an independent geometric random variable with $\mathbb{P}[R_i^t = k] = \epsilon^k \cdot (1-\epsilon)$ for all $k \in \mathbb{N}_0$.[16] Note that for $\epsilon = 0$ the process is unperturbed.

The immediate result of a given shock is that players receive a different payoff than anticipated. We shall assume that players update their aspiration levels to equal their new perturbed payoff if the payoff is positive and zero if if it is negative. If, in a given match, one of the players experiences a shock that renders his payoff negative, then we assume that the match breaks up and both players become single. Note that, if the partnership remains matched, the price does not change. The latter implies that if i does not break up or re-match in the next period ($t+1$) his unperturbed payoff in $t+1$ will again be ϕ_i^t, and the period-t shock then has no influence on subsequent state transitions.

6.1 Stochastic stability

We are interested in the long-run behavior of the process when ϵ becomes small. We shall employ the concept of *stochastic stability* developed by Foster and Young (1990), Kandori et al. (1993) and Young (1993). In particular, we follow the analysis along the lines of 'one-shot stability' as recently introduced by Newton and Sawa (2013). Note that the perturbed process is ergodic for $\epsilon > 0$ and thus has a unique stationary distribution, say Π_ϵ, over the state space Ω. We are interested in $\lim_{\epsilon \to 0} \Pi_\epsilon = \Pi_0$.

Stochastic stability. A state $Z \in \Omega$ is *stochastically stable* if $\Pi_0(Z) > 0$. Denote the set of stochastically stable states by $\mathbf{S}$.

For a given parameter ϵ, denote the probability of transiting from Z to Z' in k periods by $\mathbb{P}_\epsilon^k[Z, Z']$. The *resistance* of a one-period transition $Z \to Z'$ is the unique real number $r(Z, Z') \geq 0$ such that $0 < \lim_{\epsilon \to 0} \mathbb{P}_\epsilon^1[Z, Z']/\epsilon^{r(Z,Z')} < \infty$. For completeness, let $r(Z, Z') = \infty$ if $\mathbb{P}_\epsilon^1[Z, Z'] = 0$. Hence a transition with resistance r has probability of the order $O(\epsilon^r)$. We shall call a transition $Z \to Z'$ (possibly in multiple periods) a *least cost transition* if it exhibits the lowest order of resistance. Formally, let $\{Z = Z_0, Z_1, \ldots, Z_k = Z'\}$ (k finite) describe a path of one-period transitions from Z to Z', then a least cost transition minimizes $\sum_{l=0}^{k-1} r(Z_l, Z_{l+1})$ over all such paths. Say that a non-core state Z' is in the basin of attraction of the core state Z if Z' is pairwise stable and the unique zero-resistance transition from Z' to a core state leads to Z. For a core state $Z \in \mathbf{C}$ we shall say that a transition out of the core is a *least cost deviation* if it minimizes the resistance among all transitions from Z to any non-core state which is not in the basin of attraction of Z.

Young (1993) shows that the computation of the stochastically stable states can be reduced to an analysis of rooted trees on the set of recurrent classes of the unperturbed dynamic. Define the *resistance* between two recurrent classes Z and Z', $r(Z, Z')$, to be the sum of resistances of least cost transitions that start in Z and end in Z'. Now identify the recurrent classes with the nodes of a graph. Given a node Z, a collection of directed edges T forms a *Z-tree* if, from every node $Z' \neq Z$, there exists a unique outgoing edge in T, Z has no outgoing edge, and the graph has no cycles.

[16] For simplicity we propose this specific distribution. But note that any probability distribution can be assumed as long as there exists a parameter ϵ such that $\mathbb{P}[k+1] = \epsilon \cdot O(\mathbb{P}[k])$ for all $k \in \mathbb{N}_0$.

Stochastic potential. The resistance $r(T)$ of a Z-tree T is the sum of the resistances of its edges. The *stochastic potential* of Z, $\rho(Z)$, is given by

$$\rho(Z) = \min\{r(T) : T \text{ is a } Z\text{-tree}\}. \tag{19}$$

Theorem 4 in Young (1993) states that the stochastically stable states are precisely those states where ρ is minimized.

6.2 Analysis

With this machinery at hand we shall show that the stochastically stable states are contained in the least core. To establish this result we shall adapt Newton and Sawa (2013)'s proof technique to show that the least core is the set of states which is most stable against one-shot deviations. We shall also provide conditions on the game under which the stochastically stable set is identical with the least core.

Recall that the least core consists of states that maximize the following term:

$$\begin{aligned} e^t_{\min} &= \min_{i:\, i \text{ matched}} \{\phi^t_i - \max_{j:\, a^t_{ij}=0} (\alpha_{ij} - \phi^t_j)_+\} && (20) \\ &= \min\{\underbrace{\min_{i,j:\, a^t_{ij}=0,\, i \text{ matched}} (\phi^t_i + \phi^t_j - \alpha_{ij})}_{=:\text{ case } A} \;;\; \underbrace{\min_{i:\, i \text{ matched}} \phi^t_i}_{=:\text{ case } B}\} && (21) \end{aligned}$$

Case A holds when the minimal cost deviation is such that two players who are currently not matched experience shocks such that a match of the two players may become profitable in the next period. Case B holds when the minimal cost deviation is such that a matched agent experiences a shock that renders his payoff negative, leading to a break-up of his match.

Given two states Z and Z^*, let the *distance* between them be

$$D(Z, Z^*) = \sum_{i \in F \cup W} |\phi_i - \phi^*_i|. \tag{22}$$

Lemma 2. *Given $Z^* \in \mathbf{L}$ and $Z \in \mathbf{C} \setminus \mathbf{L}$. Let Z' be a state not in the core which is reachable from Z by a least cost deviation. Then there exists $Z_1 \in \mathbf{C}$ such that $D(Z^*, Z_1) < D(Z^*, Z)$ and $\mathbb{P}^t_0[Z', Z_1] > 0$ for some $t \geq 0$.*

Proof. By Theorem 1, the recurrent classes consist of all singleton states in $\mathbf{C}_0 \subseteq \mathbf{C}$. Thus it suffices to limit our analysis to $Z^* \in \mathbf{L} \cap \mathbf{C}_0$ and $Z \in \mathbf{C}_0 \setminus \mathbf{L}$ since other core states have zero-resistance paths to the states in $\mathbf{C}_0$.

Case A. Suppose that the least-cost deviation to a non-core state is such that two (currently not matched with each other) players experience shocks, after which a match of the two may become profitable. That is, there exists i, matched to j, and a nonempty set J' such that i, j' is least costly to destabilize for any $j' \in J'$. Note that $d_i + d_{j'} - \alpha_{ij'}$

is minimal for all $j' \in J'$ and thus constant, and that $d_i + d_{j'} - \alpha_{ij'}$ is also non-negative since we are in a core state.

Case A.1. $d_i > d_i^*$.

Case A.1a. For all $j' \in J'$, $d_i + d_{j'} > \alpha_{ij'}$.

We can construct a sequence of transitions such that i reduces his aspiration level by δ, j increases his aspiration level by δ (note that we have $d_j < d_j^*$), and all other aspiration levels stay the same. Note that D then decreased and the resulting state is again a core state given that, for all $j' \neq j$, we started out with $d_i + d_{j'} > \alpha_{ij'}$.

We shall explain this sequence in detail. Suppose a shock occurs such that i reduces his aspiration level by at least δ and i and j' match at a price such that i's aspiration level does not increase. Consequently j and i' (j''s former partner if j' is matched in the core assignment) are now single. In the following period, i and j are profitable with positive probability. With positive probability, they match at a price such that d_i decreases by δ relative to the start of this sequence. Now i' and j' are both single. With positive probability, they reduce their aspiration levels and rematch at their previous price, returning to their original aspiration levels. Thus, with positive probability, the prices are set such that d_i decreases by δ, d_j increases by δ, and all other aspiration levels do not change. Hence D decreased and, given the earlier observation, the resulting state is again in the core because, now for all j', $d_i + d_{j'} \geq \alpha_{ij'}$ and all other inequalities still hold.

For the subsequent cases, we shall omit a description of the period-by-period transitions since they are conceptually similar.

Case A.1b. For all $j' \in J'$, $d_i + d_{j'} = \alpha_{ij'}$.

It follows that $d_{j'} < d_{j'}^*$, hence a δ-reduction of i's aspiration level and δ-increases by j and all $j' \in J'$ yield a reduction in D, and this leads to a core state.

Case A.2. $d_i = d_i^*$.

Since $Z \notin \mathbf{L}$ we must have $d_{j'} < d_{j'}^*$. Otherwise, given (i, j') is least costly to destabilize, we would have $Z \in \mathbf{L}$. But then j' must be matched in the core assignment and we have, for j''s partner i', that $d_{i'} > d_{i'}^*$. Hence a δ-reduction in i''s aspiration level and δ-increases by j' and all j'' for whom $d_{i'} + d_{j''} = \alpha_{i'j''}$ yield a reduction in D, and this leads to a core state.

Case A.3. $d_i < d_i^*$.

We have $d_j > d_j^*$ and a similar argument applies again. A δ-reduction in j's aspiration level and δ-increases by i and all i' for whom $d_{i'} + d_j = \alpha_{i'j}$ yield a reduction in D, and this leads to a core state.

Case B. Suppose that the least cost deviation to a non-core state is such that one player experiences a shock and therefore wishes to break up, that is, there exists i such that d_i is least costly to destabilize.

It follows that $d_i < d_i^*$, for otherwise $Z \in \mathbf{L}$ would constitute a contradiction.

Case B.1. For all $i' \neq i$, $d_{i'} + d_j > \alpha_{i'j}$.

Again, we can construct a sequence of transitions such that i increases his aspiration level by δ, and j reduces his aspiration level by δ. Note that D then decreased and the resulting state is again in the core given that, for all $i' \neq i$, we started out with $d_{i'} + d_j > \alpha_{i'j}$.

Now, we shall explain the sequence in detail. Suppose the shock occurs such that i turns single. Consequently, j turns too and, given that we are in a core state, (i', j) is not profitable for any $i' \neq i$. Therefore, if j encounters any $i' \neq i$, he will reduce his aspiration level. Now i can rematch with his optimal match j at a new price such that i can increase his aspiration level by δ while d_j decreases his by δ. (Note that, for the latter transition, it is crucial that any matched pair has match value at least δ.) Hence D decreased and, given the earlier observation, the resulting state is again in the core, since now for all i' $d_{i'} + d_j \geq \alpha_{i'j}$, and all other inequalities still hold.

Case B.2. There exists $I' \neq \emptyset$ and $i \notin I'$ such that for all $i' \in I'$, $d_{i'} + d_j = \alpha_{i'j}$.

Similar to case B.1 we can construct a sequence such that i increased his aspiration level by δ, j reduced his by δ, and all $i' \in I'$ increased their aspiration level by δ (which will only further reduce D). The resulting state is in the core. $\square$

Theorem 3. *The stochastically stable states are maximally robust to one-period deviations, and hence* $\mathbf{S} \subseteq \mathbf{L}$.

Proof. We shall prove the theorem by contradiction. Suppose there exists $Z^* \in \mathbf{S} \backslash \mathbf{L}$. Let T^* be a minimal cost tree rooted at Z^* and suppose that $\rho(Z^*)$ is minimal. Let $Z^{**} \in \mathbf{L}$. By lemma 2, together with the fact that the state space is finite, we can construct a finite path of least cost deviations between different core states such that their distance to a core state in $\mathbf{L}$ is decreasing:

$$Z^* \to Z_1 \to Z_2 \to \ldots \to Z_k = Z^{**} \tag{23}$$

Now we perform several operations on the tree T^* to construct a tree T^{**} for Z^{**}. First add the edges $Z_1 \to Z_2, \ldots, Z_{k-1} \to Z_k$ and remove the previously exiting edges from $Z_1, \ldots, Z_{k-1}$. Note that, since the newly added edges are all minimal cost edges, the sum of resistances does not increase. Next, let us add the edge $Z^* \to Z_1$ and delete the exiting edge from Z_k. Since $Z^* \notin \mathbf{L}$, it follows that $r(Z^* \to Z_1) < r(Z_k \to \cdot)$, and hence

$$\rho(Z^{**}) \leq \rho(Z^*) + r(Z^* \to Z_1) - r(Z_k \to \cdot) < \rho(Z^*). \tag{24}$$

This constitutes a contradiction. $\square$

We can formulate natural conditions under which the stochastically stable set coincides with the least core:

Well-connected. An assignment game is *well-connected* if, for any non-core state and for any player $i \in F \times W$, there exists a sequence of rematchings in the unperturbed process such that i is single at its end.

Rich. An assignment game with match values $\boldsymbol{\alpha}$ is *rich* if, for every player $i \in F$, there exists a player $j \in W$ such that (i, j) is never profitable, that is, $\alpha_{ij} = 0$.

Corollary 4. *Given a well-connected and rich assignment game with a unique optimal matching*[17]*, the set of stochastically stable states coincides with the least core, that is* $\mathbf{S} = \mathbf{L}$.

Proof. Given two recurrent classes of the process, $Z^*, Z^{**} \in \mathbf{C}_0$, and a non-core state, $Z \notin \mathbf{C}$, which is reachable from Z^* by a least cost deviation, we shall show that $r(Z, Z^{**}) = 0$. Suppose that Z^{**} has aspiration levels $\mathbf{d}^{**}$.

The idea of the proof is to construct a finite family of well-connected sequences, such that, after going through all transitions, players have aspiration levels less than or according to $\mathbf{d}^{**}$. Once we are in such a state Z^{**} can be reached easily.

Suppose we wish to make worker j_k single. By well-connectedness there exists a sequence of rematchings which makes j_k single at its end.[18] Suppose we have a minimal-length sequence among all sequences making j_k single, say $(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)$ if j_1 is matched and $j_1, (i_2, j_2), \ldots, (i_k, j_k)$ if j_1 is single. Note that it must hold that all players along the sequence (except potentially j_1) are currently matched for otherwise the sequence is not minimal. (This is the case since any single, by richness, reduces his aspiration level to zero and would then be a natural starting point for a shorter sequence as shall become clear below.) Further, by a similar observation, the sequence does not allow for profitable matches except for a match between j_1 and i_2. Finally, note that, for the sequence allowing j_k to become single at the end we must have that $d_{i_{l+1}} < \alpha_{i_{l+1} j_l}$ for $l = 1, \ldots, k-1$. This fact, together with the former comment, implies in particular that $d_{j_l} > 0$ for $l = 2, \ldots, k-1$.

Now suppose j_1 matches with i_2 (which is possible by assumption of well-connectedness) and j_2 becomes single. Suppose that the price is set such that d_{i_2} does not increase. Then, by richness, with positive probability j_2 reduces his aspiration level to 0. Suppose next that j_2 matches again with i_2 and the price is set such that j_2 keeps aspiration level 0. Now j_1 (and i_1 if j_1 was previously matched) is single and, by richness, with positive probability reduces his aspiration level to zero. Note that j_2 and i_3 are now profitable and they match with positive probability at a price such that d_{i_3} does not increase. Hence by iterating this sequence of transitions we arrive at a state where all players along the sequence are single except j_{k-1} and i_k who are now matched to each other. Suppose that they matched at a price such that j_{k-1} kept aspiration level 0. In particular note that j_k is single.

Next, we describe transitions such that the matching along the sequence considered remains the same (that is only (j_{k-1}, i_k) are matched) and the aspiration levels of j_{k-1} and i_k are such that $d_{j_{k-1}} \leq d^{**}_{j_{k-1}}$ and $d_{i_k} \leq d^{**}_{i_k}$. This can be achieved by matching i_{k-1} with j_{k-1} at a price such that j_{k-1} keeps aspiration level 0. Next, by richness, i_k

[17] Generically the optimal matching is unique. In particular this holds if the weights of the edges are independent, continuous random variables. Then, with probability 1, the optimal matching is unique.

[18] Note that the sequence naturally needs to alternate between firms and workers in order to make players single along the way.

reduces his aspiration level to zero. Finally j_{k-1} and i_k match again at a price such that the aspiration levels are less than or equal to their core aspiration levels in $\mathbf{d}^{**}$. (This can actually occur since otherwise the aspiration level vector $\mathbf{d}^{**}$ would not be pairwise stable, contradicting that Z^{**} is a core state.)

Now, by the well-connectedness assumption, we know that from any non-core state and for every player there exists such a sequence. Hence successively applying sequence after sequence such that each player is at its end once, we can conclude that there exists a path to a state such that, for all i, $d_i \leq d_i^{**}$. But note that some players may be matched.

Next, we have to show how Z^{**} is reached from the latter state. This is achieve if we successively match all (i, j) who are matched in Z^{**}, who are not matched yet, and for whom $d_i + d_j < \alpha_{ij}$ (they are profitable) at a price such that their new aspiration levels are d_i^{**}, d_j^{**}. (Here we need the assumption that the core matching is unique. Given our construction, players may be matched at the end of the sequences described above. In particular the core might already be reached with aspiration levels d^{**}. Now, if the optimal matching is not unique, we cannot guarantee that any particular matching is in place.) This leads to a state where aspiration levels are $\mathbf{d}^{**}$. Note that these aspiration levels are good. Further note that a reduction of the sum of aspiration levels will lead to a state which is not good. Cases 1a,b and 2b of the proof of Claim 2 of Theorem 1 can now be applied iteratively (Case 2a cannot hold, otherwise aspiration levels will no longer be good). These cases concern matchings only but do not change the aspiration levels. Hence eventually the desired core state Z^{**} is reached.

The proof summarizes as follows. We have shown that once the process is in a non-core state any core state can be reached. Hence the analysis of stochastic stability reduces to the resistance of exiting a core state. But this resistance is uniquely maximized by the states in the least core, which thus coincides with the set of stochastically stable states. □

6.3 Example

We shall illustrate the predictive power of our result for the 3×3 'house trade game' studied by Shapley and Shubik (1972). Let three sellers (w_1, w_2, w_3) and three buyers (f_1, f_2, f_3) trade houses. Their valuations are as follows:

Table 1: Seller and buyer evaluations.

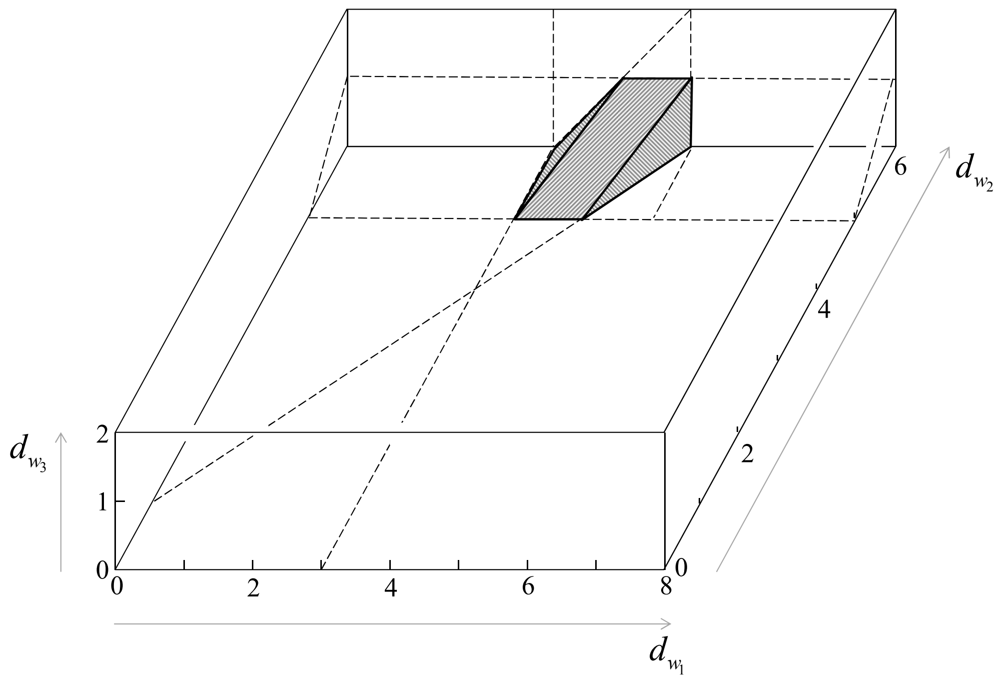
House j	Sellers willingness to accept $q_{1j}^- = q_{2j}^- = q_{3j}^-$	Buyers' willingness to pay p_{1j}^+	p_{2j}^+	p_{3j}^+
1	$18,000$	$23,000$	$26,000$	$20,000$
2	$15,000$	$22,000$	$24,000$	$21,000$
3	$19,000$	$21,000$	$22,000$	$17,000$

These prices lead to the following match values, α_{ij} (units of 1,000), where sellers are occupying rows and buyers columns:

$$\boldsymbol{\alpha} = \begin{pmatrix} 5 & \mathbf{8} & 2 \\ 7 & 9 & \mathbf{6} \\ \mathbf{2} & 3 & 0 \end{pmatrix} \tag{25}$$

The unique optimal matching is shown in bold numbers. Shapley and Shubik (1972) note that it suffices to consider the 3-dimensional imputation space spanned by the equations $d_{w_1} + d_{f_2} = 8$, $d_{w_2} + d_{f_3} = 6$, $d_{w_3} + d_{f_1} = 2$. Figure 8 illustrates the possible core allocations.

Figure 8: Imputation space for the sellers.

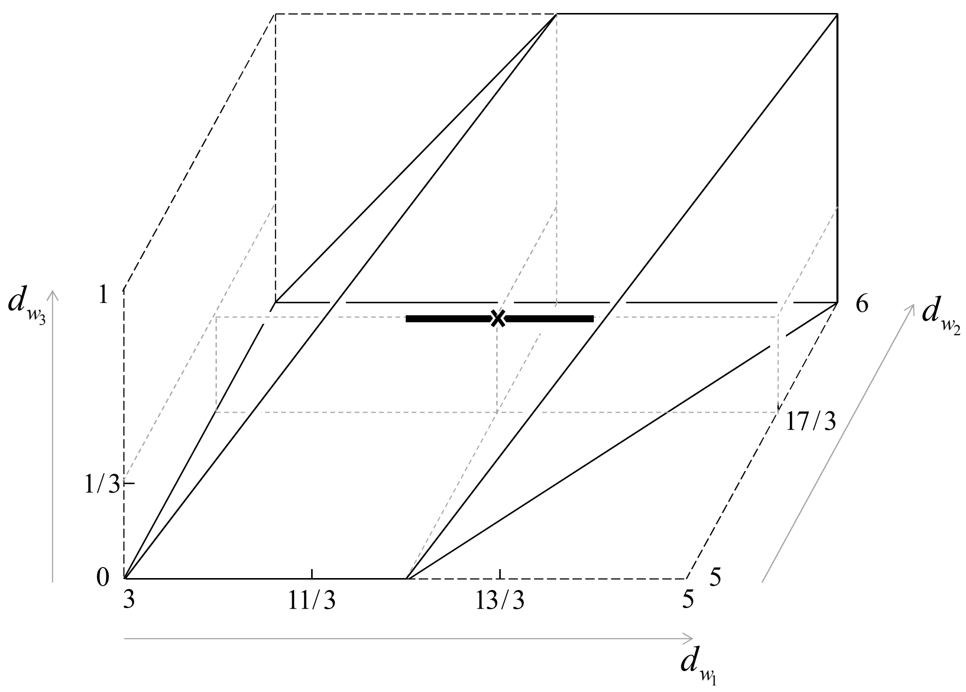


We shall now consider the least core, $\mathbf{L}$. Note that the particular states in $\mathbf{L}$ depend on the step size δ. We shall consider $\delta \to 0$ to best illustrate the core selection. By an easy calculation one finds that the states which are least vulnerable to one-period deviations are such that

$$d_{w_1} \in [11/3, 13/3],\ d_{w_2} = 17/3,\ d_{w_3} = 1/3. \tag{26}$$

The minimal excess in the least core is $e_{\min} = 1/3$. The bold line in figure 9 illustrates the set $\mathbf{L}$. The nucleolus, $d_{w_1} = 4$, $d_{w_2} = 17/3$, $d_{w_3} = 1/3$, is indicated by a cross. (One can verify, that here the kernel coincides with the nucleolus.)

Figure 9: Core selection for the sellers.



7. Conclusion

In this paper we have shown that agents in large decentralized matching markets can learn to play equitable core outcomes through simple trial-and-error learning rules. We assume that agents have no information about the distribution of others' preferences, about their past actions and payoffs, or about the value of different matches. The unperturbed process leads to the core with probability one but no authority 'solves' an optimization problem. Rather, a path into the core is discovered in finite time by a random sequence of adjustments by the agents themselves. This result is similar in spirit to that of Chen et al. (2011), but in addition our process selects equitable outcomes within the core. In particular, the stochastically stable states of the perturbed process are contained in the least core, a subset of the core that generalizes the nucleolus for assignment games. This result complements the stochastic stability analysis of Newton and Sawa (2013) in ordinal matching and of Newton (2012) in coalitional games. It is an open problem to extend the analysis to more general classes of cooperative games and matching markets.

References

Agastya, M. (1997), "Adaptive play in multiplayer bargaining situations." *The Review of Economic Studies*, 64, 411–426.

Agastya, M. (1999), "Perturbed adaptive dynamics in coalition form games." *Journal of Economic Theory*, 89, 207–233.

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

Balinski, M. L. (1965), "Integer programming: Methods, uses, computation." *Management Science*, 12, 253–313.

Balinski, M. L. and D. Gale (1987), "On the core of the assignment game." *Functional Analysis, Optimization and Mathematical Economics*, 1, 274–289.

Bayati, M., C. Borgs, J. Chayes, Y. Kanoria, and A. Montanari (2014), "Bargaining dynamics in exchange networks." *Journal of Economic Theory*, forthcoming.

Biró, P., M. Bomhoff, P. A. Golovach, W. Kern, and D. Paulusma (2012), "Solutions for the stable roommates problem with payments." *Lecture Notes in Computer Science*, 7551, 69–80.

Bush, R. and F. Mosteller (1955), *Stochastic Models of Learning*. Wiley.

Chen, B., S. Fujishige, and Z. Yang (2011), "Decentralized market processes to stable job matchings with competitive salaries." KIER Working Papers 749, Kyoto University.

Chung, K.-S. (2000), "On the existence of stable roommate matching." *Games and Economic Behavior*, 33, 206–230.

Crawford, V. P. and E. M. Knoer (1981), "Job matching with heterogeneous firms and workers." *Econometrica*, 49, 437–540.

Davis, M. and M. Maschler (1965), "The kernel of a cooperative game." *Naval Research Logistic Quarterly*, 12, 223–259.

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

Demange, G., D. Gale, and M. Sotomayor (1986), "Multi-item auctions." *Journal of Political Economy*, 94, 863–872.

Diamantoudi, E., L. Xue, and E. Miyagawa (2004), "Random paths to stability in the roommate problem." *Games and Economic Behavior*, 48, 18–28.

Driessen, T. S. H. (1998), "A note on the inclusion of the kernel in the core for the bilateral assignment game." *International Journal of Game Theory*, 27, 301–303.

Driessen, T. S. H. (1999), "Pairwise-bargained consistency and game theory: the case of a two-sided firm." *Fields Institute Communications*, 23, 65–82.

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

Foster, D. and H. P. Young (1990), "Stochastic evolutionary game dynamics." *Theoretical Population Biology*, 38, 219–232.

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

Gale, D. and L. S. Shapley (1962), "College admissions and stability of marriage." *American Mathematical Monthly*, 69, 9–15.

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

Hart, S. and A. Mas-Colell (2003), "Uncoupled dynamics do not lead to Nash equilibrium." *American Economic Review*, 93, 1830–1836.

Hart, S. and A. Mas-Colell (2006), "Stochastic uncoupled dynamics and Nash equilibrium." *Games and Economic Behavior*, 57, 286–303.

Heckhausen, H. (1955), "Motivationsanalyse der Anspruchsniveau-Setzung." *Psychologische Forschung*, 25, 118–154.

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

Hoppe, F. (1931), "Erfolg und Mißerfolg." *Psychologische Forschung*, 14, 1–62.

Huberman, G. (1980), "The nucleolus and the essential coalitions." *Analysis and Optimization of Systems, Lecture Notes in Control and Information Systems*, 28, 417–422.

Inarra, E., C. Larrea, and E. Molis (2008), "Random paths to P-stability in the roommate problem." *International Journal of Game Theory*, 36, 461–471.

Inarra, E., C. Larrea, and E. Molis (2013), "Absorbing sets in roommate problem." *Games and Economic Behavior*, 81, 165–178.

Jackson, M. O. and A. Watts (2002), "The evolution of social and economic networks." *Journal of Economic Theory*, 106, 265–295.

Kandori, M., G. J. Mailath, and R. Rob (1993), "Learning, mutation, and long run equilibria in games." *Econometrica*, 61, 29–56.

Kelso, A. S. and V. P. Crawford (1982), "Job matching, coalition formation, and gross substitutes." *Econometrica*, 50, 1483–1504.

Klaus, B. and F. Klijn (2007), "Paths to stability for matching markets with couples." *Games and Economic Behavior*, 58, 154–171.

Klaus, B., F. Klijn, and M. Walzl (2010), "Stochastic stability for roommates markets." *Journal of Economic Theory*, 145, 2218–2240.

Klaus, B. and F. Payot (2013), "Paths to stability in the assignment problem." Working paper.

Kojima, F. and M. U. Ünver (2008), "Random paths to pairwise stability in many-to-many matching problems: a study on market equilibrium." *International Journal of Game Theory*, 36, 473–488.

Kuhn, H. W. (1955), "The Hungarian Method for the assignment problem." *Naval Research Logistic Quarterly*, 2, 83–97.

Llerena, F. and M. Nunez (2011), "A geometric characterization of the nucleolus of the assignment game." *Economics Bulletin*, 31, 3275–3285.

Llerena, F., M. Nunez, and C. Rafels (2012), "An axiomatization of the nucleolus of the assignment game." Working Papers in Economics 286, Universitat de Barcelona.

Marden, J. R., H. P. Young, G. Arslan, and J. Shamma (2009), "Payoff-based dynamics for multi-player weakly acyclic games." *SIAM Journal on Control and Optimization*, 48, 373–396.

Maschler, M., B. Peleg, and L. S. Shapley (1979), "Geometric properties of the kernel, nucleolus, and related solution concepts." *Mathematics of Operations Research*, 4, 303–338.

Nash, J. (1950), "The bargaining problem." *Econometrica*, 18, 155–162.

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

Newton, J. (2012), "Recontracting and stochastic stability in cooperative games." *Journal of Economic Theory*, 147, 364–381.

Newton, J. and R. Sawa (2013), "A one-shot deviation principle for stability in matching problems." Economics Working Papers 2013-09, University of Sydney, School of Economics.

Nunez, M. (2004), "A note on the nucleolus and the kernel of the assignment game." *International Journal of Game Theory*, 33, 55–65.

Pradelski, B. S. R. (2014), "Evolutionary dynamics and fast convergence in the assignment game." Department of Economics Discussion Paper Series 700, University of Oxford.

Pradelski, B. S. R. and H. P. Young (2012), "Learning efficient Nash equilibria in distributed systems." *Games and Economic Behavior*, 75, 882–897.

Rochford, S. C. (1984), "Symmetrically pairwise-bargained allocations in an assignment market." *Journal of Economic Theory*, 34, 262–281.

Roth, A. E. and I. Erev (1995), “Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term.” *Games and Economic Behavior*, 8, 164–212.

Roth, A. E. and M. Sotomayor (1992), “Two-sided matching.” *Handbook of Game Theory with Economic Applications*, 1, 485–541.

Roth, A. E. and H. Vande Vate (1990), “Random paths to stability in two-sided matching.” *Econometrica*, 58, 1475–1480.

Rozen, K. (2013), “Conflict leads to cooperation in Nash Bargaining.” *Journal of Economic Behavior and Organization*, 87, 35–42.

Sauermann, H. and R. Selten (1962), “Anspruchsanpassungstheorie der Unternehmung.” *Zeitschrift fuer die gesamte Staatswissenschaft*, 118, 577–597.

Sawa, R. (2011), “Coalitional stochastic stability in games, networks and markets.” Working Paper, Department of Economics, University of Wisconsin-Madison.

Schmeidler, D. (1969), “The nucleolus of a characteristic function game.” *SIAM Journal of Applied Mathematics*, 17, 1163–1170.

Selten, R. (1998), “Aspiration adaptation theory.” *Journal of Mathematical Psychology*, 42, 191–214.

Selten, R. and R. Stoecker (1986), “End behavior in sequences of finite prisoner’s dilemma supergames: A learning theory approach.” *Journal of Economic Behavior and Organization*, 7, 47–70.

Shapley, L. S. and M. Shubik (1963), “The core of an economy with nonconvex preferences.” *The Rand Corporation*, 3518.

Shapley, L. S. and M. Shubik (1966), “Quasi-cores in a monetary economy with nonconvex preferences.” *Econometrica*, 34.

Shapley, L. S. and M. Shubik (1972), “The assignment game 1: The core.” *International Journal of Game Theory*, 1, 111–130.

Solymosi, T. and T. E. S. Raghavan (1994), “An algorithm for finding the nucleolus of assignment games.” *International Journal of Game Theory*, 23, 119–143.

Sotomayor, M. (2003), “Some further remarks on the core structure of the assignment game.” *Mathematical Social Sciences*, 46, 261–265.

Thorndike, E. (1898), “Animal intelligence: An experimental study of the associative processes in animals.” *Psychological Review*, 8, 1874–1949.

Tietz, A. and H. J. Weber (1972), “On the nature of the bargaining process in the Kresko-game.” *H. Sauermann (Ed.) Contribution to experimental economics*, 7, 305–334.

Young, H. P. (1993), “The evolution of conventions.” *Econometrica*, 61, 57–84.

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