Heinrich H. Nax

Equity dynamics in bargaining without information exchange

Article (Accepted version) (Refereed)

Original citation:
Nax, Heinrich H. (2015) *Equity dynamics in bargaining without information exchange.* Journal of Evolutionary Economics, 25 (5). pp. 1011-1026. ISSN 0936-9937

DOI: 10.1007/s00191-015-0405-9

Equity dynamics in bargaining without information exchange

Heinrich H. Nax[1]

Abstract
In this paper, completely uncoupled dynamics for n-player bargaining are proposed that mirror key behavioral elements of early bargaining and aspiration adjustment models (Zeuthen 1930, Sauermann & Selten 1962). Individual adjustment dynamics are based on directional learning adjustments, solely driven by histories of own realized payoffs. Bargaining this way, all possible splits have positive probability in the stationary distribution of the process, but players will split the pie almost equally most of the time. The expected waiting time for almost equal splits to be played is quadratic.

JEL classifications: C71, C73, C78, D83

Keywords: bargaining, cooperative game theory, equity, evolutionary game theory, (completely uncoupled) learning

[1]Department of Social Sciences, ETH Zürich, Clausiusstrasse 37-C3, 8092 Zürich, Switzerland. hnax@ethz.ch. Nax acknowledges support by the European Commission through the ERC Advanced Investigator Grant 'Momentum' (Grant No. 324247).

1. Introduction

Bargaining models are amongst the most important applications of game theory, spanning cooperative, noncooperative, evolutionary and experimental games. The most basic one is bilateral bargaining. Indeed, Ellingsen (1997) asks the question: *Is there any economic activity more basic than two people dividing a pie?* Dating back to Zeuthen (1930), Raiffa (1953), Luce & Raiffa (1957), Schelling (1956) and Rubinstein (1982), bilateral bargaining has been modelled as some kind of "power struggle". The proposed procedures mirror adjustments driven by admixtures of patience, threats, and/or rounds of offers and counteroffers with subsequent compromise.[2]

In this paper, we focus on infinitely repeated multilateral bargaining in a homogeneous population that takes place in an informational setting characterized by the absence of information concerning other players' utility functions, actions and payoffs. We propose an evolutionary model of bargaining without information exchange. The game-theoretic full-rationality canon is of course not germane in such an environment (Young 2004), and not even standard evolutionary models can be applied (Weibull 1995, Sandholm 2010). Other than in the dynamic bargaining models of Zeuthen (1930), Raiffa (1953), Luce & Raiffa (1957), Schelling (1956) and Rubinstein (1982), the pie is not just split once at the end of the bargaining process but repeatedly each round. Agents repeatedly demand slices of the pie without information about others' demands. Agents receive their slices when demands are globally feasible, but receive nothing when not. Without individuals going through a process of hypothesis-formation concerning other agents' actions, the model that we propose is easy as pie (pun intended):

- *an agent whose previous demand was feasible occasionally demands incrementally more, while an agent whose previous demand was infeasible reduces his demand with a probability that is increasing in his demand-payoff differential.*

Bargaining this way, all possible n-way splits of the pie have positive probability in the stationary distribution of the process, but players end up sharing the pie almost equally most of the time. Indeed, from any initial state, an almost equal split is reached in quadratic time. Equity here refers to players receiving the same (or very similar) payoffs in the long-run outcomes.[3] In our setting, due to the homogeneity of the population, the multilateral generalizations of the aforementioned, standard bargaining solutions (due to Zeuthen's, Nash's, or Rubinstein's) all coincide with this allocation. The aim of this paper is to illustrate another dynamic with which it may be reached, the distinguishing factor being the informational limits of the environment.

Even though our model of bargaining is evolutionary in a finite population, that is, the pie is split repeatedly by the same agents, our bargaining dynamics are different from the

[2] Axiomatic bargaining solutions such as Nash's (1950) explicitly consider relative bargaining strengths ("outside options"). Harsanyi (1956) shows that the solutions obtained by Zeuthen's dynamic model and Nash's axiomatization coincide.

[3] The number of adjustments needed to reach such outcomes may not be the same for all players, hence there is ground to think of some inequality in terms of bargaining efforts or more general concepts of social exchange equity (Adams 1965) depending on initial states.

standard bargaining models of this kind which we shall discuss shortly. In terms of the underlying dynamic adjustment components, our dynamics have closer antecedents in the iterative bargaining model of Zeuthen (1930) where the pie is split only once at the end of the process. In Zeuthen, bargaining starts with both parties demanding the entire pie.[4] Over time, bargaining ensues as a sequence of mutual concessions that are based on the two parties' relative willingness to risk conflict: at any infeasible intermediate proposal, the party with the lower willingness to accept breakdown, which (in the symmetric case) is the party with the higher demand, adjusts its demand to a slightly smaller demand. Concessions alternate in this way until feasible demands are made.[5] Then bargaining ends. Formally, our dynamics are a probabilistic interpretation of Zeuthen's model with repeated consumption of the pie, but the underlying behavioral motivations are also motivated differently. In Zeuthen, the party with the lower willingness to risk breakdown concedes precisely because she judges her opponent's willingness to risk breakdown to be higher (probably by interpretation of her past actions). By contrast, our model assumes that demand concessions are triggered by own demand-payoff differentials and past experiences, without hypotheses made about others.

Our individual adjustment dynamics do not rely on information about others. Instead, decisions are solely based on the histories of own realized payoffs. This means that our dynamics are "completely uncoupled" (Foster & Young 2006, Young 2009) from others' actions and payoffs. "Completely uncoupled" learning tightens the informational constraints of "uncoupled" learning (Hart & Mas-Colell, 2003, 2006), which may depend on others' past actions.[6] Completely uncoupled rules have recently been applied to noncooperative games by Karandikar et al. (1998), Foster & Young (2006), Germano & Lugosi (2007), Young (2009), Marden et al. (2011), Pradelski & Young (2012), Babichenko (2012), Marden et al. (2014), and to cooperative games and matching models by Nax (2011), Nax et al. (2013). These models have antecedents in classic learning theory dating back to Thorndike (1898), Hoppe (1931), Estes (1950), Heckhausen (1955), Herrnstein (1961) and Sauermann & Selten (1962). Reinforcement learning models (Bush & Mosteller 1955, Suppes & Atkinson 1959, Harley 1981, Cross 1983, Roth & Erev 1995, Erev & Roth 1998) are a particularly famous class of completely uncoupled learning dynamics.

Our dynamics are most closely related to the theory of aspiration adjustment due to Heckhausen (1955) and Sauermann & Selten (1962). The particular learning heuristic we adopt is based on "directional learning" (Selten & Stoecker 1986, Selten & Buchta 1998). According to the directional learning hypothesis of bargaining, agents demand either more or less dependent on whether previous demands were successful or not. This hypothesis was tested extensively in (bilateral) experiments by scientists surrounding the

[4]Other iterative bargaining procedures such as Raiffa (1953), Luce & Raiffa (1957), Kalai (1977), John & Raith (1999) start from inside the bargaining set. The differences between these approaches and ours is similar in spirit to the differences with Zeuthen that are discussed in detail here.

[5]In Raiffa (1953), Luce & Raiffa (1957), Kalai (1977), John & Raith (1999), the process moves the other way around and iterative steps towards the Pareto frontier are negotiated.

[6]See Babichenko (2010, 2012) for convergence comparisons of uncoupled and completely uncoupled dynamics.

theory's main proponents at the time (e.g. Tietz et al., 1978).[7] In fact, we could restate our adjustment dynamics with their words (Tietz et al. 1978; pp. 91, 94):

- *"the basis of the aspiration levels changes according to the economic situation and is modified by success and failure in the previous negotiation."*
- *"a subject lowers his aspiration level after a negative impulse. It is not lowered if the impulse is positive. After a neutral impulse the aspiration level is kept stable."*
- *"a subject raises his aspiration level after a positive impulse. It is not raised if the impulse is negative. After a neutral impulse the aspiration level is kept stable."*

What is new about our take on directional learning is our re-interpretation as a completely uncoupled dynamic. In the standard formulation (see, for example, Grosskopf 2003), players learn directionally because they have knowledge of counterfactuals, that is, they can assess how strategies in either direction would have performed relative to the strategy that was actually chosen. Here, we do not require such knowledge. Instead, directionality is born from the fact that players have a tendency to demand more (less) when currently receiving a payoff that matches or exceeds (falls short of) their aspirations. A similar approach has recently been taken by Nax et al. (2013), Nax and Perc (2015), Nax and Pradelski (2015), and Burton-Chellew et al. (2015).

A particular feature of these dynamics is that, after a negative impulse, an agent reduces his demand with a probability that is increasing in his demand-payoff differential. This is a phenomenon observed regularly in the aforementioned experiments (e.g. Tietz et al. 1978). More recently, experiments by Ding & Nicklisch (2013), Nax et al. (2013), Burton-Chellew et al. (2015) also provide non-bargaining evidence for this phenomenon. Nax et al. (2015), for example, consider the context of voluntary contributions games played in an experimental setting where information is neither revealed about the structure of the game nor about other players' actions and payoffs. The predominant type of adjustments identified in their study is directional in our sense, and indeed much more accentuated after negative stimuli than after positive ones. This finding is confirmed in Burton-Chellew et al. (2015), and indeed found to be a robust feature even in environments where more information is available and against competing hypotheses.[8] This suggests that negative stimuli regularly have a more immediate effect than positive stimuli, the impact of which depends on the size of the shock. This feature of our dynamic, more generally, relates to asymmetric reactions to perceived gains and losses that also lie at the heart of several of the recently proposed completely uncoupled, trial-and-error learning models (in particular, in Young 2009, Marden et al. 2011, Pradelski & Young 2012).

As mentioned previously, the differences between our approach and traditional evolutionary bargaining models, as for example in Young (1993), Ellingsen (1997), Alexander & Skyrms (1999), Saez-Marti & Weibull (1999), Binmore et al. (2003), are substantial in

[7]See also Tietz & Weber (1972), Tietz (1975), Weber (1976), Tietz et al. (1978), Tietz & Bartos (1983), Crössmann & Tietz (1983), Tietz et al. (1988). Roth (1995) discusses subsequent experiments.

[8]Actually, such directional adjustments may turn out to be strategically rationalizable in these higher information environments. (I thank an anonymous referee for pointing this out.)

terms of their behavioral and informational assumptions.[9] Take Young's (1993) model, for instance, where random pairs of agents from finite populations are repeatedly drawn to play the Nash demand game. Each player in such an interaction randomly samples demands from previous bargaining encounters and plays a best reply to his sample with high probability, but there is small probability of "noise", that is, players commit errors with small probability. The analysis of "stochastic stability" (Foster & Young 1990, Young 1993) reveals which bargaining outcomes are long-run stable as the "noise" rate goes to zero in such a dynamic.[10] Under some regularity assumptions, stochastic stability analysis reveals that the population evolves to play of the Nash bargaining solution. By contrast, our model is not based on best-reply dynamics, and its implicit noise rates do not diminish. Hence, we formally do not use the concept of stochastic stability, but a zonal notion of convergence instead.[11] Indeed, agents do not reply to others because they have no information about them. Instead, agents adjust their behavior based on own experience and continue to experiment with their own actions at fixed rates ad infinitum. In this paper, we propose an intuitive model motivated by experimental evidence as to how this is done and explore its convergence properties.

The paper is structured as follows. Next, we introduce the model's static and dynamic components. Section 3 contains the paper's convergence results. Section 4 concludes.

2. The model

2.1. Static components

The following n-player extension of the Nash demand game is played.

n-player cooperative transferable-utility bargaining. *A fixed population of players, $N = \{1, ..., n\}$, bargains over the unit pie. $G(v, N)$ is the cooperative bargaining game with characteristic function $v : 2^n \to \mathbf{R}$ such that subcoalitions are inessential ($v(S) = v(\emptyset) = 0$ for all $S \subset N$), and the grand coalition produces the unit pie (i.e. $v(N) = 1$).*

Demands. Each player $i \in N$ makes a demand $d_i \in [0, 1]$ of the unit pie. We assume that, for some $k \in \mathbb{N}^+$, each d_i is a multiple of some discrete stepsize $\delta = 1/nk$. Write $\mathbf{d} = \{d_1, ..., d_n\}$ for a demand vector, and $\mathbf{\Omega}$ for the (finite) set of possible demand vectors.

Payoffs. If demands are jointly feasible, each player receives his demand; otherwise,

[9] See also Gale et al. (1995), Nowak et al. (2000), Konrad & Morath (2015) for evolutionary models of "ultimatum bargaining" (Güth et al. 1982), or Binmore et al. (1998) for an evolutionary analysis of alternating-offer "Rubinstein bargaining" (Rubinstein 1982).

[10] "Stochastic stability" is an equilibrium refinement that is different from "evolutionary stability" based on replicator arguments (Maynard Smith & Price 1973, Maynard Smith 1974) or from "evolutionary stability" in finite populations (Schaffer 1988, Nowak et al. 2004).

[11] The difference between these convergence concepts is addressed in more detail in Young (2009), see also Babichenko (2012)

individuals receive zero. For any player $i \in N$ at any time t, his payoff is

$$\phi_i = \begin{cases} d_i & \text{if } \sum_{i \in N} d_i \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Write $\boldsymbol{\phi}$ for a vector of payoffs $\{\phi_1, ..., \phi_n\}$.

2.2. Dynamic components

The process moves in infinite continuous time. Players are "activated" by independent Poisson clocks at rate one.[12] Define a "time step" t by activation of a unique agent, the uniqueness of which is given by the independence of the Poisson clocks. A new bargaining game is played every time a new time step t begins.

Let $\mathbf{d}^t$ describe agents' demands at time t. For all $j \neq i$ not activated at time $t+1$, j remains inactive and continues with his previous demand $d_j^{t+1} = d_j^t$. For the activated agent, we assume the following demand adjustments. Recall that agents crucially have no information about other agents' demands or payoffs.

Increases. If $\sum_{j \in N} d_j^t \leq 1$, then, if $d_i^t < 1$,

$$d_i^{t+1} = \begin{cases} d_i^t + \delta & \text{with probability } r, \\ d_i^t & \text{otherwise.} \end{cases} \tag{1}$$

We assume that $r \in (0, 1)$, subsequently referred to as the *rate of experimentation*, is constant. If $\sum_{j \in N} d_j^t \leq 1$ and $d_i^t = 1$, then we assume $d_i^{t+1} = d_i^t$ with probability one.

Reductions. If $\sum_{j \in N} d_j^t > 1$, then

$$d_i^{t+1} = \begin{cases} d_i^t & \text{with probability } s(d_i^t), \\ d_i^t - \delta & \text{otherwise.} \end{cases} \tag{2}$$

We assume $s(\cdot)$, subsequently referred to as the *degree of stickiness*, to be a time-invariant linear function, constant for all players, and of the form $1 - s(d_i^t) = ad_i^t$ with $0 < a < 1$.[13] For convenience, we shall define $f(\cdot) = 1 - s(\cdot)$. Notice that $d_i^{t+1} = d_i^t$ with probability one if $d_i^t = 0$. Furthermore, in line with the empirical observation mentioned in the introduction that agents react stronger to negative than to positive stimuli, we shall assume that $r < a\delta$, i.e. that any reduction is more likely than an increase.[14]

---

[12] It will be convenient to have set up the process with these Poisson clocks when we turn to convergence times. For the meantime, it is also possible to think of agents being activated uniformly at random in discrete time.

[13] The linear function is an approximation for more general functions or a lower bound for functions that first-order dominate the linear bound (e.g. more convex or step functions). Using ad_i with $a = \frac{f(\delta)}{\delta}$ for any convex function $f(\cdot)$ with $f(0) = 0$, $f'(x) > 0$ and $f''(x) \geq 0$ for all $x > 0$, for example, "understates" the stickiness and works in the opposite direction in terms of our results.

[14] Assuming $r < a\delta$ guarantees that this assumption holds for any current $d_i^t > 0$ of any player.

3. Analysis

3.1. Recurrence class

The state of the process at any time t is described by $\mathbf{d}^t$, which implies time-t utilities for all players and also the probabilities for the time-$(t+1)$ Markov transitions (expressions 1 and 2), thus yielding a Markov chain on $\mathbf{\Omega}$. In this section, we shall show that all states with less than efficient demands and all states with demands that are infeasible by more than δ are transient, all other states are recurrent. We shall refer to state $\mathbf{d}'$ as a *neighbor* of any given state $\mathbf{d}$ if $\mathbf{d}'$ is reached with positive probability in period $t+1$ if the period-t state is $\mathbf{d}$.

Definitions. A state $\mathbf{d} \in \mathbf{\Omega}$ is *transient* if, given $\mathbf{d}^t = \mathbf{d}$ at any time t, there exists a positive probability that the process never returns to $\mathbf{d}$ at any time $t' > t$. State $\mathbf{d}$ is *recurrent* if it is not transient.

Proposition 1. *Any state* $\mathbf{d} \in \mathbf{\Omega}$ *with* $\sum_{i\in N} d_i < 1$ *or* $> 1+\delta$ *is transient. All states* $\mathbf{d} \in \mathbf{\Omega}$ *with* $\sum_{i\in N} d_i \in [1, 1+\delta]$ *are recurrent.*

Proof of proposition 1. Transience: At t, suppose $\mathbf{d}^t$ is such that $\sum_{i\in N} d_i^t \leq 1$. Starting at $\mathbf{d}^t$, the process exits with a positive probability in an "outwards" direction (to larger demands), but not "inwards" (to smaller demands). The direct neighbors of all states with $\sum_{i\in N} d_i^t < 1$ have $\sum_{i\in N} d_i^t \leq \sum_{i\in N} d_i^{t+1} \leq 1$, the states on the frontier with $\sum_{i\in N} d_i^t = 1$ have neighbors with $1 \leq \sum_{i\in N} d_i^{t+1} \leq 1+\delta$.

At t, suppose $\mathbf{d}^t$ is such that $\sum_{i\in N} d_i^t > 1$. Starting at $\mathbf{d}^t$, the process exits with a positive probability in an inwards direction, but not outwards. The direct neighbors of all states with $\sum_{i\in N} d_i^t > 1+\delta$ have $\sum_{i\in N} d_i^t \geq \sum_{i\in N} d_i^{t+1} > 1$, whereas a state with $\sum_{i\in N} d_i^t = 1+\delta$ is the neighbor of states with $1 \leq \sum_{i\in N} d_i^{t+1} \leq 1+\delta$.

Jointly, these observations imply that all states $\mathbf{d}$ with $\sum_{i\in N} d_i < 1$ (and $\sum_{i\in N} d_i > 1+\delta$) are transient because the process exits these states with a positive probability in an outward (inward) direction but, once left, they are never again reached.

Recurrence: Any recurrent state $\mathbf{d}$ is such that $\sum_{j\in N} d_j \in [1, 1+\delta]$.

Claim. There exist positive-probability transitions between any two recurrent states $\mathbf{d}$, $\mathbf{d}'$.

The claim follows directly from the following two observations.

- given $\mathbf{d^t} = \mathbf{d}$ with $\sum_{i\in N} d_i = 1$, the probability that $d_i^{t+1} = d_i^t + \delta$ for any $i \in N$ and $d_j^{t+1} = d_j^t$ for all $j \neq i$ is $r/n > 0$ if $d_i^t < 1$.
- given $\mathbf{d^t} = \mathbf{d}$ with $\sum_{i\in N} d_i = 1+\delta$, the probability that $d_i^{t+1} = d_i^t - \delta$ for any $i \in N$ and $d_j^{t+1} = d_j^t$ for all $j \neq i$ is at least $a\delta/n > 0$ if $d_i^t > 0$.

The two transitions can be used to reallocate any number of δs from any player demanding a positive amount, via any player, to any player demanding less than one in all $\mathbf{d} \in \mathbf{\Omega}$: $\sum_{i\in N} d_i \in [1, 1+\delta]$. $\square$

3.2. Embedded Pareto frontier chain

Denote by $\mathbf{\Omega}^e \subset \mathbf{\Omega}$ the states on the embedded chain of states $\mathbf{d}$ on the Pareto frontier (with $\sum_{i \in N} d_i = 1$). $\mathbf{d}'$ is a *neighbor* of any given state $\mathbf{d}$ in this chain if there exists a state $\mathbf{d}''$ not in the embedded chain such that, in the original chain, $\mathbf{d}''$ is a neighbor of $\mathbf{d}$, and $\mathbf{d}'$ is a neighbor of $\mathbf{d}''$ (i.e. $\mathbf{d}'$ is reachable from $\mathbf{d}$ in two time steps in the original chain). Recall that all states in $\mathbf{\Omega}^e$ are recurrent (proposition 1). Moreover, given any state $\mathbf{d} = (d_1, ..., d_n)$, his neighbors $\mathbf{d}_{ij}$ are of the form $(d_1, ..., d_i + \delta, ..., d_j - \delta, ..., d_n)$: i.e. between neighbouring states, $\mathbf{d}$ and $\mathbf{d}_{ij}$, in $\mathbf{\Omega}^e$ a single transfer takes place; first some player i increases his demand to $d_i + \delta$ (causing infeasibility), then some player $j \neq i$ (demanding > 0) reduces his demand to $d_j - \delta$ (restoring feasibility); all other demands remain at their previous levels. The probability of any feasible transition between any two neighbors, $\mathbf{d}$ and $\mathbf{d}_{ij}$, in $\mathbf{\Omega}^e$ is

$$\pi_{\mathbf{d}\mathbf{d}_{ij}} = \frac{1}{n} r \cdot \frac{1}{n} f(d_j). \tag{3}$$

We will view these transitions in $\mathbf{\Omega}^e$ as single time steps indicated by times with hats ($\widehat{t} = \widehat{1}, \widehat{2}, ...$). Note that these take at least two time steps in $\mathbf{\Omega}$ but may take longer if, for example, an agent demanding one is drawn on the Pareto frontier, an agent demanding zero is drawn above the Pareto frontier, etc.

3.3. Equity

Next, we shall prove that almost equal splits will be played most of the time.

Before we turn to the mathematical results, let us state the basic intuition behind this result which is best-illustrated in bilateral bargaining. (Figure 1 illustrates.) The reader should note, however, that despite the fact that bilateral bargaining is a useful (graphic) illustration of our dynamics, the same arguments do not carry over trivially to multilateral bargaining.

Suppose two players bargain over the unit-pie. If $d_1 + d_2 \leq 1$, both players receive the shares they respectively demand. At the next time step, both players are equally likely to increase their demand by δ if both demand less than one. If $d_1 + d_2 > 1$, both players receive zero. At the next time step, the player currently demanding a higher share of the pie is more likely to reduce. Eventually (by proposition 1), this increase-decrease dynamic will boil down to an ongoing process that moves on (or one δ above) the Pareto frontier; again and again, one of two transitions occur: (i) one of the two players overshoots the Pareto frontier by δ; then (ii) one of the two players (more likely the one with the higher demand) reduces by δ. Over time, this leads to equal splits.

3.4. Equity drift

To prove convergence, we track the variance of demands in the embedded chain $\mathbf{\Omega}^e$. Note that, in the recurrent class, payoffs equal demands if demands are feasible, and payoffs are zero when demands are infeasible. The variance of payoffs, too, is therefore equal to

Figure 1: Bilateral bargaining with linear boundary of slope -1.

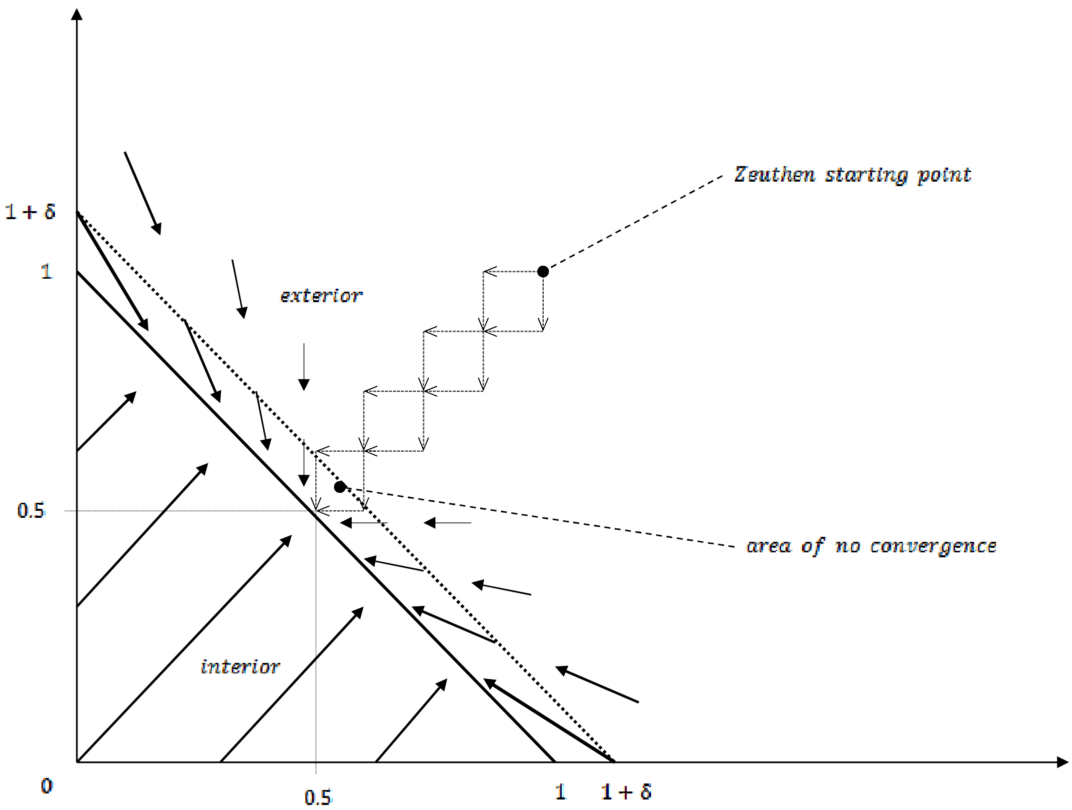

Figure 1: *The bargaining process takes place above zero. States below the Pareto frontier are transient, expected movement is outwards along 45-degree rays towards the Pareto frontier. In the external region, the process tends inwards and towards equal surplus splits. In the long run, the process moves between states with sums of demands equal to one (fat diagonal) and exterior states with sums of demands equal to* $1+\delta$ *(dashed diagonal). The zigzag in the exterior region highlights possible negotiation paths in Zeuthen's model. Long-run mass concentrates around equal surplus splits.*

the variance of demands if demands are feasible, and equal to zero when demands are infeasible.

Variance. Given any state $\mathbf{d} \in \mathbf{\Omega}$, the variance of demands is $Var(\mathbf{d}) = \frac{1}{n}\sum_{i\in N}(d_i - \mu)^2$ where $\mu = \frac{1}{n}\sum_{i\in N} d_i = \frac{1}{n}$ is the constant mean payoff in $\mathbf{\Omega}^e$. Write $\Delta(Var(\mathbf{d}^{\widehat{t+1}})) = Var(\mathbf{d}^{\widehat{t+1}}) - Var(\mathbf{d}^{\widehat{t}})$ for the change in variance between times $\widehat{t}$ and $\widehat{t+1}$ in $\mathbf{\Omega}^e$.

Variance drift.

Given any state $\mathbf{d}^{\widehat{\mathbf{t}}} = \mathbf{d}$ *at time* $\widehat{t}$ *such that* $\mathbf{d} \in \mathbf{\Omega}^e$*, we shall refer to* $\mathbb{E}[\Delta(Var(\mathbf{d}^{\widehat{t+1}}))|\mathbf{d}^{\widehat{t}} = \mathbf{d}]$ *as the variance drift.*

If $\mathbb{E}[\Delta(Var(\mathbf{d}^{\widehat{\mathbf{t+1}}}))|\mathbf{d}^{\widehat{t}} = \mathbf{d}] < 0$, there is an "equity-drift", that is, the variance of demands in $\mathbf{\Omega}^e$ diminishes in expectation.

Lemma 2. *Starting with* $\mathbf{d}^{\widehat{t}} = \mathbf{d} \in \mathbf{\Omega}^e$*, the variance drift is*

$$\mathbb{E}[\Delta(Var(\mathbf{d}^{\widehat{\mathbf{t+1}}}))|\mathbf{d}^{\widehat{t}} = \mathbf{d}] = 2ar\delta\left[\delta\frac{n-1}{n^2} - Var(\mathbf{d})\right]. \tag{4}$$

Proof of lemma 2. From any state $\mathbf{d} \in \mathbf{\Omega}^e$, we move to a given $\mathbf{d}_{ij} \neq \mathbf{d} \in \mathbf{\Omega}^e$ with probability $rf(d_j)\frac{1}{n^2}$ which is positive if $d_j > 0$. In the original chain, we leave $\mathbf{d}$ in one time step and come back to $\mathbf{d}$ in the next with probability $\sum_{i\in N}\sum_{j\neq i} rf(d_j)\frac{1}{n^2}$. Hence, with probability $1 - \sum_{i\in N}\sum_{j\neq i} rf(d_j)\frac{1}{n^2}$, we stay in $\mathbf{d}$ in $\mathbf{\Omega}^e$. The next expected sum of squares of demands in $\mathbf{\Omega}^e$ is therefore

$$\mathbb{E}[\textstyle\sum_{i\in N}(d_i^{\widehat{t+1}})^2|\mathbf{d}^{\widehat{t}} = \mathbf{d}] =$$
$$\textstyle\frac{r}{n^2}\sum_i\{\sum_{j\neq i} f(d_j)([d_i+\delta]^2 + [d_j-\delta]^2 + \sum_{k\neq i,j} d_k^2)\} + (1 - \sum_{i\in N}\sum_{j\neq i} rf(d_j)\frac{1}{n^2})\sum_i d_i^2.$$

Expanding the squares, this becomes

$$\textstyle\frac{r}{n^2}\sum_i\{\sum_{j\neq i} f(d_j)(\sum_k d_k^2 + 2\delta^2 + 2\delta[d_i - d_j])\} + (1 - \sum_{i\in N}\sum_{j\neq i} rf(d_j)\frac{1}{n^2})\sum_i d_i^2,$$

which is

$$\textstyle\sum_i d_i^2 + 2\delta r\frac{1}{n}\sum_i f(d_i)[\frac{n-1}{n}\delta - (\frac{\sum_i f(d_i)d_i}{\sum_i f(d_i)} - \frac{\sum_i d_i}{n})].$$

Substituting $f(d_i) = ad_i$ in the above equation, the drift in the sum of squares of demands is

$$\mathbb{E}[\Delta(\textstyle\sum_{i\in N}(d_i^{\widehat{t+1}})^2)|\mathbf{d}^{\widehat{t}} = \mathbf{d}] =$$
$$\textstyle 2\frac{ar\delta}{n}\sum_i d_i[\delta\frac{n-1}{n} - (\frac{\sum_i d_i^2}{\sum_i d_i} - \frac{\sum_i d_i}{n})] = 2ar\delta\sum_i d_i[\delta\frac{n-1}{n^2} - Var(\mathbf{d})],$$

which is also the drift in the variance as $\sum_{j\in N} d_j = 1$ for all $\mathbf{d} \in \mathbf{\Omega}^e$. $\square$

Note that the variance drift in $\mathbf{\Omega}^e$ is negative if, and only if,

$$Var(\mathbf{d}) > \delta\frac{n-1}{n^2}. \tag{5}$$

Furthermore, when $Var(\mathbf{d}) < \delta\frac{n-1}{n^2}$, any change in $\mathbf{\Omega}^e$ in a single time step is at most $\delta^2\frac{n-1}{n^2}$ (which occurs when $Var(\mathbf{d}) = 0$).[15]

3.5. Results

Theorem 3. *For any small $\beta > 0$ and for any large probability $1 - \gamma < 1$, there exists a step size $\delta \leq \frac{\beta\gamma}{3}$ and a time T_δ such that the variance of payoffs is less than β at least $1 - \gamma$ of the time after T_δ.*

Proof of theorem 3. First, we shall prove the following lemma:

Lemma 4. *From any state $\mathbf{d} \in \mathbf{\Omega}^e$, for any bargaining game with step size $\delta > 0$, there exists a time T_δ such that, for every $t > T_\delta$, relative inequity as measured by the variance of payoffs will in expectation be less than 2δ.*

Proof of lemma 4. It follows from proposition 1 that convergence of the process can be analyzed using the embedded chain $\mathbf{\Omega}^e$. Remember that, in $\mathbf{\Omega}^e$, payoffs and demands coincide, and recall that we move in $\mathbf{\Omega}^e$ in times $\widehat{t}$.

We prove this theorem in two steps. First, we prove that, from any state $\mathbf{d} \in \mathbf{\Omega}^e$ with $Var(\mathbf{d}) \leq 2\delta$, all expected future variances are less than 2δ, and that, from any state $\mathbf{d} \in \mathbf{\Omega}^e$ with $Var(\mathbf{d}) > 2\delta$, all expected future variances are less than $Var(\mathbf{d})$. Second, we prove that, for any initial state $\mathbf{d}^0 \in \mathbf{\Omega}^e$, it takes at most time $\widehat{T}$ for the expected variance to be less than 2δ. Jointly, these two facts imply that, starting anywhere in $\mathbf{\Omega}^e$, $\mathbb{E}[Var(\mathbf{d}^{\widehat{t}})|\mathbf{d}^0] \leq 2\delta$ after time $\widehat{T}$.

Step 1. Expression 4 is negative for all states $\mathbf{d}^{\widehat{T}} = \mathbf{d} \in \mathbf{\Omega}^e$ with $Var(\mathbf{d}) > \delta\frac{n-1}{n^2}$ (lemma 2). If $\mathbf{d}^{\widehat{T}} = \mathbf{d}$ is such that $Var(\mathbf{d}) \leq \delta\frac{n-1}{n^2} < \delta$, a maximum $\Delta(Var(\mathbf{d}^{\widehat{T+1}})) = \delta^2\frac{n-1}{n^2} < \delta$ may occur at the next time step and, thus, result in a $Var(\mathbf{d}^{\widehat{T+1}})$ no larger than 2δ.

Hence, for any state $\mathbf{d}$ with $Var(\mathbf{d}) > 2\delta$, it is true that, for all $\widehat{t'} > \widehat{t}$,

$$\mathbb{E}[Var(\mathbf{d}^{\widehat{t'}})|\mathbf{d}^{\widehat{t}} = \mathbf{d}] < Var(\mathbf{d}); \tag{6}$$

and, for any state $\mathbf{d}$ with $Var(\mathbf{d}) \leq 2\delta$,

$$\mathbb{E}[Var(\mathbf{d}^{\widehat{t'}})|\mathbf{d}^{\widehat{t}} = \mathbf{d}] < 2\delta. \tag{7}$$

Step 2. We now prove that there exists a time $\widehat{T} < \infty$ such that $\mathbb{E}[Var(\mathbf{d}^{\widehat{t}})|\mathbf{d}^0] \leq 2\delta$ indeed holds for all $\widehat{t} > \widehat{T}$ from any starting state $\mathbf{d}^0$ in $\mathbf{\Omega}^e$. Note that, for any time $\widehat{t}$

[15] Note that we may drop $2ar$ from this last expression because $2ar < 2r < 2\delta < 1$.

and for any state $\mathbf{d}^{\hat{t}} = \mathbf{d}$ with $Var(\mathbf{d}) = 2\delta$, we know that $\mathbb{E}[\Delta(Var(\mathbf{d}^{\widehat{t+1}}))|\mathbf{d}^{\hat{t}} = \mathbf{d}] < 0$. Hence, for any state $\mathbf{d}$ with $Var(\mathbf{d}) > 2\delta$, the drift can be bound by $\mathbb{E}[\Delta(Var(\mathbf{d}^{\widehat{t+1}}))|\mathbf{d}^{\hat{t}} = \mathbf{d}] = 2\frac{ar\delta}{n}[\delta\frac{n-1}{n} - Var(\mathbf{d})] < -2ar\delta^2\frac{n+1}{n^2}$. Writing $c \equiv 2ar\delta^2\frac{n+1}{n^2}$, we obtain the expression

$$\mathbb{E}[Var(\mathbf{d}^{\widehat{t+1}})|\mathbf{d}^{\hat{t}} = \mathbf{d}] \leq Var(\mathbf{d}) - c \tag{8}$$

for any $\mathbf{d} \in \mathbf{\Omega}^e$ with $Var(\mathbf{d}) > 2\delta$. Iteratively applying equation 8 as long as the variance exceeds 2δ yields, from any starting state $\mathbf{d}^0 \in \mathbf{\Omega}^e$,

$$\begin{aligned}\mathbb{E}[Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^0] &= \mathbb{E}[\mathbb{E}[Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^{\widehat{t-1}}]|\mathbf{d}_0] \\ &\leq \max\{\mathbb{E}[Var(\mathbf{d}^{\widehat{t-1}}) - c|\mathbf{d}^0], 2\delta\}.\end{aligned} \tag{9}$$

As long as $\mathbb{E}[Var(\mathbf{d}^{\widehat{t-1}}) - c|\mathbf{d}^0] > 2\delta$, we iterate expression 9 repeatedly forward to obtain

$$\mathbb{E}[Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^0] \leq \max\{Var(\mathbf{d}^0) - c\hat{t}; 2\delta\}, \tag{10}$$

which is less than or equal to 2δ for every $\hat{t} > \widehat{T}_\delta$ when $\widehat{T}_\delta \geq \frac{1}{c}(1-2\delta) \geq \frac{1}{c}(Var(\mathbf{d}^0) - 2\delta)$. □

Now we can prove theorem 3.

For any $\beta \in (0,1]$ and starting at any $\mathbf{d}^0 \in \mathbf{\Omega}^e$, lemma 4 implies that, for any $\hat{t} > \widehat{T}_\delta$,

$$\mathbb{P}([Var(\mathbf{d}^{\hat{t}}|\mathbf{d}^0] \geq \beta) \cdot \beta + \mathbb{P}([Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^0] < \beta) \cdot 0 \leq \mathbb{E}[Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^0].$$

Rearranged, for any $\mathbf{d}^0 \in \mathbf{\Omega}^e$, it holds for any $\beta > 0$ and $\gamma > 0$, that (yielding the Markov inequality)

$$\mathbb{P}([Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^0] \geq \beta) \leq \frac{\mathbb{E}[Var(\mathbf{d}^{\hat{t}})|\mathbf{d}^0]}{\beta} \leq \frac{2\delta}{\beta} \leq \gamma, \tag{11}$$

by appropriate choices of $\delta \leq \frac{\beta\gamma}{2}$ and this occurs after time

$$\hat{t} > \widehat{T}_\delta \geq \frac{1}{c}(1-2\delta). \tag{12}$$

From any state $\mathbf{d} \notin \mathbf{\Omega}^e$ with $\sum_{i\in N} d_i < 1$, $\mathbb{E}[\sum_{i\in N} d_i^t] \geq 1$ after

$$t > T'_\delta = \frac{1}{r\delta}. \tag{13}$$

For any state $\mathbf{d} \notin \mathbf{\Omega}^e$ with $\sum_{i\in N} d_i > 1$, $\mathbb{E}[\sum_{i\in N} d_i^t] \leq 1 + \delta$ after

$$t > T''_\delta = \frac{n^2}{a\delta} \tag{14}$$

Starting at any state $\mathbf{d}^0 \in \mathbf{\Omega}$, expression 11 therefore generalizes to

$$\mathbb{P}([Var(\mathbf{d}^t) + |\sum_{i \in N} d_i^t - 1| \; |\mathbf{d}^0] \geq \beta) \leq \frac{\mathbb{E}[Var(\mathbf{d}^t) + |\sum_{i \in N} d_i^t - 1| \; |\mathbf{d}^0]}{\beta} \leq \frac{3\delta}{\beta} \leq \gamma, \quad (15)$$

which holds for any $\beta > 0$ and $\gamma > 0$ by appropriate choices of $\delta \leq \frac{\beta\gamma}{3}$ and by adjustment for time; for all

$$t > T_\delta \geq \frac{n}{ar\delta}\widehat{T}_\delta + T'_\delta + T''_\delta. \quad (16)$$

□

Corollary 5. *The expected waiting time until theorem 3 holds, T_δ, is of order n^2.*

Proof of corollary 5. Expression 16 gives the expected waiting time for theorem 3:

$$T_{\beta,\gamma,\delta} = \frac{1}{r\delta} + \frac{n^2}{a\delta} + \frac{n^2}{2a^2r^2\delta^3}(1 - 2\delta). \quad (17)$$

The first term in 17, $\frac{1}{r\delta}$, follows from equation 13, which gives the maximal expected waiting time to reach a state in $\mathbf{\Omega}^e$ from any state in $\mathbf{\Omega}$ with $\sum_{i \in N} d_i < 1$. In particular, this is the expected waiting time to reach the Pareto frontier starting at $\mathbf{d}$: $d_i = 0$ for all i.

The second term, $\frac{n^2}{a\delta}$, follows from equation 14, which gives the maximal expected waiting time to reach a state in $\mathbf{\Omega}^e$ from any state in $\mathbf{\Omega}$ with $\sum_{i \in N} d_i > 1$. In particular, this is the expected waiting time to reach the Pareto frontier starting $\mathbf{d}$: $d_i = 1 + \delta$ for all i.

The third term follows from equation 12, which gives the maximal expected waiting time to reach a state in $\mathbf{\Omega}^e$ with $Var(\mathbf{d}) < 2\delta$ from any state in $\mathbf{\Omega}^e$, corrected by the maximal expected waiting time in between any two states in $\mathbf{\Omega}^e$. The correction includes one $1/r$ for the expected time spent on the Pareto frontier and another $n/a\delta$ for the maximal expected time spent one δ off the Pareto frontier until the next reduction occurs.

Jointly, this implies that $T_{\beta,\gamma,\delta} \in \mathcal{O}(n^2)$.

Note that the respective average times spent with feasible (infeasible) demands are $1/r$ (n/a). □

4. Conclusion

Zeuthen (1930) formulates a mechanistic bilateral negotiation protocol that mirrors behavioral elements of adjustments. In his model, adjustments were attributable to common knowledge about players' relative willingness to concede or to risk conflict. We propose a related dynamic based on aspiration adjustment theory and experimental evidence from directional learning for the case of a homogeneous bargaining population. Importantly, we assume that agents have information only about their own demands and payoffs but not about those of others. We have proposed a model that incorporates the underlying revision procedures in a fully dynamic n-player bargaining model. In Zeuthen, the key assumption regarding individual adjustments is that, starting from infeasible demands, the party which currently holds the higher demand incrementally reduces with probability one. This coincides with a deterministic description of our model. We assume that, during bargaining breakdown, players with higher utility loss reduce with larger probabilities than players with smaller utility loss. But, instead of consuming the pie only once, our bargaining game is infinitely repeated. Over time, the procedures implement equal splits of the surplus in a zonal rather than pinpoint way: using Brems's (1976; p. 404) famous words on Zeuthen's bargaining model to describe the final convergent area as an

- *"area around the middle in which no party is substantially more eager to secure an agreement than the other. Establishing the existence of such centripetal forces - powerful around the edges of the bargaining area but weaker towards the middle."*

Avenues for further research include multilateral bargaining experiments in the laboratory, building on the classic bargaining experiments by Tietz and Weber and on more recent non-bargaining experiments in low-information environments such as Bayer et al. (2013), Nax et al. (2013), Burton-Chellew et al. (2015). We are particularly interested in the speed with which convergence occurs, and the conditions under which such simple directional bargaining dynamics apply.

Acknowledgements

Much of this work is part of the author's D.Phil. (University of Oxford, 2011; supported by the Economic and Social Research Council [grant number PTA-031-2005-00118] and Postdoc under the Office of Naval Research Grant [grant number N00014-09-1-0751]. I also acknowledge current support by the European Commission through the ERC Advanced Investigator Grant 'Momentum' [grant number 324247]. I am particularly thankful to Peyton Young and Steffen Issleib, with whom I worked on this topic, and to Françoise Forges, Chris Wallace and Francis Dennig for comments on earlier versions of this model. The proof technique used is due to previous, joint unpublished work. I would also like to thank anonymous referees and a helpful editor for valuable comments on earlier versions of the paper.

References

Adams, J. S., 1965. Inequity in social exchange. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology 2, 267-299.

Alexander, J., Skyrms, B., 1999. Bargaining with Neighbors: Is Justice Contagious?, The Journal of Philosophy 96, 588-598.

Babichenko, Y., 2010. Uncoupled automata and pure Nash equilibria, International Journal of Game Theory 39, 483-502.

Babichenko, Y., 2012. Completely uncoupled dynamics and Nash equilibria, Games and Economic Behavior 76 (1), 1-14.

Bayer, R.-C., Renner, E., Sausgruber, R., 2013. Confusion and learning in the voluntary contributions game, Experimental Economics 16, 478-496.

Binmore, K.G., Piccione, M., Samuelson, L., 1998. Evolutionary stability in alternating-offers bargaining games, Journal of Economic Theory 80, 257-291.

Binmore, K.G., Samuelson, L., Young, H.P., 2003. Equilibrium selection in bargaining models, Games and Economic Behavior 45, 296-328.

Brems, H., 1976. From the Years of High Theory: Frederik Zeuthen (1888-1959), History of Political Economy 8, 400-411.

Burton-Chellew, M., Nax, H. H., West, S., 2015. Learning, not pro-sociality, explains the decline in cooperation in public goods games. Proceedings of the Royal Society B: Biological Sciences 282(1801), 20142678.

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

Cross, J. G., 1983. A Theory of Adaptive Economic Behavior, Cambridge University Press.

Crössmann, H., Tietz, R., 1983. Market behavior based on aspiration levels, in: Tietz (Ed.), R., Lecture Notes in Economics and Mathematical Systems 213, Berlin 1982, pp. 170-185.

Ding, J., Nicklisch, A., 2013. On the Impulse in Impulse Learning, MPI Collective Goods Preprint 13/02.

Ellingsen, T., 1997. The Evolution of Bargaining Behavior, Quarterly Journal of Economics 112, 581-602.

Erev, I., Roth, A. E., 1998. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria, American Economic Review 88, 848-881.

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

Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics, Theoretical Population Biology 38, 219-232.

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

Gale, J., Binmore, K., Samuelson, L., 1995. Learning to be Imperfect: The Ultimatum Game, Games and Economic Behavior 8, 56-90.

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

Grosskopf, B., 2003. Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition, Experimental Economics 6, 141-158.

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining, Journal of Economic Behavior and Organization 3 (4), 367-388.

Harley, C. B., 1981. Learning the evolutionarily stable strategy, Journal of Theoretical Biology 89, 611-633.

Harsanyi, J.C., 1956. Approaches to the Bargaining Problem Before and After the Theory of Games: A Critical Discussion of Zeuthen's, Hicks', and Nash's Theories, Econometrica 24, 144-157.

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

Hart S., Mas-Colell, A., 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.

John, R., Raith, M.G., 1999. Strategic Step-by-Step Negotiation, Journal of Economics 70, 127-154.

E. Kalai, 1977. Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons, Econometrica 45, 1623-1630.

Karandikar, R., Mookherjee, D., Ray, D., Vega-Redondo, F., 1998. Evolving aspirations and cooperation. J. Econ. Theory 80, 292-331.

Konrad, K. A., Morath, F., 2014. Bargaining with Incomplete Information: Evolutionary Stability in Finite Populations, Working Paper of the Max Planck Institute for Tax Law and Public Finance No. 2014-16.

Luce, R.D., Raiffa, H., 1957. Games and decisions: introduction and critical survey, Wiley, NY.

Marden, J.R., Young, H.P., Arslan, G., Shamma, J.S., 2009. Payoff-based dynamics for multiplayer weakly acyclic games. SIAM J. Control Optim. 48 (1), 373-396.

Marden, J.R., Young, H.P., Pao, L.Y., 2014. Achieving Pareto optimality through distributed learning. SIAM Journal on Control and Optimization 52 (5), 2753-2770.

Maynard Smith, J., 1974. The theory of games and the evolution of animal conflicts. Journal of Theoretical Biology 47 (1), 209-221.

Maynard Smith, J., Price, G. R., 1973. The logic of animal conflict, Nature 246 (5427), 15-18.

Nash, J., 1950. The Bargaining Problem, Econometrica 18, 155-162.

Nax, H.H., 2011. Evolutionary cooperative games, D.Phil. thesis, University of Oxford.

Nax, H. H., Pradelski, B. S. R., Young, H. P., 2013. Decentralized dynamics to optimal and stable states in the assignment game. IEEE Proceedings 52(CDC), 2398-2405.

Nax, H. H., Perc, M., 2015. Directional learning and the provisioning of public goods. Scientific Reports 5, 8010.

Nax, H. H., Pradelski, B. S. R., 2015. Evolutionary dynamics and equitable core selection in assignment games. International Journal of Game Theory, forthcoming.

Nowak, M. A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. Nature 428 (6983), 646-650.

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

Raiffa, H., 1953. Arbitration schemes for generalized two-person games, in: Kuhn, H., Tucker, A., Dresher, M. (Eds.), Contributions to the Theory of Games, Vol. 2, Princeton University Press, pp. 361-387.

Roth, A.E., 1995. Bargaining experiments, in: Kagel, J., Roth, A.E. (Eds.), Handbook of Experimental Economics, Princeton University Press, pp. 253-348.

Roth, A.E., Erev, I., 1995. Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term, Games and Economics Behavior 8, 164-212.

Saez-Marti, M., Weibull, J.W., 1999. Clever Agents in Young's Evolutionary Bargaining Model, Journal of Economic Theory 86, 268-279.

Sandholm, W., 2010. Population Games and Evolutionary Dynamics, MIT Press.

Sauermann, H., Selten, R., 1962. Anspruchsanpassungstheorie der Unternehmung, Zeitschrift für die Gesamte Staatswissenschaft 118, 577-597

Schaffer, M. E., 1988. Evolutionary stable strategies for a finite population and a variable contest size, Journal of Theoretical Biology 132(4), 469-478.

Schelling, T. C., 1956. An Essay on Bargaining. American Economic Review 46(3), 281-306

Selten, R., Buchta, J., 1998. Experimental Sealed Bid First Price Auction with Directly Observed Bid Functions, in: Budescu, D., Zwick, I.E. & R. (Eds.), Games and Human Behavior, Essays in Honor of Amnon Rapoport.

Selten, R., Stoecker, R. 1986. End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach, Journal of Economic Behavior and Organization 7, 47-70.

Suppes, P., Atkinson, A. R., 1959. Markov Learning Models for Multiperson Situations, Stanford University Press.

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

Tietz, R., 1975. An Experimental Analysis of Wage Bargaining Behavior, Zeitschrift für die gesamte Staatswissenschaft 131, 44-91.

Tietz, R., Bartos, O., 1983. Balancing of aspiration levels as fairness principle in negotiations, in: Tietz, R. (Ed.), Lecture Notes in Economics and Mathematical Systems 213, pp. 52-66.

Tietz, R., Weber, H., 1972. On the nature of the bargaining process in the Kresko-game, in: Sauermann, H. (Ed.), Contributions to experimental economics, Vol. 3, pp. 305-334.

Tietz, R., Weber, H., 1978. Decision behavior in multi-variable negotiations, in: Sauermann, H. (Ed.), Contributions to experimental economics, Vol. 7, pp. 60-87.

Tietz, R., Weber, H., Vidmajer, U., Wentzel, C., 1978. On Aspiration Forming Behavior in Repetitive Negotiations, in: Sauermann, H. (Ed.), Contributions to experimental economics, Vol. 7, pp. 88-102.

Weber, H., 1976. On the theory of Adaptation of Aspiration Levels in a Bilateral Decision Setting, Zeitschrift für die gesamte Staatswissenschaft 132, 582-591.

Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57-84.

Young, H.P., 2004. Strategic Learning and Its Limits, Oxford University Press.

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

Zeuthen, F., 1930. Problems of Monopoly and Economic Warfare, Routledge, London.