Perseverance, Information and Stochastically Stable Outcomes

Murali Agastya*
Center for Rationality and Interactive Decision Theory
Hebrew Univeristy, Givat Ram
Jerusalem 91904
ISRAEL
e-mail: murali@math.huji.ac.il

First Draft: May 1994
This Version: December 1994

*I thank Sergiu Hart, Larry Samuelson and Avner Shaked for helpful discussions.

Abstract

One bargainer from a finite population X, is matched at random with a bargainer from another finite population Y. They simultaneously precommit to "minimal" shares of a unit surplus. Populations differ in their degree of perseverance, parameterized by $\lambda \in (0,1)$. If the players precommit to x and y such that $x + y \leq 1$, then player i gets his demand x_i as well as a fraction λ_i of the unbargained surplus $(1 - x - y)$. If $x + y > 1$, they get nothing.

When players play adaptively and sometimes make errors as in Young (1993b), in the long run, a single division of surplus is observed most often. This is close to the asymmetric Nash bargaining solution with the weights $(1 - \lambda_x)$ and $(1 - \lambda_y)$. The surprise here is that the population that seemingly does well in the one shot encounters loses in the long run.

Keywords: Adaptive Play, Stochastic Stability, Information, Perseverance, Assymetric Nash solution.

J.E.L Classification Nos: C78, C79

1 Introduction

In a recent paper, Young (1993b) provides an evolutionary justification for the asymmetric Nash bargaining solution. He considers a model in which bargainers' attitudes are shaped by precedent. In the most basic of the several models considered, one representative from one of the two populations is matched at random with another from the second population. They then play the Nash demand commitment game. That is, each player simultaneously precommits to a share of a unit surplus. If their demands are compatible, they get their demands; otherwise they get nothing. Assuming that agents are myopic and rely on precedent to determine their optimal responses at any instant, potentially , any split of the pie may emerge as the norm in the long run. However, if one allows for the possibility that players err in their decisions, then it is possible that one norm is displaced and another eventually emerges. In the long run, the norm the is observed most often is the one that is in some sense the hardest to displace. This norm is referred to as the *Stochastically Stable Convention* (SSC) and corresponds closely to the asymmetric Nash bargaining solution. The asymmetry here is induced by the information gathering propensity of the members of each population. The greater the information, larger is the share of surplus.

This paper concentrates on a different kind of asymmetry. The results are somewhat surprising and at least initially, non-intuitive. As in the above model, players simultaneously precommit to their demands. It these demands are not compatible, they get nothing. If however, the demands are compatible, they get their demands *as well as a non-negative fraction of the unbargained surplus.* For example, if the delegates from populations X and Y precommit to demands x and y such that $x + y \leq 1$, then player X gets his demand x as well as a non-negative fraction λ_x of the unbargained surplus $(1 - x - y)$. By way of interpretation, one may think of each player's initial demand as his *minimal* demand. The variable λ_i determines how much of the remaining surplus they obtain over and beyond their minimal demands. It may be thought of as a measure of the player's *perseverance.* A higher λ means a relatively more perseverate player. Furthermore, it will not be assumed that all the surplus is appropriated, that is $\lambda_x + \lambda_y \leq 1$.

As in the Young (1993b), every split of the pie may become the norm in the long run. It is in the characterization of the SSCs that the model differs. When the players err slightly in choosing their optimal responses, the set one divisions that one expects to observe is considerably refined. The SSCs of the revised model also closely correspond to an asymmetric Nash bargaining solution. The asymmetry is induced through both their information gathering propensities as well as their degree of perseverance. Assuming for the moment that the populations are identical in terms of their information gathering propensities, it turns out that the SSCs are close to the *asymmetric Nash bargaining*[1] *solution with the weights for players X and Y being $(1 - \lambda_x)$ and $(1 - \lambda_y)$ respectively.*

The above result is at least initially surprising because a more perseverant player who gets a relatively larger share of the unbargained surplus in individual encounters when the players underbargain, ends up with a much smaller share in the long run. The intuition for the result is as follows. To fix ideas, consider the case when $\lambda_x = 1$ and $\lambda_x = 0$. That is if the players should underbargain, player X takes all the

[1]For example, if the players have linear utilities for the surplus, then player X obtains x^* where $x^* = \arg\max x^{(1-\lambda_x)}(1-x)^{(1-\lambda_y)}$.

remaining surplus. Note that any split of the pie $(x, 1-x)$ may be a long run equilibrium. That is, if this split is somehow arrived upon by chance and played for a while, and if the players memories are somehow bounded, then regardless of what portion of the recent history that player X samples, he observes that player Y demands $(1-x)$ and must therefore demand x as a best response. Similar arguments hold for player Y and the split $(x, 1-x)$ is hence self-perpetuating.

Now, when we allow for the possibility of players making mistakes, it is possible to move away from $(x, 1-x)$ split. Now suppose that there is a possibility that player Y commits one error in the conventional split is $(x, 1-x)$, and demands a little more, say $1-x+\delta$, where $\delta > 0$ and if player X were to sample this demand, it becomes a weakly dominant strategy to demand $x-\delta$. Indeed, there is a positive probability that he my do this as long as that one error is in his information set. If this were to indeed happen, Player Y on the other hand recognizes that player X is now demanding $x-\delta$, instead of x and hence demand $1-x+\delta$ and in due course the new norm would be $(x-\delta, 1-x+\delta)$. Moving to a norm such as $(x+\delta, 1-x-\delta)$ is much more unlikely as this would require a series of errors on the part of player X for player Y to start revising his demand from $1-x$ to $1-x-\delta$.

The rest of the paper is organized as follows. Section 2 describes the model and thereupon quickly embarks on the proving the results formally. Section 3 contains a short conclusion. Proofs of some ancillary results used in the proofs in Section 2 are collected in the Appendix.

2 The Model

Time is discrete and is denoted by $t = 1, 2, \ldots$. At each date one player from a large population X is randomly matched with another player from another large population Y. It the two strike a bargain, there is a unit of surplus available for sharing between the two. Before they are matched, both players simultaneously precommit to a certain share of the surplus. If X demands x and Y demands y so that these demands are incompatible, that is $x + y > 1$, then they both get zero. However, if they underbargain, that is $x + y \leq 1$, then player i gets his demand as well as a fraction $\lambda_i \geq 0$ of the unbargained surplus, *i.e.* $(1-x-y)$. It will be assumed that $\lambda_x + \lambda_y \leq 1$. This is clearly more general than assuming that $\lambda_x + \lambda_y = 1$. Let α_x and α_y be two rational numbers. A integer m is said to be admissible if $\alpha_i m$ is an integer for both $i = x, y$. The variable λ_i measures the degree of *perseverance* of the player and α_i is a measure of his *information.* It will be assumed that the players are homogeneous within a population in these two variables, *i.e.* $\alpha_x = \alpha_{x'}$ and $\lambda_x = \lambda_{x'}$ for all x, x' in population X. Apart from certain interpretational issues, one may then assume that it is the same player from population X that is being matched with a particular player from population Y.

Players are assumed to be myopic and resort to historical records to determine an optimal demand at each date. More specifically, at date t, player X picks a random sample, say η_x containing $\alpha_x m$ from the m most recent historical records $\mathbf{s}_t = ((x_{t-m+1}, y_{t-m+1}), (x_{t-m+1}, y_{t-m+1}), \ldots (x_{t-1}, y_{t-1}))$. We will refer to $\mathbf{s}_t$ as the state at time t. Let $f_x(y|\eta_x)$ denote the frequency with which the demand y of player Y is observed in the sample η. Then, the expected payoff of a demand x for player X is

$$U_x(x|\eta_x) = \int_0^{(1-x)} v_x((1-\lambda_x)x + \lambda_x(1-y)) f_x(y|\eta_x) dy, \tag{1}$$

where $v_x(x)$ is the VNM utility of consuming the x of the surplus and has been scaled so that $v_x(0) = 0$. The functions $v_i(.)$ are assumed to be concave and continuously differentiable.

Now, let $\delta = 10^{-p}$, where p is a finite integer. Let $\mathcal{D}$ denote that the set of all $p-$ decimal fractions that are positive and less than or equal to one. We will assume that players demands are restricted to be $\mathcal{D}$. Player X is assumed to be myopic and maximize his one period payoff given by Eq. 1. Hence, his demand x_t must satisfy

$$x_t = \arg\max_{x \in \mathcal{D}} U_x(x|\eta_x) \tag{2}$$

for some sample η_x of size $\alpha_x m$ from the state $\mathbf{s}_t$. Since $\mathcal{D}$ is finite, the solutions to Eq. 2 exist. If there are several solutions to Eq. 2 then each of them is assumed to be played with a strictly positive probability. A similar set of equations hold for player Y.

Hence, the solution to Eq. 2 is well defined. The best response rules of the players together with the above restriction on the strategies determine a finite state Markov process. The state space Ω consists of all finite sequences of demands from $\mathcal{D} \times \mathcal{D}$ of length m. Letting $p_i^*(x_i|\mathbf{s})$ denote the probability with which player i demands x_i in state $\mathbf{s}$. A state $\mathbf{s}'$ is said to be the successor of a state $\mathbf{s}$ if $\mathbf{s}'$ has been obtained from the state $\mathbf{s}$ be deleting the first entry and concatenating an additional record to the residual sequence of length $m-1$. The transition probability from a state $\mathbf{s}$ to another state $\tilde{\mathbf{s}}$, $P_{\mathbf{s}\tilde{\mathbf{s}}} = 0$ if $\tilde{\mathbf{s}}$ is not a successor of $\mathbf{s}$ and

$$P_{\mathbf{s}\tilde{\mathbf{s}}} = \begin{cases} p_x^*(x|\mathbf{s})p_y^*(y|\mathbf{s}) & \text{if } \tilde{\mathbf{s}} \text{ is a successor of } \mathbf{s} \text{ and} \\ & (x,y) \text{ are the rightmost demands in } \tilde{\mathbf{s}}. \\ 0 & \text{otherwise} \end{cases}$$

Let P denote the matrix of the above transition probabilities. The Markov Process with the state space Ω, the transition probability matrix P is said to be an *evolutionary bargaining process* (EBP) with memory m, precision δ, behavioral parameters $\{\alpha_i, \lambda_i\}$. As mentioned in the introduction, the only addition made to Young's model to this point is the assumption that when players underbargain, they get a fraction of the unbargained surplus together with their original. The above EBP with $\lambda_i = 0$, for both $i = x, y$ is the special case studied by Young (1993). The reader who finds the description of the model presented here somewhat terse is referred to Young (1993b).

2.1 Convergence of the EBP

Definition 1 *(Convention) A state $\mathbf{s}$ is said to be a convention if and only if it is an absorbing state of the EBP, i.e. $P_{\mathbf{ss}} = 1$.*

Note that from a given state there is a positive probability of reaching only an immediate successor. Hence, if $\mathbf{s}$ is a convention and is the state at time t, it is the state at time $t+1$ as well. For this to be true, it must be the case that $\mathbf{s}$ must be a sequence of m identical demands. Now suppose that these identical demands are (x^*, y^*). Now regardless of which sample η_y that player Y picks from this state $\mathbf{s}$,

$$f_y(x|\eta_x) = \begin{cases} 1 & \text{if } x = x^* \\ 0 & \text{otherwise.} \end{cases}.$$

Now, since $\mathbf{s}$ is an absorbing state, it follows from a simple observation of Eq.1 that $y^* = (1 - x^*)$. Hence if $\mathbf{s}$ is a convention, it must be a sequence of m repetitions of demands $(x, 1-x)$ where $x \in \mathcal{D}$. Furthermore it is easy to see that the converse is also true. We will denote by $\mathbf{x}$ a convention in which the demands $(x, 1-x)$ are being repeated.

Once the players reach a particular convention, it becomes self perpetuating. Moreover, in the light of the above observations, if the players should arrive at a state in which a certain division is self-perpetuating, then it must be one in which all the surplus is appropriated. Theorem A, the proof of which is almost identical to Theorem 1 in Young (1993), provides sufficient conditions under which a convention can be reached regardless of the starting point.

Theorem A: *S*uppose that $\alpha_i < \frac{1}{2}$ for both $i = x, y$. Then the EBP converges to a convention with probability one.

Proof. Almost identical to Theorem 1 in Young (1993b) and is hence omitted.

3 Stochastically Stable Conventions

It is in the characterization of the *Stochastically Stable Conventions* (SSC) that the model presented here differs from that of Young (1993b). For a motivation of the notion of the SSC, the reader is referred to Young (1993a) or Young (1993b) and Agastya (1994). For the most part, I will focus on the formal details.

Definition 2 *(Mistake) Suppose that the EBP is in state* $\boldsymbol{s} = ((x_1, y_1), \ldots$
$\ldots (x_2, y_2), \ldots$
$\ldots (x_m, y_m))$ *at time t and in state* $\boldsymbol{s}' = ((x_2, y_2), (x_3, y_3), \ldots (x_m, y_m), (x, y))$ *at date* $t+1$. *The transition from* $\boldsymbol{s}$ *to* $\boldsymbol{s}'$ *is said to involve a mistake on the part of player X if there is no sample of size* $\alpha_x m$ *in* $\boldsymbol{s}$ *such for which* x *is a best response, i.e.* $p_x^*(x|\boldsymbol{s}) = 0$. *A mistake on the part of player Y is similarly defined.*

Clearly, the number of mistakes involved in a transition from a state $\mathbf{s}$ to its successor $\mathbf{s}'$ can only be $0, 1$ or 2 depending on which of the two players have made a mistake. Suppose that the probability with which player i makes a mistake is given by $\epsilon\gamma_i > 0$. Conditional on the fact that player i has made a mistake, let $q_i(x_i|\mathbf{s})$ denote the probability with which he demands the wage x_i. Clearly, q_i will be different from p_i^*. The parameter ϵ is the absolute probability with which players commit errors and $\frac{\gamma_i}{\gamma_j}$ is the relative probability of each player making an mistake. The event that player X makes a mistake is assumed to be independent of the event that player Y makes a mistake.

Now allowing for the possibility of players making mistakes, we obtain a new Markov process with transition probabilities given by $P^{\epsilon}_{\mathbf{s}\tilde{\mathbf{s}}}$ where

$$\begin{aligned} P^{\epsilon}_{\mathbf{s}\tilde{\mathbf{s}}} &= (1-\epsilon\gamma_x)(1-\epsilon\gamma_y)p_x^*(x|\mathbf{s})p_y^*(y|\mathbf{s}) + \epsilon\gamma_x(1-\epsilon\gamma_y)q_x(x|\mathbf{s})p_y^*(x|\mathbf{s}) \\ &\quad + \epsilon\gamma_y(1-\epsilon\gamma_x)p_x^*(x|\mathbf{s})q_y(y|\mathbf{s}) + \epsilon^2\gamma_x\gamma_y q_x(x|\mathbf{s})q_y(y|\mathbf{s}). \end{aligned}$$

Now, if one assumes that when a player makes a mistake, he may make any demand from $\mathcal{D}$ with a strictly positive probability, that is $q_i(x_i|\mathbf{s}) > 0$ for all $x_i \in \mathcal{D}$, then the new transition new Markov process P^{ϵ} is irreducible. It is easy to see that it is also aperiodic. Therefore, it has a unique invariant distribution, say μ^{ϵ}. For a given state $\mathbf{s}$, $\mu_{\mathbf{s}}^{\epsilon}$ is the limit of the relative frequency with which this state is observed in the first t periods as $t \to \infty$. Since the invariant distributions of the perturbed process converge to an invariant distribution of the original process, when the likelihood of making mistakes is small, the conventions that are observed most often are the ones in the support of this limit distribution. This method allows for a refinement of the set of conventions and correspondingly the set of divisions that one would expect to observe.

Definition 3 *(Stochastically Stable Convention) A convention $\boldsymbol{x}$ is said to be stochastically stable if* $\lim_{\epsilon \to 0} \mu_{\mathbf{x}}^{\epsilon}$ *exists and is positive. A state is strongly stable if* $\lim_{\epsilon \to 0} \mu_{\mathbf{s}}^{\epsilon} = 1$.

Identification of the SSCs is considerably simplified by a certain equivalence theorem initially due to Friedlin and Wentzell (1984) and adapted for finite processes by Young (1993a). In order to introduce this, certain definitions are required.

Definition 4 ($\mathbf{x}$-tree) *Fix a convention $\boldsymbol{x}$. A $\boldsymbol{x}$-tree is a directed graph with the set of conventions as its vertices such that from each convention $\boldsymbol{x}' \neq \boldsymbol{x}$, there is a unique path directed to $\boldsymbol{x}$ and there are no cycles.*

Definition 5 (Resistance) *Let $\boldsymbol{s}'$ be a successor of $\boldsymbol{s}$. The resistance between these two states, denoted by $\boldsymbol{r(s,s')}$, is the minimum number of mistakes required in the one period transition $\boldsymbol{s} \longrightarrow \boldsymbol{s}'$. Similarly, for any two states $\boldsymbol{s}^1$ and $\boldsymbol{s}^2$, $\boldsymbol{r(s^1,s^2)}$ is the minimum number of mistakes required to reach $\boldsymbol{s}^2$ from $\boldsymbol{s}^1$ through a sequence of one period transitions.*

The resistance of a $\mathbf{x}$-tree $\mathbf{r}(\mathcal{G})$ is naturally defined as the total resistance of each of its edges. Let $\mathcal{T}_{\mathbf{x}}$ denote the set of all $\mathbf{x}$-trees.

Definition 6 (Stochastic Potential) *The stochastic potential of a convention $\boldsymbol{x}$ is the least resistance among all $\boldsymbol{x}$-trees:*

$$\gamma(\boldsymbol{x}) = \min_{T \in \mathcal{T}_w} \sum_{(\mathbf{w}^1,\mathbf{w}^2) \in T} \mathbf{r}(\mathbf{w}^1,\mathbf{w}^2).$$

Theorem B (Young (1993a)) *The sequence of stationary distributions μ^{ϵ} converge to a stationary distribution μ^0 of P^0 as $\epsilon \to 0$. Moreover, a state $\boldsymbol{s}$ is stochastically stable if and only if $\boldsymbol{s} = \boldsymbol{x}$ is a convention and has the minimum stochastic potential among all conventions.*

In order to allow for a parsimonious construction of the tree of minimum stochastic potential, I make the following assumption.

Assumption 1 *Players only make mistakes that are a distance δ away from a best response. That is, if $\mathbf{s}$ is the state at time t, then for every i, $q_i(x_{i,t+1} | \mathbf{s}) > 0$ iff there is a $\hat{x}$ such the*

- $p_i^*(\hat{x} | \mathbf{s}) > 0$
- $|\hat{x} - x_i| \leq \delta$.

Two remarks are in order. First, note that Assumption 1 is made on the players' strategies rather than the payoffs. However, if the EBP is in a convention $\mathbf{x}$, then player X is assumed to demand both $x+\delta$ and $x-\delta$ as well as x. (Similarly player Y is assumed to demand $(1-x-\delta)$ and $(1-x+\delta)$ as well as $(1-x)$. In terms of payoffs, $x+\delta$ is a large mistake as it will be rejected for sure while $x-\delta$ is a small mistake as it will be met for sure and will involve only a small change in his payoff. So, in a sense, Assumption 1 allows only for extremes in terms of payoffs. Of course, we do not obtain a similar interpretation when the EBP is not in a convention as the payoffs in this model are not continuous in strategies.

Second, recall that the definition of stochastic stability was based on the assumption that the perturbed transition probability matrix had a unique invariant distribution. With the above assumption, it is no longer clear that the P^ϵ is irreducible. Consequently, it is now not immediate that a unique invariant distribution exists and hence stochastic stability may be an ill-defined concept. Theorem 2 below, which uses a special bound on the extent of players' information establishes the validity of the solution concept.

Theorem 1 *Under the hypotheses of Theorem A, P^ϵ admits a unique invariant distribution for every $\epsilon > 0$ even if Assumption 1 is imposed. Moreover, the support of this invariant distribution contains the set of all conventions.*

Proof. See Appendix.

For each δ define the following real valued functions $L_i(.;\delta)$, $H_i(.;\delta)$ and $R(.;\delta)$ as follows:

$$L_i(x;\delta) = \frac{[v_i(x) - v_i(x-(1-\lambda_i)\delta)]}{[v_i(x) + v_i(x-\delta) - v_i(x-(1-\lambda_i)\delta)]}, \tag{3}$$

$$H_i(x;\delta) = \frac{v_i(x)}{v_i(x+\delta) + v_i(x) - v_i(x+\lambda_i\delta)}, \tag{4}$$

$$R(x;\delta) = \min\left\{\alpha_x L_x(x;\delta), \alpha_y L_y(1-x;\delta)\right\}. \tag{5}$$

Theorem 2 *There exists a level of precision δ^* such that for all $0 < \delta \leq \delta^*$ a convention $\boldsymbol{x}^*$ is stochastically stable for every admissible m if and only if x maximizes the function $R(.,\delta)$ over all $x \in \mathcal{D}$.*

The proof of Theorem 2 uses Lemma 2 stated below. Its proof appears in the Appendix.

Lemma 1 *Let $\delta^* = \frac{\min\{\alpha_x,\alpha_y\}}{4\max\{\alpha_x,\alpha_y\}}$. For all $0 < \delta \leq \delta^*$, the minimum number of mistakes for either X or Y to have a best response different from the one their conventional demand in a typical convention $\boldsymbol{x}$ is at least $[mR(x;\delta)]$.*

Proof of Theorem 2: Let $\mathbf{x}^*$ be a maximum of $R(.;\delta)$. Now suppose we can construct a $\mathbf{x}^* - tree$, say $\mathcal{G}$, such that the resistance of the edge directed away from a convention $\mathbf{x} \neq \mathbf{x}^*$ is exactly equal[2] to $[mR(x;\delta)]$, then the resistance of this graph is equal to

$$\mathbf{r}(\mathcal{G}) = \sum_{\substack{x\in\mathcal{D}\\ x\neq x^*}} [mR(x;\delta)] \tag{6}$$

[2]For a real number α, $[\alpha]$ denotes it integer part.

Now consider any other $\hat{\mathbf{x}}-tree$, say $\hat{\mathcal{G}}$. By Lemma 2, we know that for any edge in this tree, say $\mathbf{x}^1 \rightarrow \mathbf{x}^2$, the resistance $\mathbf{r}(\mathbf{x}^1,\mathbf{x}^2) \geq \lceil mR(x^1;\delta)\rceil$. Hence the resistance of the graph $\hat{\mathcal{G}}$ is

$$\begin{aligned}
\mathbf{r}(\hat{\mathcal{G}}) &= \sum_{(\mathbf{x}^1,\mathbf{x}^2)\in\mathcal{G}} \mathbf{r}(\mathbf{x}^1,\mathbf{x}^2) \\
&\geq \sum_{\substack{x\in\mathcal{D}\\ x\neq\hat{x}}} \lceil mR(x;\delta)\rceil \\
&= \mathbf{r}(\mathcal{G}) + \{\lceil mR(x^*;\delta) - \lceil mR(\hat{x};\delta)\rceil\} \\
&\geq \mathbf{r}(\mathcal{G}). \qquad (7)
\end{aligned}$$

Since $\hat{\mathcal{G}}$ is an arbitrary tree it follows that the tree $\mathcal{G}$ has the least stochastic potential and by Theorem B, $\mathbf{x}^*$ is a stochastically stable outcome. If there are several maxima of $R(.;\delta)$ then there may be several stochastically stable conventions. Since the inequality in Eq.7 is strict if $\hat{x}$ is not a maximum of $R(.;\delta)$, it follows that $\hat{\mathbf{x}}$ cannot be a stochastically stable, again by Theorem B. It remains to show that one can construct such a graph $\mathcal{G}$ with the above properties when x^* of $R(.;\delta)$.

First note that the function $L_i(.;\delta)$ is a decreasing function of the first argument given our assumption that $v_i(.)$ is a concave. It follows then that there is an $x^* \in \mathcal{D}$ such that

$$R(x;\delta) = \begin{cases} \alpha_x L_x(x;\delta) & \text{if } x \geq x^* \\ \alpha_y L_y(1-x;\delta) & \text{if } x < x^*. \end{cases}$$

Note that $R(.;\delta)$ attains a maximum at x^*. Depending on the discretization, it may be that x^* is the only maximum or both x^* and $x^*-\delta$ are the maxima[3]. Let us first consider the case where x^* is the unique maximum.

Let $\mathcal{G}$ denote the $\mathbf{x}^* - tree$ where for every convention $\mathbf{x}$ with $x < x^*$, there is a edge $\mathbf{x} \rightarrow \mathbf{x}+\delta$ and for every convention $\mathbf{x}$ with $x > x^*$, there is an edge $\mathbf{x} \rightarrow \mathbf{x}-\delta$. We will now prove that the edge leading out of the convention $\mathbf{x}$ has resistance $\lceil mR(x;\delta)\rceil$.

Let $\mathbf{x}$ be a convention and for concreteness assume that $x < x^*$. Suppose that the EBP is in the convention $\mathbf{x}$ at date t. Between dates t and $t+\hat{t}$, suppose that player X demands $x+\delta$, where $\hat{t} = \lceil m\alpha_y L_y(1-x;\delta)\rceil - 1$. Let $\mathbf{s}$ denote the state at date $T_0 = t + \hat{t} + 1$. By our choice of $\hat{t}$, it follows that $\mathbf{r}(\mathbf{x},\mathbf{s}) = \lceil m\alpha_y L_y(1-x;\delta)\rceil = \lceil mR(x;\delta)\rceil$. The proof is complete if we construct a path between $\mathbf{s}$ and $\mathbf{x}+\delta$ such that at every date, from date T onwards, involves a best response for both players .

Let η_y denote the sample drawn from $\mathbf{s}$ consisting of $\alpha_y m$ most recent demands. In this sample, $\lceil m\alpha_y L_y(1-x;\delta)\rceil$ demands of player X are $x+\delta$ while the remaining are x. From the arguments in the proof Lemma 1, it is easy to see that $(1-x-\delta)$ is a best response for player Y to the sample η_y. By our assumptions that $\max\{\alpha_x,\alpha_y\} < 1/2$ and that every sample of size $\alpha_y m$ from the m most recent demands is sampled with a strictly positive probability, it follows that the sample η_y is among the m most recent records for the next $\alpha_x m$ dates and will be sampled with a strictly positive probability. Hence, there is a positive probability that between dates T and $T+\alpha_x m - 1$, player Y will demand $(1-x-\delta)$. Suppose this event did occur and the resulting state at date $T_1 = T_0 + \alpha_x m$ is $\tilde{\mathbf{s}}$.

Let η_x denote the sample of size $\alpha_x m$ consisting of the most recent demands from $\tilde{\mathbf{s}}$. Note that every demand of player Y in the sample η_x is $(1-x-\delta)$. Hence, the

[3]The second case occurs when $R(x^*-\delta;\delta) = \alpha_y L_y(1-x^*+\delta;\delta) = \alpha_x L_x(x^*;\delta) = R(x^*;\delta)$.

unique best response of player X to this sample is $x+\delta$. Again, by our assumptions on the players' memory size, there is a positive probability that player X will sample η_x for the next $\max\{\alpha_x, \alpha_y\}m$ dates and hence $x_t = x+\delta$, for all $t = T_1, T_1 + 1, \ldots T_1 + \max\{\alpha_x, \alpha_y\}m - 1$. Furthermore, since the sample η_y is present in $\tilde{\mathbf{s}}$, there is a positive probability that player Y will demand $(1-x-\delta)$ at date T_1. So there is a positive probability of observing $(x+\delta, 1-x-\delta)$ at date T_1. At date T_1+1, however, depending on m, either the entire sample η_y is available or the earliest record from η_y has been eliminated from memory. In case of the former, there is a positive probability that η_y may yet be sampled by player Y and $y_{T_1+1} = (1-x-\delta)$. In case of the latter, there is a positive probability the player Y picks the sample consisting of the $\alpha_y m - 1$ remaining demands from η_y and $(x_{T_1}, y_{T_1}) = (x+\delta, 1-x-\delta)$. Call this sample η_y^1.

Since $(1-x-\delta)$ was a best response to the sample η_y, it follows that $(1-x-\delta)$ is a best response to η_y^1 as well. Hence, there is a positive probability that $(x_{T_1+1}, y_{T_1+1}) = (x+\delta, 1-x-\delta)$ as well. Arguing similarly, we find that there is a positive probability of observing $(x+\delta, 1-x-\delta)$ from dates T_1 through $T_1+\alpha_y m - 1$. Suppose that the above sequence of events were to occur, we are lead to a state $\bar{\mathbf{s}}$ at time $T_2 = T_1 + \alpha_y m$ where the most $\alpha_y m$ most recent demands of player X are $x+\delta$ while $\alpha_y m$ most recent demands of player Y are $(1-x-\delta)$. Hence, from T_2 onwards there is a positive probability that player i samples the most recent $\alpha_i m$ demands to duly establish the convention $\mathbf{x}+\delta$.

A symmetric argument shows that $\mathbf{r(x,x}-\delta\mathbf{)}= [\alpha_x m L_x(x;\delta)] = [mR(x;\delta)]$. The arguments for the case where $R(.;\delta)$ has a maximum at x^* and $x^*-\delta$ should be clear as the edge $\mathbf{x}^* \to \mathbf{x}^* - \delta$ and the edge $\mathbf{x}^* - \delta \to \mathbf{x}^*$ have the same resistance.

Theorem 3 *Let $\{\boldsymbol{x}_\delta\}$ be a sequence of stochastically stable conventions. Then, $\lim_{\delta\to 0} x_\delta = x^*$ where*

$$x^* = \arg\max v_x(x)^{\alpha_x(1-\lambda_x)} v_y(1-x)^{\alpha_y(1-\lambda_y)}.$$

Proof. Let $\{\mathbf{x}_\delta\}$ be a sequence of stochastically stable conventions. Then, for each $\delta > 0$, x_δ maximizes the function $R(.;\delta)$ over all demands in $\mathcal{D}$ and hence also maximizes the function $R(.;\delta)/\delta$ over the same domain. Let $\bar{x}_\delta$ be the argmax of $R(.;\delta)/\delta$ over all *real numbers*. Then for every $\delta > 0$,

$$\alpha_x L_x(\bar{x}_\delta;\delta) = \alpha_y L_y(\bar{x}_\delta;\delta). \tag{8}$$

Furthermore, along every convergent subsequence of $\{\bar{x}_\delta\}$, if x^* is the limit, then,

$$\begin{aligned}\lim_{\delta\to 0} \alpha_x L_x(\bar{x}_\delta;\delta) &= \lim_{\delta\to 0} \frac{\alpha_x(1-\lambda_x)}{v(\bar{x}_\delta) + v_x(\bar{x}_\delta) - v_x(\bar{x}_\delta - (1-\lambda_x)\delta)} \\ &\quad \times \lim_{\delta\to 0} \frac{v_x(\bar{x}_\delta) - v_x(\bar{x}_\delta - (1-\lambda_x)\delta)}{(1-\lambda_x)\delta} \\ &= \frac{\alpha_x(1-\lambda_x)v_x'(x^*)}{v_x(x^*)}\end{aligned} \tag{9}$$

and

$$\lim_{\delta\to 0} \alpha_y L_y(1-\bar{x}_\delta;\delta) = \lim_{\delta\to 0} \frac{\alpha_y(1-\lambda_y)}{v(1-\bar{x}_\delta) + v_y(1-\bar{x}_\delta) - v_y(1-\bar{x}_\delta - (1-\lambda_y)\delta)}$$

$$\times \lim_{\delta \to 0} \frac{v_y(1-\bar{x}_\delta) - v_y(1-\bar{x}_\delta - (1-\lambda_y)\delta)}{(1-\lambda_y)\delta}$$
$$= \frac{\alpha_y(1-\lambda_y)v_y(1-x^*)}{v_y(1-x^*)} \tag{10}$$

Given our assumptions that v_i is continuously differentiable concave function, it is now easy to conclude, by taking limits on both sides of Eq. 8 and using equations 9 and 10 that x^* is the unique solution to the problem

$$\arg\max v_x^{\alpha_x(1-\lambda_x)}(x) v_y^{\alpha_y(1-\lambda_y)}(1-x).$$

Since the above arguments are true for an arbitrary sub-sequence of $\{\bar{x}_\delta\}$, it follows that $\{\bar{x}_\delta\}$ is a convergent sub-sequence. We leave it to the reader to convince himself that

$$|x_\delta - \bar{x}_\delta| \leq 2\delta$$

and conclude that hence $\{x_\delta\}$ converges to x^* as well, thus completing the proof.

4 Conclusion

Unlike in Young (1993b), greater information does not suffice to secure a larger share of the pie. How perseverant a player is, also matters. The surprise is, as articulated in the introduction, that the long run stable outcome is a decreasing function of the λ's.

Young (1993b) also considers several extensions of the basic model to heterogeneous populations. Corresponding extensions of the basic model presented here yield corresponding modifications of Theorems 2 and 3. I will not pursue them here because of the limited novelty. It suffices to point out that the fifty-fifty split arrived at when there is some mixing between the populations remains intact.

5 Appendix

Proof of Theorem 1:

It is easy to see that the Markov Process associated with P^ϵ is aperiodic. Now, for an aperiodic Markov chain, the set of all states Ω can be uniquely decomposed into disjoint closed sets, $C_1, C_2, \ldots C_K$ such that from any given state in the set C_i, all other states in this set and only these states can be reached with a strictly positive probability. In particular, if we can show that the decomposition must involve a unique closed set, say Ω^*, it follows that all the states in $\Omega \backslash \Omega^*$ must be transient and the states corresponding to Ω^* form an irreducible sub-chain. This sub-chain admits a unique stationary distribution. Since a transient state cannot be in the support of a stationary distribution, it then follows that there is a unique invariant distribution to the original chain. This stationary distribution is the invariant distribution of the sub-chain corresponding to the states in Ω^* extended with zeros corresponding to the states in $\Omega \backslash \Omega^*$.

It remains to show that by restricting to the mistakes satisfying Assumption 1, we can construct such a set Ω^*. Indeed, let $C_1, C_2, \ldots C_K$ be a decomposition. Let $\mathbf{s}$ be a state in C_k. By Theorem A, there is a positive probability of reaching some convention, say $\mathbf{x}$. Hence, C_k must contain $\mathbf{x}$. Starting at $\mathbf{x}$, following arguments almost identical to those in the proof of Theorem 3 involved in constructing the graph $\mathcal{G}$, it can be seen that any other convention can be reached with a strictly positive probability. Hence, C_k must contain the set of conventions. Since both $\mathbf{s}$ and C_k were arbitrarily chosen, it follows that $C_k = C_{k'} = \Omega^*$.

Lemma 2 *The number of mistakes on the part of player Y that are required for player X to have a best response different from x in the convention $\boldsymbol{x}$ is at least*

$$\lceil m \min\{\alpha_s L_x(x;\delta), \alpha_x H_x(x;\delta)\}\rceil.$$

Similarly, the number of mistakes on the part of player X that are required for player Y to have a best response different from $(1-x)$ in the convention $\boldsymbol{x}$ is at least

$$\alpha_y \lceil m \min\{L_y(1-x;\delta), H_y(1-x;\delta)\}\rceil.$$

Proof. I shall only prove the first part of the lemma. The second follows from symmetry. Let $\mathbf{s}$ be the <u>first</u> state, starting from the convention $\mathbf{x}$, in which X is the <u>first</u> of the two players to have a best response $x^* \neq x$. Hence there is a subsample say η of $\mathbf{s}$ of size $\alpha_x m$ such that $U_x(x^*|\eta) \geq U_x(x|\eta)$.

By definition of the state $\mathbf{s}$, every demand of Y in η that differs from $(1-x)$ is a mistake. The only mistakes that are permitted are $(1-x+\delta)$ and $(1-x-\delta)$ by Assumption 1. Among the demands of Y in η, suppose that a fraction π_l are $(1-x-\delta)$, a fraction π_h are $(1-x+\delta)$ and the remaining $(1-\pi)$ are $(1-x)$ where $\pi = \pi_l + \pi_h$. The total number of mistakes in η is equal to $\pi\alpha_x m$.

It is clear that x^* can only be $x+\delta$ or $x-\delta$. If player X continues with the demand x on picking the sample η, his payoff is

$$U_x(x|\eta) = (1-\pi)v_x(x) + \pi_l v_x(x+\lambda_x\delta) \tag{11}$$

But if he demands $x-\delta$ instead, his expected payoff is

$$\begin{aligned} U_x(x-\delta|\eta) &= \pi_h v_x(x-\delta) + (1-\pi)v_x(x-(1-\lambda_x)\delta) + \\ &\quad \pi_l v_x(x+\lambda_x\delta-(1-\lambda_x)\delta) \\ &= \pi v_x(x-\delta) + (1-\pi)v_x(x-(1-\lambda_x)\delta) + \\ &\quad \pi_l[v_x(x+\lambda_x\delta-(1-\lambda_x)\delta) - v_x(x-\delta)]. \end{aligned} \tag{12}$$

Hence, $x^* = x-\delta$ if and only if $U_x(x-\delta|\eta) \geq U_x(x|\eta)$. On comparing Eq.11 and Eq.12 this holds if and only if

$$\begin{aligned} \pi v_x(x-\delta) + (1-\pi)v_x(x-(1-\lambda_x)\delta) &\geq (1-\pi)v_x(x) + \pi_l v_x(x-\delta) + \\ &\quad \pi_l[v_x(x+\lambda_x)\delta) - v_x(x+\lambda_x\delta-(1-\lambda_x)\delta)] \\ &\geq (1-\pi)v_x(x) \end{aligned} \tag{13}$$

Straightforward algebraic manipulation of Eq.13 yields

$$\begin{aligned} \pi &\geq \frac{[v_x(x) - v_x(x-(1-\lambda_x)\delta)]}{[v_x(x) + v_x(x-\delta) - v_x(x-(1-\lambda_x)\delta)]} \\ &= L_x(x;\delta) \end{aligned} \tag{14}$$

Similarly, the expected payoff of player X on demanding $x+\delta$ is $U_x(x+\delta|\eta) = \pi_l v_x(x+\delta)$. Hence, if $x^* = x+\delta$, then we must have $U_x(x+\delta|\eta) \geq U_x(x|\eta)$ or

$$\begin{aligned}
\pi_l v_x(x+\delta) &\geq (1-\pi)v_x(x) + \pi_l v_x(x+\lambda_x\delta) \\
\pi v_x(x+\delta) &\geq (1-\pi)v_x(x) + \pi v_x(x+\lambda_x\delta) \\
&\quad +\pi_h[v_x(x+\delta) - v_x(x+\lambda_x\delta)] \\
&\geq (1-\pi)v_x(x) + \pi v_x(x+\lambda_x\delta) \qquad (15)
\end{aligned}$$

Once again, carrying out the usual algebraic manipulations, we obtain

$$\begin{aligned}
\pi &\geq \frac{v_x(x)}{v_x(x+\delta) + v_x(x) - v_x(x+\lambda_x\delta)} \\
&= H_x(x;\delta) \qquad (16)
\end{aligned}$$

Now, using Eq. 14 and Eq. 16 we have that the number of mistakes in η, namely $\pi\alpha_x m$ is at least as large as $\min\{L_x(x;\delta), H_x(x;\delta)\}$.

Proof of Lemma 2. From Lemma 2, we know that the number of mistakes required before either X or Y changes his demand must be at least

$$[m\min\{\alpha_x L_x(x;\delta), \alpha_x H_x(x;\delta), \alpha_y L_y(1-x;\delta), \alpha_y H_y(1-x;\delta)\}]$$

It then suffices to show that for all δ satisfying the hypothesis of the Lemma, that

$$\min\{\alpha_x H_x(x;\delta), \alpha_y H_y(1-x;\delta)\} \geq \begin{cases} \alpha_x L_x(x;\delta) & \text{if } x \geq 1/2 \\ \alpha_y L_y(1-x;\delta) & \text{if } x \leq 1/2. \end{cases}$$

To see this, first note that

$$\begin{aligned}
H_i(x;\delta) &\geq \frac{v_i(x)}{v_i(x+\delta)} \\
&\geq \frac{x}{x+\delta} \qquad (17) \\
&\geq 1/2. \qquad (18)
\end{aligned}$$

Inequality 17 follows from concavity of v_i while inequality in 18 follows from the fact that $x \geq \delta$.

Simple algebraic manipulation also yields the following sequence of inequalities.

$$\begin{aligned}
L_i(x;\delta) &= \frac{[v_i(x) - v_i(x-\delta)] - [v_i(x-(1-\lambda_i)\delta) - v_i(x-\delta)]}{v_i(x) - [v_i(x-(1-\lambda_i)\delta) - v_i(x-\delta)]} \\
&\leq \frac{[v_i(x) - v_i(x-\delta)]}{v_i(x)} \\
&\leq \frac{\delta}{x} \qquad (19)
\end{aligned}$$

Equation 19 follows from concavity of v_i. Now, if $x \geq 1/2$, then from Eq. 19 it follows that $L_x(x;\delta) \leq 2\delta$. It is now immediate, on using equation 18, that

$$\alpha_x L_x(x;\delta) \leq \min\{\alpha_x H_x(x;\delta), \alpha_y H_y(1-x;\delta)\}$$

whenever $\delta \leq \delta^*$. If $x \leq 1/2$, then $(1-x) \geq 1/2$ and hence $L_y(1-x;\delta) \leq 2\delta$. Again, it is easy to check that

$$\alpha_y L_y(1-x;\delta) \leq \min\{\alpha_x H_x(x;\delta), \alpha_y H_y(1-x;\delta)\}$$

if $\delta \leq \delta^*$.

References

[1] AGASTYA, M (1994): "An Evolutionary Bargaining Model," Discussion Paper, Center for Rationality and Interactive Decision Theory, Hebrew University, JERUSALEM 91904, ISRAEL.

[2] FRIEDLIN, M. AND WENTZELL, A. *Random Perturbations of Dynamical Systems,* Springer Verlag, New York, 1984.

[3] YOUNG, H.P. (1993a): "The Evolution Of Conventions," *Econometrica,* 61, 57-84.

[4] YOUNG, H.P. (1993b): "An Evolutionary Model of Bargaining," *Journal of Economic Theory,* 59, 1, 145-168