Notes for

Implementing the Nash Program in Stochastic Games

Dilip Abreu (Princeton) and David Pearce (NYU)

April 2009.

Preliminary and incomplete. Not for circulation.

1 Introduction

Nash (1953) considers a scenario in which two players may choose their strategies independently, but in which contractual enforcement is available both for strategic agreements the two players may come to, and for threats each player makes about what she will do if agreement is not reached. Nash gives two analyses of this problem, and shows that the two solutions coincide. One builds upon Nash (1950) in giving an axiomatic treatment, while the other devises what is now called a "Nash demand game" whose payoffs are perturbed to yield a unique refined Nash equilibrium payoff pair. Carrying out this dual axiomatic/noncooperative approach to strategic problems with contracts is what has been dubbed "the Nash program".

This paper attempts to implement the Nash program in a broad class of two-player stochastic games. Leaving behind the static world of Nash (1953), it admits problems in which the state of the world (for example, firms' marginal costs, capital stocks, inventories and so on) may evolve over time, perhaps influenced by the players' actions. Like a game without state variables, a stochastic game with contracts is, in essence, a bargaining problem. One wants to know how players are likely to divide the surplus afforded by their stochastic environment.

Since the passage of time is crucial in a stochastic game, whereas it plays no role in Nash (1953), it is not immediately clear how to do an exercise in the spirit of Nash in these dynamic settings. For this reason, we begin in Section 2 by recasting the atemporal game of Nash as a strictly repeated discounted game. At the beginning of each period, players select actions for that period, and have an opportunity to bargain over how to split the surplus for the rest of the infinite-horizon game. If agreement is not reached in period 1, there is another opportunity to bargain in period 2, and so on. All stationary perfect equilibria of the intertemporal game approach (as slight stochastic perturbations as in Nash (1953) tend to zero) the same division of surplus as the static Nash bargaining with threats (NBWT) solution. The result is independent of the rate of interest.

After the stochastic game model is introduced in Section 3, Section 4 develops the proposed solution for a broad class of these games. At the heart of the analysis is a family of interlocking Nash bargaining problems. With each state ω is associated a bargaining set (the convex hull of the set of all pairs of expected present discounted values of strategy profiles for the game starting in ω) and a disagreement point. The disagreement point is determined partly by the "threat" actions played in ω, and partly by the solution values of possible successor states of ω. The solution value at ω is generated by the feasible set and disagreement point at ω by the maximization of the "Nash product" just as it is in Nash (1950, 1953). At least one solution (giving action pairs and value pairs in each state) exists, and we give sufficient conditions for all solutions to have the same value pair starting at state ω: call this value pair $v^*(\omega)$.

Consider perturbing the game G so that it is not perfectly predictable whether a given pair of demands is feasible at ω. Section 5 establishes that all Markov perfect equilibrium payoffs have the same limit as the perturbation approaches 0; for the game starting at ω, this limit equals $v^*(\omega)$, the solution value suggested by the family of NBWT problems from the preceding paragraph.

Thus, the solution $v^*(\omega)$ has been given a noncooperative interpretation. Section 6 demonstrates that, applying the axiomatic approach of Nash (1953) to the family of NBWT problems of Section 3, one gets unique predictions of how surplus will be divided starting in any state ω. Showing that this prediction coincides with $v^*(\omega)$ completes the Nash program for stochastic games.

Given the flexibility of the stochastic game model, applications of the solution are almost limitless. Section 7 sketches one example that illustrates the ability of a relatively weak competitor to extort surplus from a stronger party.

Section 8 concludes, and relates the results to ongoing work on reputationally perturbed stochastic games.

2 Strictly Repeated Games

This Section translates the noncooperative treatment Nash (1953) gives his bargaining problem, from his static setting to a stationary, infinite-horizon environment. Making assumptions analogous to those of Nash, we derive identical results regarding the proportions in which surplus is divided, and the actions that should be employed as threats.

Nash takes as exogenous a finite game $G = (S_1, S_2; U_1, U_2)$ in strategic form (with associated mixed strategy sets M_1 and M_2) and a bargaining set $B \subseteq \mathbb{R}^2$. The set of feasible payoffs of G, namely $\Pi = co\,\{u(s) : s \in S\}$ (where co denotes "convex hull of"), represents all the payoffs players can attain without cooperation (ignoring incentives). The set B includes all payoffs available to players through cooperation, that is, through enforceable contracts. Nash assumes that B is convex and compact, and that $\Pi \subseteq B$. The interpretation is that if players are willing to cooperate, they may be able to attain payoff combinations not possible from playing G. (For example, if a couple are willing to sign a marriage contract, they gain additional legal rights and perhaps receive a tax break.)

For any arbitrary nonempty, compact, convex bargaining set $X \subseteq \mathbb{R}^2$ and "threat point" or "disagreement point" $d \in X$, $N(d)$ denotes the associated Nash bargaining solution. The latter is the unique solution to $\max_{x \in B} (x_1 - d_1)(x_2 - d_2)$ if there exists $x \in B$ such that $x \gg d$ and otherwise uniquely satisfies $N(d) \in X$ and $N(d) \geq x$ all $x \in X$ such that $x \geq d$. Let the functions $V_i : M_1 \times M_2 \rightarrow \mathbb{R}$ be defined by $V_i(m) = N_i(U(m))$.

In the strategic setting described by (G, B) as in the preceding paragraph, there is a bargaining set, but no exogenous threat point. In constructing his proposed solution, Nash imagines that players choose respective threats $m_i \in M_i$, $i = 1, 2$, knowing that the Nash bargaining solution will result (relative to the threat point (m_1, m_2) and B). That is, he defines the game $\widehat{G} = (M_1, M_2; V_1, V_2)$. Nash shows that this game $\widehat{G}$ whose pure strategies are the mixed strategies of G, has equilibria that are interchangeable and equivalent. Their value, denoted v^*, is the Nash bargaining with threats

(NBWT) solution.

Notice that the game $\widehat{G}$ is just a construction in the formulation of the solution, NOT the noncooperative implementation of that solution. The construction mixes the idea of Nash equilibrium with the Nash product, which was justified axiomatically in Nash (1950).

To obtain an entirely strategic justification for his proposed solution, free of any axiomatic assumptions, Nash devised a two-stage game as follows. In the first stage, each player i simultaneously chooses $m_i \in M_i$. Thus, the pure actions of the first stage game are the mixed strategies of G. In the second stage, having observed the actions (m_1, m_2) from the first stage, each player i makes a utility demand u_i. If the pair (u_1, u_2) is feasible in B (more precisely, B^+ as defined below), then it is implemented. Otherwise, the utility pair received by the players is $U(m_1, m_2)$, the threat point determined by first period choices. Since the threat pair is typically NOT a Nash equilibrium of G, the players often have an interest in not carrying it out; external enforcement is needed to ensure that the threats are not abandoned ex post.

There is in general a great multiplicity of (subgame perfect) equilibria of the two-stage game, so Nash introduces random perturbations to the feasible set, making players slightly unsure about whether a given pair of demands would be feasible or not. This allows him (after taking limits of sequences of equilibria, as the perturbations become vanishingly small) to isolate a particular equilibrium, whose value pair coincides with the feasible pair that maximizes the Nash product.

We follow Nash in assuming free disposal: if $u \in B$ and $v \leq u$ then v is feasible. Let $B^+ = \{v \mid v \leq u \text{ for some } u \in B\}$. In the unperturbed problem, if players demand $v = (v_1, v_2)$, the probability it is feasible is 1 if $v \in B^+$ and 0 if $v \notin B^+$. In a perturbed game, a perturbation function h specifies the probability that v will be feasible.

We consider perturbation schemes as defined by probability functions of the following form:

A perturbation is a function $h\colon \mathbb{R}^2 \to [0,1]$ with

(i) $h(v) = 1$ if $v \in B^+$ and $h(v) < 1$ if $v \notin B^+$.

(ii) h is continuously differentiable. Furthermore, $h(v) \in (0,1) \Rightarrow h_i(v_1, v_2) < 0$.

We are interested in limits of SPEs of a sequence of perturbed games, where the perturbation functions approach the unperturbed game in a natural way.

Nash shows that there is only one equilibrium that survives all local perturbations. It is unfortunately still possible that for any particular perturbation, there may be many equilibria with dramatically different values. This cannot be the case for any regular perturbation, as defined below.

A sequence of perturbations $\{h^n\}_{n=1}^{\infty}$ is regular if:

(i) A compact and $A \cap B^+ = \varnothing \Rightarrow \exists$ integer $\overline{n}$ s.t. $v \in A \Rightarrow h^n(v) = 0$ $\forall\, n \geq \overline{n}$.

Let $O^n = \{v \mid h^n(v) \in (0,1)\}$. For $(v_1, v_2) \in O^n$,

$$\psi^n(v) \equiv -\frac{h_1^n(v)}{h_2^n(v)}$$

is the slope of the iso-probability line at v.

Let $\overline{s}(v)$ and $\underline{s}(v)$ be the supremum and infimum respectively of slopes of supporting hyperplanes of B at v.

Let $\overline{B}^+$ denote the boundary of B^+.

(ii) $\forall v \in \overline{B}^+$ & $\forall \varepsilon > 0$, $\exists \delta > 0$ & $\overline{n}$ s.t.

$$\begin{aligned} v' &\in C^n \;\&\; |v' - v| < \delta \Longrightarrow \\ \underline{s}(v) - \varepsilon &\leq \psi^n(v') \leq \overline{s}(v) + \varepsilon \end{aligned}$$

The first condition imposes a uniformity on the way in which points outside B are assigned certain infeasibility as n grows. The second requirement is that asymptotically, the iso$-$probability sets must respect (approximately, for points near the frontier of B^+) the trade-offs between players' demands that are expressed in the slope of the frontier of B^+.

Let $\underline{v}_i$ denote player i's minmax payoff in G. Let $\bar{b}_i$ be player i's highest payoff in B (or equivalently B^+).

To avoid some tedious qualifications in the proofs, we assume that:

ASSUMPTION: $\underline{v}_i < \bar{b}_i$, $i = 1, 2$ and $(\bar{b}_1, \bar{b}_2) \notin B$.

Recall that v^* denotes the equilibrium payoff profile and let m^* denote a profile of mixed strategy equilibrium threats of the standard NBWT game associated with (G, B).

Let $\underline{m}_i \in M_i$ denote a strategy of i which minmaxes $j \neq i$.

Lemma 1 *If* $v_j^* = \bar{b}_j$ *then* $\underline{m}_i$ *is an optimal strategy for* i *in the NBWT game* $\widehat{G}$.

Proof. Let m_i^* be an optimal strategy for i in the NBWT game $\widehat{G}$ and furthermore equal to $\underline{m}_i$ if $v_j^* = \bar{b}_j$. Since by definition $U_j(m_i^*, m_j) \leq v_j^*$, it follows that whether or not $v_j^* = \bar{b}_j$, $U_j(m_i^*, m_j) < \bar{b}_j$ for all $m_j \in M_j$. ■

Even in a perturbed demand game, there may be degenerate equilibria in which each player i demands so much that if $j \neq i$ demands at least as much as his value at the threat point, the probability of feasibility is zero.

All our results are for equilibria that are nondegenerate in this sense on all subgames.

Proposition 1 says that the values of nondegenerate SPE's converge, as you move along a regular sequence of perturbations, to the NBWT value v^*.

Proposition 1 *Let $\{h^n\}$ be a regular sequence of perturbations and $\{\sigma^n\}$ any sequence of nondegenerate SPEs of the respective perturbed games. Then*

$$\lim_{n\to\infty} U(\sigma^n) = v^* \qquad \text{(NBWT solution).}$$

Proof. If the conclusion is false there exists a subsequence (which we again denote by n) of nondegenerate subgame perfect equilibria σ^n with corresponding equilibrium threats and demands m^n, v^n and equilibrium payoffs w^n which satisfy

$$w^n = v^n h^n(v^n) + (1 - h^n(v^n))d^n$$

where $d^n = U(m^n)$ and such that m^n, v^n, d^n and w^n converge to corresponding limits m, v, d, w and $w \neq v^*$.

It must be the case that v lies on the boundary of B^+ (denoted $\overline{B}^+$). If $v \notin \overline{B}^+$ then for large n, $h^n(v^n) = 0$, contradicting the nondegeneracy assumption. If $v \in B^+$ and $v \notin \overline{B}^+$ then the optimality of players' choice of demands is contradicted for large n.

Suppose $v_1 < v_1^*$. We argue that for large enough n if Player 1 chooses m_1^*, then in the subgame defined by m_1^* and Player 2's equilibrium threat m_2^n, Player 1's payoff will strictly exceed v_1^n, a contradiction. Denote by $\widehat{v}_i^n$ and $\widehat{w}_i^n$ Player i's equilibrium demand and payoff respectively in the subgame indexed by (m_1^*, m_2^n). Let $\widehat{d}^n \equiv U(m_1^*, m_2^n)$.

If $U_1(m_1^*, m_2^n) = \bar{b}_1$, then (for large n) $\widehat{d}_1^n > w_1^n$. Since $\widehat{d}_1^n$ is a lower bound for Player 1's payoff in the subgame, this yields a contradiction. As argued earlier, $U_2(m_1^*, m_2^n) < \bar{b}_2$. If $U(m_1^*, m_2^n)$ is (strictly) Pareto efficient then $U(m_1^*, m_2^n) = N(U(m_1^*, m_2^n))$. By the definition of m_1^*, $N(U(m_1^*, m_2^n)) \geq v_1^*$. Hence $U(m_1^*, m_2^n) \geq v_1^*$. Again we conclude that $\widehat{d}_1^n > w_1^n$, and obtain a contradiction.

It remains to deal with the case $U(m_1^*, m_2^n) \ll b$ (and consequently $U(m_1^*, m_2^n) \ll \bar{b}$) for some $b \in B$. Consider a (sub)-subsequence (for simplicity denote this also by n) such that $\widehat{v}^n$and $\widehat{d}^n$ converge to some $\widehat{v}$ and $\widehat{d}$.

In the subgame $\widehat{v}_1^n$ solves

$$\max_{v_1^n} \left\{ \widehat{v}_1^n h^n(\widehat{v}_1^n, \widehat{v}_2^n) + (1 - h^n(\widehat{v}_1^n, \widehat{v}_2^n))\widehat{d}_1^n \right\}$$

The FONC are:

$$\widehat{v}_1^n h_1^n + h^n - h_1^n \widehat{d}_1^n = 0 \ \ or \ \ (\widehat{v}_1^n - \widehat{d}_1^n) h_1^n = -h^n$$

By the nondegeneracy assumption $h^n(\widehat{v}^n) > 0$. It follows that $(\widehat{v}_1^n - \widehat{d}_1^n) > 0$, and $h_1^n < 0$.

Since the corresponding conditions apply to Player 2,

$$\frac{\widehat{v}_2^n - \widehat{d}_2^n}{\widehat{v}_1^n - \widehat{d}_1^n} = \frac{h_1^n(\widehat{v}_1^n, \widehat{v}_2^n)}{h_2^n(\widehat{v}_1^n, \widehat{v}_2^n)}.$$

As noted earlier it must be the case that $\widehat{v}$ lies on the boundary of B^+ (denoted $\overline{B}^+$). Note that in this case $d \ll b$ for some $b \in B$.

It follows that for all $\varepsilon > 0$, there exists $\overline{n}$ such that for all $\psi^n(\widehat{v}^n) \equiv -\frac{h_1^n(\widehat{v}_1^n, \widehat{v}_2^n)}{h_2^n(\widehat{v}_1^n, \widehat{v}_2^n)}$ (the slope of the iso-probability line at $\widehat{v}^n$) satisfies $\underline{s}(\widehat{v}) - \varepsilon \leq \psi^n(\widehat{v}^n) \leq \overline{s}(\widehat{v}) + \varepsilon$.

It follows that $\frac{\widehat{v}_2 - \widehat{d}_2}{\widehat{v}_1 - \widehat{d}_1} = -s$ for some $s \in [\underline{s}(\widehat{v}), \overline{s}(\widehat{v})]$. By Nash (1950, 1953), if $\widehat{v}$ is on the boundary of B and $\widehat{d} \ll b$ for some $b \in B$, then the preceding condition is satisfied if and only if $\widehat{v} = N(\widehat{d})$. Futhermore $\widehat{v} \gg \widehat{d}$. We now argue that $\widehat{w} = \widehat{v}$. If $h^n(v^n) \to 1$ then clearly $\widehat{w} = \widehat{v}$. Now suppose $h^n(v^n) \nrightarrow 1$. By assumption, for all $b \in B^+$ either $b_1 < \bar{b}_1$ or $b_2 < \bar{b}_2$. Since $\widehat{v} = N(\widehat{d})$, $\widehat{v} \in B^+$. If $\widehat{v}_j < \bar{b}_j$ then for large n Player i can guarantee feasibility by reducing $\widehat{v}_i^n$ slightly, which will be a profitable deviation if

$\widehat{v}_i > \widehat{d}_i$ (which is the case) given that $h^n(v^n) \nrightarrow 1$ as we have assumed. Thus $h^n(v^n) \nrightarrow 1$ leads to a contradiction.

By the definition of m_1^* and Nash's geometric characterization of $N(.)$,

$$\frac{v_2^* - U_2\left(m_1^*, m_2^n\right)}{v_1^* - U_1\left(m_1^*, m_2^n\right)} \geq -s \text{ } for \text{ } all \text{ } s \in [\underline{s}(v^*), \overline{s}(v^*)]$$

Since w lies to the left of v^* on the frontier of B^+ it follows that

$$\frac{w_2 - U_2\left(m_1^*, m_2^n\right)}{w_1 - U_1\left(m_1^*, m_2^n\right)} > -s \text{ } for \text{ } all \text{ } s \in [\underline{s}(w), \overline{s}(w)]$$

Consequently

$$\frac{w_2 - \widehat{d}_2}{w_1 - \widehat{d}_1} > -s \text{ } for \text{ } all \text{ } s \in [\underline{s}(w), \overline{s}(w)]$$

It follows that $(\widehat{w}_1) = \widehat{v}_1 = N_1(\widehat{d}) > w_1$, as required. ■

Proposition 1 does something slightly different from what Nash (1953) shows, establishing convergence of ALL nondegenerate equilibria (along ANY regular sequence of perturbations). Nash instead argues that only one equilibrium survives ALL nearby perturbations, and he admits a broader class of perturbation functions than we do. A second distinction is that Nash does the limit analysis only in the second- stage game, using the limiting values to determine first-period behavior.

This completes our analysis of the static world of Nash (1953). We turn now to the description of an infinite horizon model whose SPE's yield the same (limiting) results. In each period (if agreement has not yet been reached), the two players play the perturbed two-stage game described earlier: each player i chooses a threat m_i from M_i, and having observed her opponent's threat, chooses a demand $v_i \in \mathbb{R}$. With probability $h(v)$, the demands are feasible, and the game is essentially over: each player i receives v_i in each subsequent period. With complementary probability, the demands are infeasible, and play proceeds to the next period. In every period before agreement is reached the same perturbation function h is used, but the draws are independent across time. Payoffs are discounted at the rate of interest $r > 0$.

Notice that the utility pair $U(m_1, m_2)$ serves as a temporary threat point: it will determine the period-t payoffs if the demand pair is infeasible. In contrast to Nash (1953), infeasibility causes a delay to cooperation rather than irreversible breakdown.

We are interested in the Markov perfect equilibria (MPE) of the repeated game. An MPE is a stationary subgame perfect equilibrium in which neither player's behavior in period t depends on the history of actions or demands in earlier periods.

The proposition below is the analog of the result Nash (1953) derives for his two-stage noncooperative game (in which a choice of threats is followed by a Nash demand game). It proves that along any sequence of perturbed games (and MPE's thereof) with the perturbations converging to 0, the demands made by the players converge to the NBWT solution (Nash (1953). Thus, the repeated game is an alternative to Nash's original two-stage game as a setting in which to give noncooperative expression to the NBWT solution.

We omit the proof. The repeated environment is a special case of the stochastic environment introduced in the next section and Proposition 2 is an implication of Theorem 1 of Section 6.

Proposition 2 *Let $\{h^n\}$ be a regular sequence of perturbations of the "repeated bargaining game" and $\{\sigma^n\}$ any sequence of corresponding nondegenerate Markov perfect equilibria of the respective perturbed games. Then*

$$\lim_{n\to\infty} U(\sigma^n) = v^*$$

An axiomatic foundation for the NBWT solution is easily given in the repeated game setting of this section, but it is covered in the more general treatment of Section 6.

3 The Stochastic Model

In the stationary model of Section 2, the noncooperative game G summarizes the payoff pairs that are feasible (ignoring incentives), and the bargaining set B specifies a weakly larger set of payoffs available to players if they sign binding contracts. This section specifies the game and the bargaining sets (one for each state) for the stochastic environment studied in Sections 4, 5 and 6.

The role of G will be played by $\mathcal{G} = (\Omega, S_i(\omega), U_i(.;\omega), \rho(.;\omega, s(\omega)), s(\omega) \in S(\omega), \omega \in \Omega, i = 1, 2, \omega_0, r)$, where Ω is the finite set of states, ω_0 is the initial state, $S_i(\omega)$ is the finite set of pure strategies available to player i in state ω, U_i specifies i's utility in any period as a function of the state ω prevailing in that period and the action pair $s \in S(\omega)$ played in that period, $\rho(\omega'.;\omega, s)$ is the probability that if state ω prevails in any period t, and s is the action pair in $S(\omega)$ played in t, state ω' will prevail in period $t + 1$. Let $M_i(\omega)$ be the mixed strategy set associated with $S_i(\omega)$. For any $m(\omega) \in M(\omega)$, $\rho(\omega';\omega, m(\omega)) = \sum_{s_1 \in S_1(\omega)} \sum_{s_2 \in S_2(\omega)} \rho(\omega';\omega, s)m_1(s_1;\omega)m_2(s_2;\omega)$. Finally r is the strictly positive rate of interest at which both players discount their infinite stream of payoffs.

The interpretation is that in period 1, each player i selects a strategy from $S_i(\omega_0)$ or from its associated mixed strategy set $M_i(\omega_0)$, and the strategy pair results in an immediate payoff and a probability of transiting to each respective state in period 2, and so on. Starting in any period t and state ω one can compute the feasible (average) payoffs from t onward; let this set be denoted $\Pi(\omega)$.

Let $B(\omega)$ denote the set of discounted average payoffs that the players could attain from period t onward starting in state ω, by signing contracts. Just as Nash assumed $\Pi \subseteq B$ (see Section 2), we assume for each ω that $\Pi(\omega) \subseteq B(\omega)$: contractual cooperation can achieve anything that independent action can achieve. Further, anything players can accomplish by acting independently today and then signing contracts tomorrow, they can achieve today by simply signing one contract today. Formally, we assume:

$co\{(1-\delta)u(m(\omega);\omega)+\delta\sum_{\omega'}\rho(\omega'\omega,m(\omega))v(\omega')$ s.t. $m(\omega)\in M(\omega), v(\omega')\in B(\omega')$ all $\omega'\} \subseteq B(\omega)$.

To establish uniqueness of a fixed point arising in the proposed solution in Section 4, either of the following conditions is sufficient.

Eventual Absorption (EA): The set of states can be partitioned into K classes Ω_k, $k = 1, ..., K$ such that Ω_K is an absorbing set of states and from any states in Ω_k, $k = 1, ..., K-1$, play can transit only to states in $\Omega_{k'}$ for $k' > k$.

Uniformly Transferable Utility (UTU): The efficiency frontiers of all $B(\omega)$, $\omega \in \Omega$ are linear and have the same slope.

Because of the availability of long-term contracts, it is not crucial to work with infinite-horizon stochastic games. Note that Eventual Absorption places no restrictions whatever on finite-horizon stochastic games. Transferable utility is most plausible when players are bargaining over something that is "small" relative to their overall wealth.

We will refer to the game $\mathcal{G}$ and the collection of bargaining sets $\mathcal{B}$, as a stochastic bargaining environment.

4 The Proposed Solution

Here we develop a solution for stochastic games with contracts, that will be given noncooperative and axiomatic justifications, respectively, in Sections 5 and 6. The goal is to formulate a theory that explains players' behavior in a state ω by analyzing the bargaining situation they find themselves in at ω.

What bargaining problem do players face at ω, if they have not yet signed a contract? The available strategies for player i are those in $M_i(\omega)$, and the bargaining set is $B(\omega)$. We want to follow Nash by maximizing the Nash product in $B(\omega)$ relative to the disagreement point. But if players choose the threat pair (m_1, m_2), the corresponding one-period payoff $U(m(\omega); \omega)$ is just the temporary disagreement point, familiar from Section 2. Taking a dynamic programming perspective, a player who observes that bargaining has failed today in state ω expects that after getting $U(m(\omega); \omega)$ today, she will get the value assigned by the solution to whatever state ω' arises tomorrow. Thus, the dynamic threat point $D(\omega; m)$ associated with threats m and proposed value function $v(.; m)$, is given by the formula:

$$\begin{aligned} D(\omega; m) &= (1-\delta)U(m(\omega); \omega) + \\ &\quad +\delta \sum_{\omega'} \rho\left(\omega' | \omega, m(\omega)\right) V\left(\omega'; m\right) \end{aligned}$$

which naturally depends on the rate of interest and on the endogenous transition probabilities.

Notice the simultaneous determination of the values $D(\omega; m)$ and $V(\omega'; m)$: we wish each $V(\omega; m)$ to maximize the Nash product relative to $D(\omega; m)$, but at the same time $D(\omega; m)$ is partly determined by the $V(\omega'; m)$. Thus, even holding fixed the threats $m(\omega)$, finding a solution involves a fixed point calculation. The uniqueness of the fixed point is guaranteed by either eventual absorption (EA) or by uniformly transferable utility (UTU) (see section 3).

Some useful definitions and notation follow.

Let b be a $|\Omega|$ $-$dimensional vector such that $b_\omega \in B(\omega)$. For given $m \in M$ define

$$\widetilde{D}(\omega; m(\omega), b) = (1-\delta)U(m(\omega); \omega) + \delta \sum_{\omega'} \rho(\omega'|\omega, m(\omega))\, b_{\omega'}.$$

Let $\overline{\overline{B}}(\omega)$ denote the efficient frontier of $B(\omega)$. By the consistency conditions relating $B(\omega)$ to the other $B(\omega')'s$ and $\mathcal{G}$, $\widetilde{D}(\omega; m(\omega), b) \in B(\omega)$. Let $\overline{\overline{B}} \equiv \prod_\omega \overline{\overline{B}}(\omega)$. Let the function $\xi_\omega(.; m(\omega)) : \overline{\overline{B}} \to \overline{\overline{B}}$ be defined by $\xi_\omega(b; m(\omega)) = N(\widetilde{D}(\omega; m(\omega), b); B(\omega))$. Define $\xi(.; m) : \overline{\overline{B}} \to \overline{\overline{B}}$ where $\xi(b; m) \equiv (\xi_\omega(b; m(\omega)))_\omega$.

Lemma 2 *Assume EA or UTU. Then for any $m \in M$, there exists a unique function $V(\cdot; m)$ defined on Ω, such that for all $\omega \in \Omega$, $V(\omega; m)$ is the Nash bargaining solution to the bargaining problem $(B(\omega), D(\omega; m))$, where*

$$D(\omega; m) = (1-\delta)U(m(\omega); \omega) + \delta \sum_{\omega'} \rho(\omega'|\omega, m(\omega))\, V(\omega'; m).$$

Proof. Fix $(m_1, m_2) \in M_1 \times M_2$ and first consider the case of EA. Suppose that the conclusion is true for $\omega \in \Omega^n$ for $n = k+1, k+2, ..., K$. We will argue that the conclusion is then true for $\omega \in \Omega^k$. By the EA assumption, if $\omega' \neq \omega$ and $\rho(\omega'|\omega, m(\omega)) > 0$ then $\omega' \in \Omega^n$ for some $n \in \{k+1, k+2, ..., K\}$. Consequently we may rewrite $D(\omega)$ as

$$D(\omega; m) = (1 - \delta P)A + \delta PV(\omega; m)$$

where $P = 1 - \sum_{\omega', \omega' \neq \omega} \rho(\omega'|\omega, m(\omega))$ and $(1 - \delta P)A = (1 - \delta)U(m; \omega) + \delta \sum_{\omega', \omega' \neq \omega} \rho(\omega'|\omega, m(\omega)) V(\omega'; m)$ where A is specified "exogenously" by the inductive hypothesis.

By the consistency conditions relating $B(\omega)$ to the other $B(\omega')'s$ and $\mathcal{G}$, $A \in B(\omega)$.

Since $A \in B(\omega)$, $N(A; B(\omega))$ is well defined. Since $V(\omega; m) - D(\omega; m) = (1 - \delta P)(V(\omega; m) - A)$ it follows that $V(\omega; m)$ is the Nash bargaining solution to the bargaining problem $(B(\omega), D(\omega; m))$ if and only if $V(\omega; m)$ is the Nash bargaining solution to the bargaining problem $(B(\omega), A)$.This establishes the induction. Finally note that the hypothesis is true for $\omega \in \Omega^K$: this corresponds to $P = 1, A = U(m; \omega)$.

Now suppose that UTU is satisfied. Recall the definitions preceding the statement of the lemma. Let s be the common slope of $\overline{\overline{B}}(\omega)'s$ and define $\varsigma = \begin{bmatrix} -s \\ 1 \end{bmatrix}$. Let $b_\omega \equiv (b_{1\omega}, b_{2\omega})$. Then for $b_\omega, b'_\omega \in \overline{\overline{B}}(\omega)$, $b'_\omega = b_\omega + (b'_{1\omega} -$

$b_{1\omega})\varsigma$. For $b, b' \in \prod_{\omega} \overline{\overline{B}}(\omega)$, let $\vartheta(b, b') = \max_{\omega} |b_1(\omega) - b_1'(\omega)|$ deine a metric on $\prod_{\omega} \overline{\overline{B}}(\omega)$. The mapping $\xi(.; m)$ is a contraction mapping with modulus δ. Clearly $(\prod_{\omega} \overline{\overline{B}}(\omega), \vartheta)$ is a complete metric space. By the contraction mapping theorem, $\xi(.; m)$ has a unique fixed point. Denote the latter b^*. Then setting $V(\omega; m) = b^*_{\omega}$ yields a unique solution to the collection of bargaining problems (associated with the given $m \in M$). ■

NOTE: Everything that follows depends only on the existence and uniqueness, for all $m \in M$, of the functions $V(\cdot; m)$ or equivalently that for all $m \in M$, the function $\xi(.; m)$ has a unique fixed point; the assumptions EA & UTU *per se* do not play any role in the argument below. Remember also for later use that if $b = \xi(b; m)$ then $V(\omega; m) = b_{\omega}$.

The above exercise was done for a fixed action pair. Now that value consequences for action pairs are established, we can ask for each state ω, what actions (threats, in Nash's interpretation, 1953) players would choose if they were in ω. In other words, we imagine players playing modified versions of $\mathcal{G}$, where for state ω, the payoffs will be given by $V(\omega, \cdot)$. This is called the threat game. It is indexed by the "initial" state ω and is denoted

$$\widehat{\mathcal{G}}(\omega) = (M_i, V_i(\omega, \cdot); i = 1, 2)$$

Again, we mimic Nash in thinking of players in ω choosing m_1 and m_2, to maximize $V_1(\omega; m)$ and $V_2(\omega; m)$ respectively. As in Nash(53), $\widehat{\mathcal{G}}(\omega)$ is a zero sum game: for all $m, m' \in M$, if $V_1(\omega, m) >$ (resp. $<$ and $=$)$V_1(\omega, m')$ if and only if $V_2(\omega, m)$ $<$(resp.$>$and$=$)$V_2(\omega, m')$. (Notice that we are not considering mixtures over the strategies in the M_i's and we look for 'pure' equilibria in the underlying strategy space M). This game's equilibria are interchangeable and equivalent, so it has a value $v^*(\omega)$.

Let b_i denote $(b_{i\omega})_\omega$ and recall the definitions preceding the previous lemma.

We have:

Lemma 3 *For any $m \in M$, $b, b' \in B$ and $i \in \{1, 2\}$, if $b_i' \geq b_i$, then $\xi_i(b'; m) \geq \xi_i(b; m)$.*

Proof. If $\widetilde{D}_{i\omega}(b') \geq D_{i\omega}(b)$ and $\widetilde{D}_{j\omega}(b') \leq \widetilde{D}_{j\omega}(b)$ for all $\omega \in \Omega$, then clearly $N_{i\omega}(\widetilde{D}(m; b'); B(\omega)) \geq N_{i\omega}(\widetilde{D}(m; b); B(\omega))$ for all $\omega \in \Omega$. ■

For $n = 2, 3, ...$, let $\xi^n(b; m) = \xi(\xi^{n-1}(b; m); m)$.

Lemma 4 *For $i = 1$ or 2 and $b \in \overline{\overline{B}}$, if $\xi_i(b; m) \geq b_i$ then there exists $b^* \in \overline{\overline{B}}$ such that $b^* \geq b$ and $b^* = \xi(b^*; m) = (V(\omega; m))_\omega$. Moreover for $n = 2, 3, ...$, $\xi_i^n(b; m) \geq \xi_i^{n-1}(b; m)$ and $b^* = \lim_{n\to\infty} \xi^n(b; m)$.*

Proof. Let $b^n \equiv \xi^n(b; m)$. By the preceding lemma,

$b_i^{n+1} \geq \xi_i(b^n; m) \geq b_i^{n-1}$

Clearly $\lim b^n$ exists. Since $\xi_i(.; m)$ is continuous, $\lim \xi_i(b^n; m) = \xi_i(b^*; m)$.

Hence $b_i^* \geq \xi_i(b^*; m) \geq b_i^*$, and $b^* = \xi(b^*; m)$. Of course, $b^* \geq b$. ■

Lemma 5 *Equilibria of $\widehat{\mathcal{G}}(\omega)$ are equivalent and interchangeable.*

Proof. This follows directly from the fact that $\widehat{\mathcal{G}}(\omega)$ is a "zero sum" game as explained above. ■

Let $b_\omega(m) \equiv V(\omega; m)$ and $b(m) = (b_\omega(m))_\omega$

Definition 1 *The strategy profile $m \in M$ is locally optimal if for all $m_i'(\omega) \in M_i(\omega), \omega \in \Omega, i = 1, 2$,*

$$\xi_{i\omega}(b(m); (m_i'(\omega), m_j(\omega))) \leq \xi_{i\omega}(b(m); (m_i(\omega), m_j(\omega))) = b_{i\omega}(m).$$

Lemma 6 *The strategy profile $m \in M$ is an equilibrium of $\widehat{\mathcal{G}}(\omega)$ for all $\omega \in \Omega$ if and only if m is locally optimal.*

Proof. Suppose m is locally optimal. Then for all $m_i' \in M_i, \omega \in \Omega, i = 1, 2$ $\xi_{i\omega}(b(m); (m_i'(\omega), m_j(\omega))) \leq \xi_{i\omega}(b(m); (m_i(\omega), m_j(\omega))) = b_{i\omega}(m)$. Hence $\xi_i(b(m); (m_i', m_j)) \leq b_i(m)$. By Lemma 4 it follows that $V(\omega; (m_i', m_j)) \leq b_{i\omega}(m) = V(\omega; m)$. It follows that $m \in M$ is an equilibrium of $\widehat{\mathcal{G}}(\omega)$ for all $\omega \in \Omega$. Conversely suppose there exists $\omega' \in \Omega$, $m_i'(\omega') \in M_i(\omega')$ such that $\xi_{i\omega}(b(m);(m_i'(\omega'),m_j(\omega'))) > \xi_{i\omega}(b(m);(m_i(\omega'),m_j(\omega')))$. Consider the strategy m_i' such that $m_i'(\omega') = m_i(\omega')$ and $m_i'(\omega) = m_i(\omega)$ for all $\omega \neq \omega'$. Again by Lemma 4 it follows that m_i' is a profitable deviation for Player i against m_j in $\widehat{\mathcal{G}}(\omega')$. ■

Lemma 7 *(Existence) There exists a strategy profile $m^* \in M$ such that m^* is an equilibrium of $\widehat{\mathcal{G}}(\omega)$ for all $\omega \in \Omega$.*

Proof. Say that $m_i'(\omega) \in M_i(\omega)$ is a "local best response to $m^0 \in M$" if

$$\xi_{i\omega}(b(m^0);(m_i'(\omega),m_j^0(\omega))) \geq \xi_{i\omega}(b(m);(m_i(\omega),m_j^0(\omega))).$$

for all $m_i(\omega) \in M_i(\omega)$. Consider the mapping $\eta : M \to M$ where

$$\eta_i\left(m_i^0,m_j^0\right) = \left\{m_i' \mid \text{for all } \omega,\ m_i'(\omega) \text{ is a "local best response to" } m^0\right\}.$$

By the definition of η a fixed point of η must be locally optimal. The result then follows from the preceding lemma. That η is non-empty valued and upper hemicontinuous follows from the continuity of the underlying functions and in particular the continuity of the NBWT solution in the disagreement payoff. We now argue that η is indeed convex valued. Suppose that $m_i', m_i'' \in \eta_i\left(m_i^0,m_j^0\right)$. For any $\alpha \in (0,1)$ we show that $\alpha m_i' + (1-\alpha) m_i'' \in \eta_i\left(m_i^0,m_j^0\right)$.

Recall the definitions preceding Lemma 2. Let $m' = (m'_i, m^0_j)$, $m'' = (m''_i, m^0_j)$ and $\overline{m} = \alpha m' + (1-\alpha) m''$. Then

$$\widetilde{D}(\omega; \overline{m}(\omega), b(m^0)) = \alpha \widetilde{D}(\omega; m'(\omega), b(m^0)) + (1-\alpha) \widetilde{D}(\omega; m''(\omega), b(m^0)).$$

Then $\xi_\omega(b(m^0); \overline{m}) \equiv N(\widetilde{D}(\omega; \overline{m}(\omega), b(m^0)); B(\omega)) = \xi_\omega(b(m^0); m'(\omega)) = \xi_\omega(b(m^0); m''(\omega))$. Hence $\overline{m}_i \in \eta_i\left(m^0_i, m^0_j\right)$, so that η is indeed convex valued. By Kakutani's fixed point theorem η has a fixed point m^* and we are done. ■

Notice that in addition to existence, the lemma asserts a nice time consistency property.

Let the function $v^*\colon \Omega \to \mathbb{R}^2$ be defined by $v^*(\omega) = V(\omega; m^*)$. This is the proposed solution.

In the framework of Nash (1953), the pair $(m^*_1, m^*_2) = m^*$ is the (state-contingent) pair of threats associated with the stochastic game with initial state ω, and $V(\omega; m^*_1, m^*_2)$ is the associated equilibrium value pair. These may be viewed as generalizations of the NBWT solution to stochastic environments.

5 Noncooperative Treatment

Section 4 developed a proposed solution for any stochastic game that satisfies "eventual absorption" or that has transferable utility. Here we provide support for the proposed solution by doing a noncooperative analysis of the stochastic game in the spirit of Nash (1953). As in Section 2, we perturb the demand game (in any state) and study the equilibria as the perturbations become vanishingly small. All Markovian equilibria have values in any state ω converging to $v*(\omega)$, the demand pair recommended by the proposed solution. Similarly, the limit points of any sequence of Markovian equilibrium action pairs at ω (as perturbations vanish) are in the interchangeable and equivalent set of temporary threat pairs at ω specified by the proposed solution. In other words, a noncooperative perspective points to the same state-contingent values and threat actions as the proposed solution.

We begin by describing the (unperturbed) noncooperative game to be analyzed. Based on the stochastic bargaining environment of Section 3, it involves the bargainers playing a threat game, followed by a demand game, in any period if no contract has yet been agreed upon. In period 1, the state is ω_0, so each player i chooses a threat $x \in M_i(\omega_0)$. Having observed the threats, players make demands (v_1, v_2). If $(v_1, v_2) \in B(\omega_0)$, the rewards are enforced contractually and the game is essentially over. Otherwise, the threat payoff is realized in period 1, and the state transits to ω' with probability $\rho(\omega'|\omega, x)$. In period 2, threats are again chosen (from sets that depend on the prevailing state), and so on.

This unperturbed game, denoted $\mathcal{G}$, naturally has many perfect Bayesian equilibria, so one looks at a sequence of perturbed games approaching $\mathcal{G}$. The nth element of the sequence is a stochastic game in which feasibility of a demand pair $(v_1, v_2) \in B(\omega)$ is given by $h^{nw}(v_1, v_2)$, where the outcomes are independent across periods. For any ω, the perturbation function h^{nw} satisfies the same conditions as in Section 3, and regularity of the sequence (with index n) is defined as before.

Before stating the convergence result precisely we provide some rough intuition for the case of "eventual absorption" (with K classes of states). In any absorbing state ω, players are in the situation covered by Section 2, where the "Nash bargaining with threats" convergence results were established. If instead ω is in class $K-1$, incentives are different, both because the game in the current period differs from the game to be played from tomorrow onward, and because threats today affect the state transition matrix. But the dynamic threat point defined in the construction of the proposed solution in Section 4 mimics these phenomena exactly, so convergence to the generalized NBWT threats and demands (the proposed solution) also occurs in these states. The same argument applies by induction to all states.

ASSUMPTION: There exists $\underline{m}_1(\omega) \in M_1(\omega)$ such that $\bar{b}_2(\omega) < (1-\delta)U((\underline{m}_1(\omega), m_2(\omega)); \omega) + \delta \sum_{\omega'} \rho(\omega' \mid \omega, (\underline{m}_1(\omega), m_2(\omega)))\bar{b}_2(\omega')$ for all $m_2(\omega) \in M_2(\omega)$.

Theorem 1 *Let* $\{h^{n\omega}\}_{n,\omega}$ *be a regular sequence of perturbations of the stochastic bargaining game and* $\{\sigma^n\}$ *any sequence of corresponding nondegenerate Markov Perfect equilibria of the respective perturbed games. Then*

$$\lim_{n\to\infty} U(\sigma^n(\omega)) = v^*(\omega)$$

Proof. If the conclusion is false there exists a subsequence (which we again denote by n) of nondegenerate Markov Perfect equilibria σ^n with corresponding equilibrium threats and demands m^n, v^n and equilibrium payoffs w^n which satisfy

$$D^n(\omega) = (1-\delta)U(m^n(\omega); \omega) + \delta \sum_{\omega'} \rho(\omega' \mid \omega, m(\omega)) w^n(\omega')$$

$$w^n(\omega) = v^n(\omega) h^{n\omega}(v^n(\omega)) + (1 - h^{n\omega}(v^n(\omega)))D^n(\omega)$$

such that $w^n \to w \neq v^*$. We may w.l.o.g. assume that the sequences v^n and D^n converge also. Let v, D and w denote the corresponding limits.

Step 1: $v(\omega)$ and $w(\omega)$ lie on the boundary of $B^{+}(\omega)$.

It must be the case that $v(\omega)$ lies on the boundary of $B^{+}(\omega)$ (denoted $\overline{B}^{+}(\omega)$). If $v(\omega) \notin \overline{B}^{+}(\omega)$ then for large n, $h^{n\omega}(v^{n}(\omega)) = 0$, contradicting the nondegeneracy assumption. If $v(\omega) \in B^{+}(\omega)$ and $v \notin \overline{B}^{+}(\omega)$ then the optimality of players' choice of demands is contradicted for large n. We now argue that $w(\omega) \in \overline{B}^{+}(\omega)$. If $h^{n\omega}(v^{n}(\omega)) \to 1$ then $w(\omega) = v(\omega)$ and consequently $w(\omega) \in \overline{B}^{+}(\omega)$. Now suppose $h^{n\omega}(v^{n}(\omega)) \nrightarrow 1$. By assumption if $v(\omega) \in \overline{B}^{+}(\omega)$ (which we have established above) either $v_1(\omega) < \bar{b}_1(\omega)$ or $v_2(\omega) < \bar{b}_2(\omega)$. If $v_j(\omega) < \bar{b}_j(\omega)$ then for large n Player i can guarantee feasibility by reducing $v_i^{n}(\omega)$ slightly, which will be a profitable deviation if $v_i(\omega) > D_i(\omega)$ given that $h^{n\omega}(v^{n}(\omega)) \nrightarrow 1$ as we have assumed. Suppose $v_1(\omega) < \bar{b}_1(\omega)$ and $v_2(\omega) < \bar{b}_2(\omega)$. Since by assumption $h^{n}(v^{n}) \nrightarrow 1$, it follows by the preceding argument that $v_i = D_i$, $i = 1, 2.$ and $w = v = D$. Consequently $w \in \overline{B}^{+}(\omega)$. Finally suppose that $v_j(\omega) < \bar{b}_j(\omega)$ and $v_i(\omega) = \bar{b}_i(\omega)$. Then $v_i(\omega) = D_i(\omega)$ as argued earlier. But in this case also $w(\omega) \in \overline{B}^{+}(\omega)$ (although we may have $w_j(\omega) < v_j(\omega)$). This completes Step 1.

We show below that $w(\omega) = V(\omega; m)$.

Suppose w.l.o.g. that $w_1(\omega) < V_1(\omega; m)$. We argue that for large enough

n Player 1's equilibrium payoff must strictly exceed $w_1^n(\omega)$, a contradiction. The argument proceeds by considering various cases starting with the central one below.

Step 2: If $D(\omega) \ll b$ for some $b \in B$ then $w(\omega) = V(\omega; m)$.

In the subgame $v_1^n(\omega)$ solves

$$\max_{v_1'^n(\omega)} \{v_1'^n(\omega)h^{n\omega}(v_1'^n(\omega), v_2^n(\omega)) + (1 - h^{n\omega}(v_1'^n(\omega), v_2^n(\omega)))D_1^n(\omega)\}$$

where $D^n(\omega) = (1-\delta)U\left(m_1^n(\omega), m_2^n(\omega); \omega\right) + \delta w^n(\omega)$.

The FONC are:

$$v_2^n(\omega)h_1^{n\omega}(v^n(\omega)) + h^{n\omega}(v^n(\omega)) - h_1^n D_1^n(\omega) = 0$$

$$\text{or} \quad (v_2^n(\omega) - D_1^n(\omega))h_1^n = -h^n$$

By the nondegeneracy assumption $h^{n\omega}(v^n(\omega)) > 0$. It follows that $v_1^n(\omega) - D_1^n(\omega) > 0$, and $h_1^{n\omega}(v^n(\omega)) < 0$.

Since the corresponding conditions apply to Player 2,

$$\frac{v_2^n(\omega) - D_2^n(\omega)}{v_1^n(\omega) - D_1^n(\omega)} = \frac{h_1^{n\omega}(v^n(\omega))}{h_2^{n\omega}(v^n(\omega))}$$

As noted earlier it must be the case that v lies on the boundary of $B^+(\omega)$ (denoted $\overline{B}^+(\omega)$).

It follows that for all $\varepsilon > 0$, there exists $\overline{n}$ such that for all $\psi^{n\omega}(v^n(\omega)) \equiv -\frac{h_1^{n\omega}(v^n(\omega))}{h_2^{n\omega}(v^n(\omega))}$ (the slope of the iso-probability line at $v^n(\omega)$) satisfies $\underline{s}(v^n(\omega)) - \varepsilon \leq \psi^{n\omega}(v^n(\omega)) \leq \overline{s}(v^n(\omega)) + \varepsilon$.

It follows that $\frac{v_2(\omega)-D_2(\omega)}{v_1(\omega)-D_1(\omega)} = -s$ for some $s \in [\underline{s}(v), \overline{s}(v)]$. By Nash (1950, 1953), if v is on the boundary of $B(\omega)$ and $D(\omega) \ll b$ for some $b \in B(\omega)$, then the preceding condition is satisfied if and only if $v(\omega) = N(D(\omega), B(\omega))$. The argument used earlier to establish that $v(\omega)$ and $w(\omega)$ lie on the boundary of $B(\omega)$ imply that when $D(\omega) \ll b$ for some $b \in B$ then $h^{n\omega}(v^n) \to 1$ and consequently $v(\omega) = w(\omega)$.

Step 3: If $D(\omega) < \overline{b}(\omega)$ then $w(\omega) = N(D(\omega), B(\omega))$.

If $D(\omega)$ is efficient (that is, $D(\omega) \in \overline{\overline{B}}(\omega)$) then $D(\omega) = N(D(\omega), B(\omega))$ and $w(\omega) = D(\omega)$. If not, there exists some $b \in B$ such that $D(\omega) \ll b$. Now invoke Step 2.

Step 4: $D(\omega) < \overline{b}(\omega)$.

If $D_1(\omega) = \overline{b}_1(\omega)$ ($\geq V_1(\omega; m)$) then (for large n) $D_1^n(\omega) > w_1^n(\omega)$. Since $D_1^n(\omega)$ is a lower bound for Player 1's payoff in the game with initial state ω this yields a contradiction to the initial supposition that $w_1(\omega) < V_1(\omega; m)$. Now suppose $D_2(\omega) = \overline{b}_2(\omega)$. Then $w_2(\omega) = \overline{b}_2(\omega) = N_2(D(\omega), B(\omega))$. Suppose $w_1(\omega) < N_1(D(\omega), B(\omega))$. Consider deviation by 1 to $\underline{m}_1(\omega)$ and consider a subsequence along which all relevant quantities converge. Denote the new limit disagreement payoff $\underline{D}(\omega)$. Then $\underline{D}_2(\omega) < \bar{b}_2(\omega)$. If $\underline{D}_1(\omega) \geq N_1(D(\omega), B(\omega))$ we have obtained our contradiction. If not, there exists b (in particular, we may use $b = N(D(\omega), B(\omega))$ such that $b > \underline{D}(\omega)$. Now we may use the same argument as in Step 2 to obtain a contradiction.

We have therefore established that for all ω, $w(\omega) = V(\omega; m)$.

Recall the notation from the preamble to Lemma 6.

Step 5: m is 'locally optimal' for all ω.

Let $b(m) = V(\omega; m)$.

Suppose not and suppose w.l.o.g. that

$$\xi_{1\omega}(b(m); (m_1'(\omega), m_2(\omega))) > \xi_{1\omega}(b(m); (m_1^1(\omega), m_2(\omega))) = b_{i\omega}(m) = V_1(\omega; m)$$

for some $m_1'(\omega) \in M_1(\omega)$.

We argue that for large enough n and in state ω if Player 1 chooses $m_1'(\omega)$ then in the subgame defined by $m_1'(\omega)$ and Player 2's equilibrium threat $m_2^n(\omega)$ Player 1's payoff will strictly exceed $w_1^n(\omega)$, a contradiction. Define $\widetilde{m}^n(\omega) \equiv (m_1'(\omega), m_2^n(\omega))$. Denote by $\widetilde{v}_i^n$ Player i's equilibrium demands in the subgame indexed by $\widetilde{m}^n(\omega)$. Let

$$\widetilde{D}^n(\omega) = (1-\delta)U(\widetilde{m}^n(\omega); \omega) + \delta \sum_{\omega'} \rho(\omega' \mid \omega, \widetilde{m}^n(\omega)) w^n(\omega')$$

$$\widetilde{w}^n(\omega) = \widetilde{v}^n(\omega)\, h^{n\omega}(\widetilde{v}^n(\omega)) + (1 - h^{n\omega}(\widetilde{v}^n(\omega)))\widetilde{D}^n(\omega)$$

denote the disagreement and equilibrium payoff respectively in the subgame (contingent on the– appropriate– assumption that equilibrium behavior will be reverted to in the next round.)

Consider a (sub)-subsequence (for simplicity denote this also by n) such that $\widetilde{v}^n(\omega), \widetilde{D}^n(\omega)$ and $\widetilde{w}^n(\omega)$ converge to some $\widetilde{v}(\omega)$, $\widetilde{D}(\omega)$ and $\widetilde{w}(\omega)$. Of course, $m_2^n(\omega)$ converges to $m_2(\omega)$.

By the earlier argument for large n, $w_1^n(\omega)$ is close to $V_1(\omega; m)$. We mimic this argument to show that similarly $\widetilde{w}_1^n(\omega)$ is close to $\xi_{1\omega}(b(m); (m_1'(\omega), m_2(\omega)))$. The latter, of course, equals $N_1(\widetilde{D}(\omega), B(\omega))$.(See Lemma 2 and preceding definitions).

Step 1 from above applies directly to $\widetilde{v}(\omega)$ and $\widetilde{w}(\omega)$. If $\widetilde{D}(\omega) < b$ for some $b \in B(\omega)$ then repeat Step 2 to obtain the desired conclusion. Similarly Step 3 may be replicated. For Step 4 the case $\widetilde{D}_1(\omega) = \overline{b}_1(\omega)$ yields a contradiction as before and the case $\widetilde{D}_2(\omega) = \overline{b}_2(\omega)$ contradicts the initial hypothesis as in this case we have $\xi_{2\omega}(b(m); (m_1'(\omega), m_2(\omega))) = N_2(\widetilde{D}(\omega), B(\omega)) = \widetilde{D}_2(\omega) = \overline{b}_2(\omega)$, and therefore $V_1(\omega; m) \geq \xi_{1\omega}(b(m); (m_1'(\omega), m_2(\omega)))$. This completes the proof. ■

6 Cooperative Treatment

Nash (1953) gives us an axiomatic theory of how a bargaining problem will be resolved. A bargaining problem consists of a nonempty, compact and convex set B of feasible utility pairs, nonempty finite sets S_1 and S_2 of pure strategies (or "threats") players can employ (they can mix over those pure strategies), and a utility function U mapping $S_1 \times S_2$ into $\mathbb{R}^2$. A theory associates with each bargaining problem a unique solution, an element of the feasible set. Nash proposes a set of axioms such a theory should satisfy; he shows there is exactly one theory consistent with this set.

At first glance, it would appear that a much more elaborate set of axioms is required to address the complexities of a stochastic game with contracts. But adopt the perspective of Section 4: the players in the stochastic game beginning in state ω implicitly face a bargaining problem. Their feasible set is the set of all present discounted expected payoff pairs they can generate by signing contracts today concerning their actions in all contingencies. Their sets of threats are the sets of actions available at ω. How do the players evaluate a pair of threats (m_1, m_2)? They get a flow payoff pair $U(m_1, m_2)$ until the state changes and there is some new opportunity to bargain. At that point, they have encountered a new bargaining problem (the stochastic game beginning in some state ω'), and the theory we are trying to axiomatize says what players should get in that situation. Since the pair (m_1, m_2) determines the arrival rates of transition to other states, one can compute the expected discounted payoff consequences of (m_1, m_2) for each player.

To summarize, a theory assigns to each stochastic game with contracts, a solution pair from its feasible set. If the players believe the theory, these values determine a payoff pair that players expect to result if they adopt a particular threat pair and agreement is not reached. Analogues of Nash's axioms can be applied directly to this family of bargaining problems. The difference between this family and that of Nash (1953) is that for Nash, the threat pair utilities are fully specified by a pair of actions, whereas here they are partially determined by the proposed theory, as explained in the preceding paragraph. This gives rise to a fixed point problem. While we can show existence in great generality, for uniqueness we assume either transferable utility or eventual absorption, as in Sections 4 and 5.

A stochastic bargaining environment $\mathcal{E}$ is defined by a stochastic game $\mathcal{G}$ and a collection of state-dependent bargaining set $B(\omega), \omega \in \Omega$ where $\Pi(\omega) \subset B(\omega)$.

Fix Ω, ρ. We may associate a variety of stochastic bargaining environments E with the above fixed elements [By varying $B(.)$'s, M's, etc.]

Definition 2 *A value v^* for a stochastic bargaining environment E specifies for each $\omega \in \Omega$ a unique element $v^*(\omega) \in B(\omega)$.*

Definition 3 *A solution specifies a unique value for each $\mathcal{E}$.*

Axioms on Solution:

Axiom 1 *Pareto optimality*

Axiom 2 *Independent of Cardinal Representation.*

Consider $\mathcal{E}$ and $\mathcal{E}'$ where $\mathcal{E}'$ is identical to $\mathcal{E}$ except that for some $a_i > 0$ and b_i, $i = 1, 2$, utility values u_i in $\mathcal{E}$ are transformed to

$$u_i' = a_i u_i + b_i \text{ in } \mathcal{E}'.$$

Then

$$v_i^*(\omega; \mathcal{E}') = a_i v_i^*(\omega; \mathcal{E}) + b_i \qquad \forall \omega, i = 1, 2.$$

Axiom 3 *"Local" determination / IIA*

Suppose $\mathcal{E}$ and $\mathcal{E}'$ are stochastic bargaining environments and are identical except that $B'(\omega) \subseteq B(\omega)$ $\forall \omega$. If for all ω $v^*(\omega; \mathcal{E}) \in B'(\omega)$ then

$$v^*(\omega; \mathcal{E}') = v^*(\omega; \mathcal{E}) \quad \forall \omega$$

For bargaining environments $\mathcal{E}$ with a single threat pair (m_1, m_2), the disagreement payoff at state ω is denoted $D(\omega)$ and is defined endogenously in terms of the solution as follows:

$$D(\omega) = (1-\delta)u(m_1, m_2; \omega) + \delta \sum_{\omega'} \rho(\omega' \mid \omega, m(\omega))v^*(\omega')$$

where v^* is the value specified by the solution for $\mathcal{E}$.

Axiom 4 *SYMMETRY*

Suppose a bargaining environment $\mathcal{E}$ has a single threat pair (m_1, m_2) and at some state ω, $B(\omega)$ is symmetric and $D_1(\omega) = D_2(\omega)$. Then $v_1^*(\omega) = v_2^*(\omega)$.

Axiom 5 *Suppose $M_1^{'} \subseteq M_1$. Then $v_1(\omega; M_1', M_2) \leq v_1(\omega; M_1, M_2) \quad \forall \omega$.*

Axiom 6 *For all $m_1 \in M_1$ there exists $m_2 \in M_2$ s.t.*

$$v_1^*(\omega; \{m_1\}, \{m_2\}) \leq v_1^*(\omega; M_1, M_2)$$

The first four axioms are the most familiar, as they appear in Nash (1950) as well as Nash (1953). The final two axioms are analogous to two Nash added in 1953 to handle endogenous threat points. Axiom 5 says that a player is (weakly) strengthened by having access to more threats. Axiom 6 says that if Player 1's set of threats is reduced to a singleton $\{m_1\}$, and 2's threat set is reduced to a singleton in the most favorable way for 2, then 2 is not hurt by the changes. This is compelling if, in some sense, threats don't exert influence "as a group" against a singleton threat of an opponent.

(To be completed)

7 Example

8 Conclusion

When two persons have different preferences about how to cooperate, what should each of them threaten to try to gain advantage, and what will the ultimate outcome be? For static bargaining situations, Nash (1953) proposes a solution, and presents both axiomatic and noncooperative strategic analyses that isolate his solution. We translate his results into a real-time setting, and then allow for dynamic phenomena such as random changes in the environment, learning by doing, investment in physical and human capital, and so on. Our extensions of Nash's axiomatic and noncooperative approaches agree on a unique division of surplus in a wide class of stochastic games with contracts, and on what actions to take to influence the outcome in one's favor.

As a simple example of the strategic dynamics that can be captured, we show that a weak rival can extort a surprising amount of money from a stronger competitor by threatening to enter the market (even if this would be at great loss to the weaker party). If gaining access to the market is costly to the potential entrant, the theory offers a prediction about the optimal rate of investment in the technology needed for entry.

Our adaptation of Nash's perturbed demand game to the stochastic game setting is perhaps more convincing than his original story in the static case: when an accidental failure of bargaining occurs (because of random perturbations), we don't need to insist that the inefficient threat actions will be carried out in perpetuity. Rather, they will be reconsidered when another opportunity to bargain arises. Nonetheless, we think there is a still more plausible noncooperative story that justifies our proposed solution. In ongoing work we show that small behavioral perturbations of the stochastic game lead to "war of attrition" equilibria whose expected payoffs coincide with those proposed here.

References

[1] Nash, J. (1950a), "The Bargaining Problem," Econometrica, 18: 155—162.

[2] Nash, J. (1953), "Two-Person Cooperative Games," Econometrica, 21: 128—140.