6 Repeated Games with Complete Information

Motivating example: The Prisoners' dilemma, again.

Note: payoffs here have been increased by 5 to avoid negative numbers

Player 2

Player 1		Not Fink	Fink
	Not Fink	3, 3	0, 4
	Fink	4, 0	1, 1

Dominant strategy solution is (Fink, Fink)

Cooperation may seem more plausible. One reason: participants may expect to interact more in the future. If you "Fink" today, your opponent may retaliate in the future.

Simple device for getting at these issues: imagine that the game is repeated.

Two important cases: (i) potentially infinite repetitions, (ii) finite repetitions.

6.1 Infinitely repeated games

6.1.1 Some preliminaries

Terminology: The game played in each period is called the *stage game*. The dynamic game formed by infinite repetitions of a stage game is called a *supergame*.

Observations:

(i) Even if the stage game is finite, the associated supergame is not.

(ii) There are no terminal nodes. How do we assign payoffs?

Evaluating payoffs:

What we need: a mapping from strategy profiles into expected payoffs. (For finite games, this is given by the composition of the mapping from strategy profiles into distributions over terminal nodes, with the mapping from terminal nodes to payoffs.)

For repeated games, one can assume that payoffs are distributed immediately after each play of the stage game. Then any strategy profile maps to a distribution of paths through the game tree, and any path through the game tree maps to a sequence of payoffs for each player i, $v_i = (v_i(1), v_i(2), ...)$.

Remaining issue: We need a mapping from strategy profiles to scalar payoffs for each player. How do we get from payoff sequences to scalar payoffs?

General answer: assume that players have utility functions mapping sequences of payoffs into utility: $u_i(v_i)$

Some specific answers:

(i) Use discounted payoffs: $u_i(v_i) = \sum_{t=1}^{\infty} \delta^{t-1} v_i(t)$. Note: the discount factor may reflect both time preference and a probability of continuation.

(ii) For the case of no discounting, we can use the average payoff criterion:

$$u_i(\pi_i) = \lim_{T \to \infty} \left(\frac{1}{T} \right) \sum_{t=1}^{T} v_i(t)$$

Another possibility for the case of no discounting: the *overtaking criterion*.

Strategies:

The nature of strategies will depend upon assumptions about what is observed each time the game is played.

For the time being, we will assume that, each time the stage game is played, all players can observe all previous choices.

With this assumption, each sequence of choices up to (but not including) period t corresponds to a separate information set (for each player) in period t. Consequently, we proceed as follows.

Let $a_i(t)$ denote the action taken by player i in period t

Let $a(t) = (a_1(t), ..., a_I(t))$ be the profile of actions chosen in period t

A t-history, $h(t)$, is a sequence of action profiles $(a(1), ..., a(t-1))$, summarizing everything that has occurred prior to period t.

Since, by assumption, $h(t)$ is observed by all players, there is, for each player, a one-to-one correspondence between t-histories and period t information sets.

Consequently, a strategy is a mapping from all values of $t \in \{1, 2, ...\}$ and all possible $t-$histories to period t actions (for period 1, the set of 1-histories is degenerate).

6.1.2 Nash equilibria with no discounting

The prisoners' dilemma

Illustration of a t-history: $((NF, NF), (F, NF), (F, F), ..., (NF, F))$

Example of a strategy: $\sigma_i^N(t, h(t)) \equiv F$ for all t, $h(t)$.

Note: this repeats the equilibrium of the stage game in every period.

Claim: If players use the average payoff criterion, $\left(\sigma_1^N, \sigma_2^N\right)$ is a Nash equilibrium.

Demonstration: Check to see whether a player can gain by deviating from this strategy, given that his opponent plays this strategy.

If the player sticks to the strategy, the sequence of payoffs will be $v_i(t) = 1$ for all t. The average payoff is 1.

If player i deviates to any other strategy while j sticks to σ_j^N, the sequence of payoffs for i contains zeros and one, so the average payoff cannot exceed 1.

Conclusion: Repeating the equilibrium of the stage game is a Nash equilibrium for the supergame.

Remark: The same proposition obviously holds with discounting, and without discounting using the overtaking criterion (a sequence of ones always beats a sequence of ones and zeros).

Exercise: Prove that this point is completely general (it holds for all stage games).

Question: Can we get anything other than repetitions of the stage game equilibrium?

Another possible strategy:

$$\text{In period 1, play } NF$$

$$\text{In period } t > 1\text{, play } \begin{cases} F \text{ if } h(t) \text{ contains an } F \\ \\ NF \text{ otherwise} \end{cases}$$

With these strategies, the game would unfold as follows: Both players would play NF forever. If any player ever deviated from this path, then subsequently both players would play F forever.

Let's imagine that both players select this strategy.

Claim: If players use the average payoff criterion to evaluate payoffs, this is a Nash equilibrium.

Demonstration: Check to see whether a player can gain by deviating from this strategy, given that his opponent plays this strategy.

If the player sticks to the strategy, the sequence of payoffs will be $v_i(t) = 3$ for all t. The average payoff is 3.

Now consider a deviation to some other strategy. Let t' be the first period t for which this strategy dictates playing F when $h(t)$ does not contain an F (if there is no such t', then the deviation also generates an average payoff of 3). If the player deviates to this strategy, the sequence of payoffs will be

$$v_i(t) \begin{cases} = 3 \text{ for } t < t' \\ = 4 \text{ for } t = t' \\ \leq 1 \text{ for } t > t' \end{cases}$$

(For $t > t'$, this follows because the opponents strategy will always dictate playing F). The associated average payoff is not larger than 1, and therefore certainly less than 3. Consequently, this is a Nash equilibrium.

Remarks: (i) This example demonstrates that cooperation is possible. Cooperation is sustained through the threat of punishment.

(ii) We cannot necessarily claim that the players will cooperate in this way, because there are many other Nash equilibria.

Illustration: For all t, let $h^*(t)$ be the t-history such that (i) $a_1(t') = F$ for $t' < t$ odd, and $a_1(t') = NF$ for $t' < t$ even; (ii) $a_2(t') = NF$ for all $t' < t$. In words: player 1 has alternated between NF and F, while player 2 has always chosen NF.

Consider the following strategies:

$$\sigma_1(t, h(t)) = \begin{cases} NF \text{ if } h(t) = h^*(t) \text{ and } t \text{ is even} \\ F \text{ otherwise} \end{cases}$$

$$\sigma_2(t, h(t)) = \begin{cases} NF \text{ if } h(t) = h^*(t) \\ \\ F \text{ otherwise} \end{cases}$$

In other words, player 1 alternates between NF and F, while player 2 always plays NF (the result is $h^*(\infty)$). However, if either deviates from this path, both subsequently play F forever.

Claim: When players use the average payoff criterion, there is also a Nash equilibrium wherein both players select the preceding strategy.

Demonstration: As long as no player deviates, player 1's payoffs alternate between 3 and 4, while 2's payoffs alternate between 3 and 0. Average payoffs are 3.5 for player 1, and 1.5 for player 2.

Now imagine that player i considers deviating to some other strategy. If this deviation has any effect on the path of outcomes (and hence on payoffs), there must be some period t' in which the actions taken diverge from $h^*(\infty)$. Assuming that player j sticks with its equilibrium strategy, $v_i(t) \leq 1$ for $t > t'$. Consequently, player i's average payoff is 1. This is less than the payoff received by both players in equilibrium.

Question: What other outcomes are consistent with equilibrium?

We will proceed by the process of elimination.

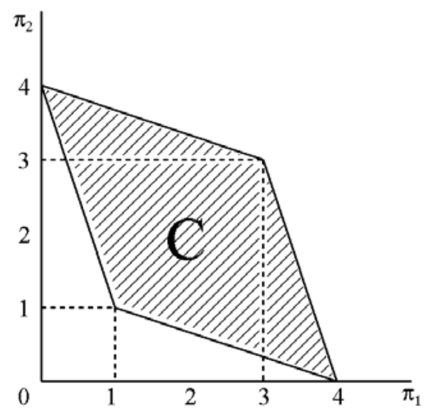
Some definitions

First, we identify the set of feasible payoffs (including things that can be achieved through arbitrary randomizations).

Let $C = \{w \mid w$ is in the convex hull of payoff vectors from pure strategy profiles in the stage game$\}$

Remark: C is potentially larger than the set of payoffs achievable through mixed strategies in the stage game, since we allow for correlations.

For our example (the prisinors' dilemma):



Can anything in C occur in equilibrium as an average payoff? No. Each player can assure himself of a payoff of at least unity each period by playing F all of the time. Therefore, we know that no player can get a payoff smaller than unity.

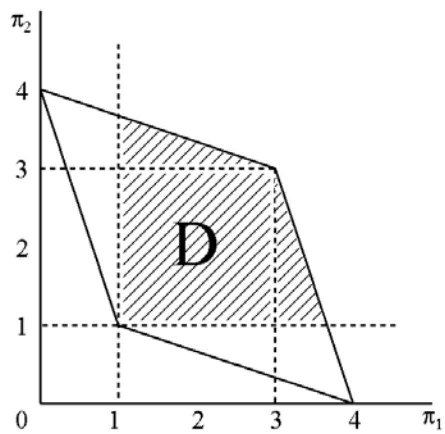
More generally, we define player i's *minmax* payoff:

$$\pi_i^m = \min_{(\delta_1,\ldots,\delta_{i-1},\delta_{i+1},\ldots,\delta_I)} \max_{\delta_i} \pi_i(\delta)$$

This is the average payoff that player i can assure himself, simply by making a best response to what everyone else is supposed to do (according to their strategies) in every period. Player i cannot receive an average payoff less than π_i^m in equilibrium.

Consequently, define $D = \{w \mid w \in C \text{ and } w_i \geq \pi_i^m \text{ for all } i\}$

For our example (the prisoners' dilemma):



D is called the set of *feasible and individually rational payoffs.*

Finally, define $E = \{w \mid \text{there is a Nash equilibrium with average payoff vector } w\}$

Question: How does E compare with D? It's reasonably clear that E isn't larger, but can it be smaller?

The folk theorem

The folk theorem: Consider a supergame formed by repeating a finite stage game an infinite number of times. Suppose that players use the average payoff criterion to evaluate outcomes. Then $E = D$.

Sketch of proof: The proof consists of three steps.

Step 1: $E \subseteq D$. We have already covered this.

Step 2: Consider any sequence of actions $h'(\infty)$ yielding average payoffs $w \in D$. Then there exists a pure strategy Nash equilibrium for the supergame where this sequence of actions is taken on the equilibrium path. We show this by contruction.

Consider the following strategies. If play through period $t-1$ has conformed to $h'(t)$, players continue to follow $h'(\infty)$ in period t. If play has not conformed to $h'(t)$, inspect the actual history $h(t)$ to find the first lone deviator (in other words, ignore any period in which there are multiple deviators). If no lone deviator exists in any period prior to t, then revert to following $h'(\infty)$ in period t. If the first lone deviator is i, then all $j \neq i$ play

$$\delta_{-i}^m = \arg \min_{(\delta_1,\ldots,\delta_{i-1},\delta_{i+1},\ldots,\delta_I)} \max_{\delta_i} \pi_i(\delta)$$

while i plays some arbitrarily assigned action.

It is easy to check that this is a Nash equilibrium. If all players choose their equilibrium strategies, the outcome is $h'(\infty)$, and the average payoff for i is $w_i \geq \pi_i^m$ (since $w \in D$ by assumption). If player i deviates, then, assuming all others play their equilibrium strategies, i will be the first deviator, and subsequently can do not better than π_i^m in any period. This means that i's average payoff will be no greater than π_i^m. Thus, the deviation does not benefit i.

Step 3: For all $w \in D$, there exists a sequence of actions yielding average payoffs of w.

Idea: alternate actions to produce the same frequencies as the randomization. This is easy if the randomization involves rational frequencies. If it involves irrational frequencies, one varies the frequency in the sequence to achieve the right frequency in the limit.

Interpretation of the folk theorem: (i) Anything can happen. Comparative statics are problematic.

(ii) The inability to write binding contracts is not very damaging. Anything attainable through a contract is also obtainable through a self-enforcing agreement, at least with no discounting. The equilibrium that gets played is determined by a process of negotiation. It is natural to expect players to settle on some self-enforcing agreement that achieves the efficient frontier. The precise location may depend upon bargaining strengths.

6.1.3 Nash equilibria with discounting

Now imagine that players evaluate payoffs according the utility function $u_i(v_i) = \sum_{t=1}^{\infty} \delta^{t-1} v_i(t)$. For simplicity, take the rate of discounting $\delta \in (0,1)$ to be common for all players.

Remark: We can think of δ as the product of a pure rate of time preference, ρ, and a continuation probability, λ (measuring the probability of continuing the game in period $t+1$, conditional upon having reached t): $\delta = \rho\lambda$. In particular, assume that, if the game ends, subsequent payoffs are zero (this is just a normalization). Let T be the realized horizon of the game. Then expected payoffs are

$$
\begin{aligned}
& \sum_{k=1}^{\infty} \left[prob(T = k) \sum_{t=1}^{k} \rho^{t-1} v_i(t) \right] \\
= & \sum_{k=1}^{\infty} \left[\lambda^{k-1}(1-\lambda) \sum_{t=1}^{k} \rho^{t-1} v_i(t) \right] \\
= & \sum_{k=1}^{\infty} \sum_{t=1}^{k} \left[\lambda^{k-1}(1-\lambda) \rho^{t-1} v_i(t) \right] \\
= & \sum_{t=1}^{\infty} \sum_{k=t}^{\infty} \left[\lambda^{k-1}(1-\lambda) \rho^{t-1} v_i(t) \right] \\
= & \sum_{t=1}^{\infty} \left[\rho^{t-1} v_i(t) \, (1-\lambda) \, \lambda^{t-1} \sum_{k=t}^{\infty} \lambda^{k-t} \right] \\
= & \sum_{t=1}^{\infty} (\lambda\rho)^{t-1} \, v_i(t)
\end{aligned}
$$

The magnitude of δ in any context will depend upon factors such as the frequency of interaction, detection lags, and interest rates.

The prisoners' dilemma

Imagine again that both players use the following strategies:

$$\text{In period 1, play } NF$$

$$\text{In period } t > 1, \text{ play } \begin{cases} F \text{ if } h(t) \text{ contains an } F \\ NF \text{ otherwise} \end{cases}$$

If players discount payoffs at the rate δ, is this a Nash equilibrium?

If the player sticks to the strategy, the sequence of payoffs will be $v_i(t) = 3$ for all t. The discounted payoff is

$$\sum_{t=1}^{\infty} 3\delta^{t-1} = \frac{3}{1-\delta}$$

Now consider a deviation to some other strategy. Without loss of generality, imagine that player i deviates to F in period 1. Player i knows that j will play F in all subsequent periods (since this is dictated by j's strategy). Consequently, it is optimal for i to play F in all subsequent periods, having deviated in the first. Thus,

$$v_i(t) \begin{cases} = 4 \text{ for } t = 1 \\ = 1 \text{ for } t > 1 \end{cases}$$

Player i's discounted payoffs are

$$4 + \sum_{t=2}^{\infty} \delta^{t-1} = 4 + \frac{\delta}{1-\delta}$$

Player i therefore finds the best deviation unprofitable when

$$\frac{3}{1-\delta} \geq 4 + \frac{\delta}{1-\delta}$$

This is equivalent to

$$\delta \geq \frac{1}{3}$$

Thus, the strategies still constitute an equilibrium provided that the players do not discount the future too much.

This is also true for the other equilibrium considered above.

The folk theorem

Analysis of prisoners' dilemma suggests that it becomes possible to sustain cooperative outcomes as δ gets larger.

A natural conjecture: it is possible to sustain all feasible, individually rational cooperative outcomes as $\delta \to 1$.

Problem: as $\delta \to 1$, discounted payoffs become unbounded. To discuss what occurs in the limit, we need to renormalize payoffs.

Renormalize payoffs as follows:

$$u_i(v_i) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v_i(t)$$

Notice we can now think of utility as a weighted average of the single period payoffs,

$$u_i(v_i) = \sum_{t=1}^{\infty} \mu_t v_i(t),$$

where $\mu_t = (1-\delta)\delta^{t-1}$, and $\sum_{t=1}^{\infty} \mu_t = 1$.

With this normalization, the set of feasible discounted payoffs is C.

Moreover, as $\delta \to 1$, this converges to the average payoff criterion.

In this setting, the folk theorem needs to be restated slightly: for any $w \in D$ with $w_i > \pi_i^m$ for all i, there exists $\delta^* < 1$ such that for all $\delta \in (\delta^*, 1)$, w is the payoff vector for some Nash equilibrium.

The proof is very similar to that of the original folk theorem.

6.1.4 Subgame perfect Nash equilibria

The equilibria constructed to establish the folk theorem may not be subgame perfect. Punishing a player through a minmax strategy profile may not be credible, since the punishers may suffer. Can we sustain cooperation in $SPNE$?

The prisoners' dilemma

Claim: All of the Nash equilibria considered above for the repeated prisoners' dilemma are $SPNE$.

Demonstration: All Nash equilibrium strategies necessarily consistute Nash equilibria in all subgames that are reached along the equilibrium path. For any subgame off the equilibrium path, the prescribed strategies are $\left(\sigma_1^N, \sigma_2^N\right)$. All subgames are identical to the original game, and $\left(\sigma_1^N, \sigma_2^N\right)$ is a Nash equilibrium for the original game. Thus, we have a Nash equilibrium in every subgame.

Nash reversion

Generalization: The players can attempt to support cooperation by using repetitions of a stage-game equilibrium as a punishment. These punishments are known as *Nash reversion* (the players attempt to cooperate, but revert to a static Nash outcome if someone deviates).

Formally, consider some arbitrary stage game, as well as the supergame consisting of the infinitely repeated stage game. Assume that the stage game has at least one Nash equilibrium. For any particular stage-game Nash equilibrium, a^*, consider the following strategies:

$$\sigma_i^N(t, h(t)) = a_i^* \text{ for all } t,\ h(t)$$

Now imagine that we want to support some particular outcome, $h^*(\infty)$, as an equilibrium path. Let's try to do this with the following strategies: If play through period $t-1$

has conformed to $h^*(t)$, players continue to follow $h^*(\infty)$ in period t. If play has not conformed to $h^*(t)$, then players use σ^N.

Claim: If the aforementioned strategies constitute a Nash equilibrium, then the equilibrium is subgame perfect.

Demonstration: We use precisely the same argument as for the prisoners' dilemma. All Nash equilibrium strategies necessarily constitute Nash equilibria in all subgames that are reached along the equilibrium path. For any subgame off the equilibrium path, the prescribed strategy profile is σ^N. All subgames are identical to the original game, and σ^N is a Nash equilibrium of the original game. Thus, we have a Nash equilibrium in every subgame.

Remark: Since σ^N is an equilibrium of the supergame irrespective of whether the players use discounted payoffs, average payoffs, or the overtaking criterion, this claim is equally valid for all methods of evaluating payoffs.

Implication: When one uses Nash reversion to punish deviations, it is particularly simple to build $SPNE$ and check subgame perfection: one simply makes sure that the equilibrum is Nash (equivalently, that no player has an incentive to deviate from a prescribed choice on the equilibrium path).

The folk theorem

For some games, Nash reversion is as severe as minmax punishments.

Examples:

(i) The prisoners' dilemma

(ii) Bertrand competition (minmax payoffs and Nash payoffs are both zero)

However, Nash reversion is frequently much less severe than minmax punishments.

Example: Cournot competition (minmax payoffs are 0, while Nash profits are strictly positive)

Question: Does the validity of the folk theorem depend, in general, on the ability to use non-credible punishments (at least for stage games with the property that Nash payoffs exceed minmax)?

Answer: Subject to some technical conditions, one can prove versions of the folk theorem (with and without discounting) for $SPNE$. The proofs are considerably more difficult.

Implication: If the stage-game Nash equilibrium payoffs exceed minmax payoffs, then, for δ sufficiently close to unity, there exist more severe punishments than Nash reversion.

Example:

		Player 2	
		Not Fink	Fink
Player 1	Not Fink	3, 3	0, 4
	Fink	4, 1	1, 0

This game has only one Nash equilibrium: (F, NF). Note that this gives player 1 the maximum possible payoffs. It is therefore impossible to force player 1 to do anything through Nash reversion.

Exercise: For this example, construct a $SPNE$ in which (NF, NF) is chosen on the equilibrium path. Either use the average payoff criterion, or assume an appropriate value for δ.

We will see another explicit example of punishments that are more severe than Nash reversion when we analyze the dynamic Cournot model.

6.1.5 A short list of other topics

1. Repeated games with imperfect observability of actions.
2. Repeated games with incomplete information (reputation)
3. Heterogeneous horizons (models with overlapping generations, or both short and long-lived players)
4. Renegotation

6.2 Finitely repeated games

One might think that finitely repeated games get to look a lot like infinitely repeated games when the horizon is sufficiently long. This is correct for Nash equilibria (where credibility is not required), but not for subgame perfect equilibria.

Theorem: Consider any finitely repeated game. Suppose that there is a unique Nash equilibrium for the stage game. Then there is also a unique $SPNE$ for the repeated game, consisting of the repeated stage game equilibrium

Proof: By induction (with T denoting the number of repetitions).

For $T = 1$, the repeated game is the stage game, which has a unique Nash equilibrium

Now assume the theorem is true for $T - 1$. Consider the T-times repeated game. All subgames beginning in the second period simply consist of the of the $(T - 1)$-times repeated game, which, by assumption, has a unique $SPNE$. Thus, in a $SPNE$, actions taken in the first period have no effect on choices in subsequent periods. In equilibrium first period choices must therefore be mutual best responses for the stage game. This means that the first period choices must be the Nash equilibrium choices for the stage game. Q.E.D.

Remarks:

(i) It is often said that a finitely repeated game "unravels" from the end, much like the centipede game.

(ii) Cooperation may be possible when the stage game has multiple Nash equilibria.

Example:

		Player 2		
		A	B	C
	a	4, 4	0, 0	0, 5
Player 1	b	0, 0	3, 3	0, 0
	c	5, 0	0, 0	1, 1

There are two Nash equilibria: (b, B) and (c, C). (a, A) is Pareto superior, but it is not a Nash equilibrium.

Imagine that the game is played twice in succession without discounting.

Strategies: Play a (A) in the first period. If the outcome in the first period was (a, A), play b (B) in the second period; otherwise, play c (C).

This is plainly a Nash equilibrium: any other strategy yields a gain of at most 1 unit in the first period, and a loss of at least 2 in the second period.

It is also a $SPNE$, since it prescribes a Nash equilibrium in every proper subgame.

Remark: There are folk theorems for finite horizon games formed by repetitions of stage games that possess multiple equilibria.

6.3 Applications

6.3.1 The repeated Bertrand model

Stage game: $N \geq 2$ firms simultaneously select price. Customers purchase from the firm with the lowest announced price, dividing equally in the event of ties. Quantity purchased is given by a continuous, strictly decreasing function $Q(P)$. Firms produce with constant marginal cost c. Let

$$\pi(p) \equiv (p - c)Q(p)$$

We assume that $\pi(p)$ is increasing in p on $[c, p^m]$ (where p^m is the monopoly price).

Observation: (i) Nash reversion involves setting $p = c$, which generates 0 profits. This is also the minmax profit level. Thus, Nash reversion generates the most severe possible punishment. Anything that can be sustained as an equilibrium outcome can be sustained using Nash reversion as punishments. Therefore, we can, without loss of generality, confine attention to equilibria that make use of Nash reversion.

(ii) The static Bertrand solution is unique. Thus, we know that no cooperation can be sustained in $SPNE$ for finite repetitions. Henceforth, we focus on infinite repetitions.

Analysis of equilibria:

Consider the following $h(\infty)$: both firms select some price $p^* \in [c, p^m]$ in every period.

Assuming that players discount future utility, when can we sustain this path as the outcome of a $SPNE$?

Given the preceding observation, we answer this question by determining the conditions under which this outcome can be supported as a Nash equilibrium using Nash reversion.

In equilibrium each firm receives a payoff of

$$\sum_{t=1}^{\infty} \frac{\pi(p^*)}{N} \delta^{t-1}$$

If a firm deviates to a price higher than p, it obviously earns nothing. If it deviates to a price below p, it will earn nothing in subsequent periods (since price will be driven to marginal cost), and its current period profits are bounded above by $\pi(p^*)$. Thus, no firm has an incentive to deviate provided that

$$\sum_{t=1}^{\infty} \frac{\pi(p^*)}{N} \delta^{t-1} \geq \pi(p^*)$$

This is equivalent to

$$\frac{1}{1-\delta} \geq N$$

which in turn implies

$$\delta \geq \frac{N-1}{N}$$

Provided discounting is not too great, cooperation is possible.

Implications:

(i) Cooperation becomes more difficult with more firms. For $N = 2$, cooperation is sustainable iff $\delta \geq \frac{1}{2}$. As $N \to \infty$, the threshhold discount factor converges to unity.

(ii) The equilibrium condition is independent of π, and therefore independent of p^* (the price we are attempting to sustain). For the Bertrand game, either everything is sustainable, or nothing is sustainable.

(iii) There is no longer a sharp discontinuity between one firm and two, as in the static Bertrand model. However, given (ii), there is still a sharp discontinuity between some N and $N+1$, where the best cooperative equilibrium shifts from monopoly to perfect competition.

6.3.2 The repeated Cournot model

Stage game: $N = 2$ firms simultaneously select quantities. The market clearing price is given by $Q(P) = a - bQ$. Firms produce with constant marginal cost c.

Let $Q^m = \frac{a-c}{2b}$ denote monopoly quantity, and let $\pi^m = \frac{(a-c)^2}{4b}$ denote monopoly profits.

Let $q^c = \frac{a-c}{3b}$ denote Cournot duopoly quantity, and let $\pi^c = \frac{(a-c)^2}{9b}$ denote Cournot duopoly profits (both per firm).

Observation: If the static Cournot equilibrium is unique, we know that it is impossible to sustain cooperation in $SPNE$ for finitely repeated games. Henceforth we focus on infinitely repeated games.

Analysis of equilibria using Nash reversion:

Consider the following $h(\infty)$: each firm sets $\frac{Q^m}{2}$ in every period.

Assuming that players discount future utility, when can we sustain this path as the outcome of a $SPNE$, using Nash reversion?

In equilibrium each firm receives a payoff of

$$\sum_{t=1}^{\infty} \frac{\pi^m}{2} \delta^{t-1} = \frac{1}{1-\delta} \frac{(a-c)^2}{8b}$$

Imagine instead that the firm makes a static best response to $\frac{Q^m}{2}$ (this is its best possible deviation). Best response profits given that the rival plays $\frac{Q^m}{2}$ are $\frac{9}{64} \frac{(a-c)^2}{b}$. In every subsequent period (after the deviation occurs), the deviating firm earns the static Cournot profits, π^c. The deviation therefore yields discounted profits of

$$\frac{9}{64}\frac{(a-c)^2}{b} + \frac{\delta}{1-\delta}\frac{(a-c)^2}{9b}$$

The proposed strategies therefore form an equilibrium iff

$$\frac{9}{64}\frac{(a-c)^2}{b} + \frac{\delta}{1-\delta}\frac{(a-c)^2}{9b} \leq \frac{1}{1-\delta}\frac{(a-c)^2}{8b}$$

Rearranging this expression yields

$$\delta \geq \frac{9}{17} > \frac{1}{2}$$

Implication: Using Nash reversion, it is easier to get cooperation with Bertrand than with Cournot. In the static setting, Bertrand is more competitive. Consequently, Nash reversion punishments are more severe.

Exercise: We know that Cournot profits decline with the number of firms. This means that, for the repeated Cournot game, Nash reversion punishments become more severe with more firms. Does the preceeding "implication" mean that, for Cournot, cooperation is easier to sustain with more firms? If not, why not?

Remark: The preceeding concerns the sustainability of the monopoly outcome. One can perform a similar calculation for other quantities. In contrast to the Bertrand model, it turns out that it is easier to sustain less cooperative outcomes (that is, the threshhold discount factors are lower). Indeed, for Cournot, it is possible to sustain some degree of cooperation (profits in excess of π^c) for all $\delta > 0$. This is a consequence of the envelope theorem: as one reduces quantities starting at the Cournot equilibrium, the improvement in profits is first-order, but the change in the difference between profits and best-deviation profits is second order.

Exercise: For the linear Cournot model, solve for the most profitable symmetric equilibrium sustained by Nash reversion, as a function of the discount factor, δ.

Alternative punishments

Motivation: From the folk theorem, it is obvious that more severe punishments may be available than Nash reversion. In principle, the associated strategies could be extremely complex, which would make them difficult to analyze.

Under some circumstances, however, it is possible to characterize the most severe punishments within large classes of strategies, and to show that the associated strategies have a relatively simple "stick and carrot" structure. We illustrate using the Cournot model.

Definitions: A symmetric stick-and-carrot equilibrium for the repeated Cournot model is characterized by two levels of quantity, q^L and q^H, with $q^L < q^H$.

Let $g_i(q_i, q_j)$ denote firm i's profits when it produces q_i and j produces q_j. Assume we have chosen q^L and q^H so that $g_i(q^L, q^L) > g_i(q^H, q^H)$.

We define a stick-and-carrot strategy, $\sigma^{sc}(t, h(t))$, by induction on t:

(i) $\sigma^{sc}(1, h(1)) = q^L$ (start by playing q^L).

(ii) Having defined $\sigma^{sc}(t-1, h(t-1))$ for all feasible histories $h(t-1)$, we define $\sigma^{sc}(t, h(t))$ as follows. If $q_i(t-1) = \sigma^{sc}(t-1, h(t-1))$ for $i = 1, 2$, then $\sigma^{sc}(t, (h(t-1), q(t-1))) = q^L$. Otherwise, $\sigma^{sc}(t, (h(t-1), q(t-1))) = q^H$.

In words, the choice between q^L and q^H is always determined by play in the previous period. If firms have played their prescribed choices in the previous period, then they play q^L. If one or both deviated in the previous period, they play q^H.

When both players select stick-and-carrot strategies, play evolves as follows. On the equilibrium path, (q^L, q^L) is played every period. If a firm deviates in a single period $t-1$, both players play (q^H, q^H) in the following period as a punishment, after which they return to (q^L, q^L) forever. Notice that, if a player deviates, this strategy requires the player to participate in its own punishment in the following period by playing q^H. If it refuses and instead deviates in the punishment period, the punishment is prolonged. If, on the other hand, it cooperates in its punishment, the punishment period ends and cooperation is restored. Thus, there is both a "stick" (a one-period punishment) and a "carrot" (a reward for participating in the punishment). Use of the carrot can lead players to willingly participate in a very severe one-period punishment.

Analysis of equilibrium: We now derive the conditions under which $(\sigma^{sc}, \sigma^{sc})$ is a $SPNE$ for the infinitely repeated Cournot game. Using the standard dynamic programming argument, it suffices to check single-period deviations.

Given the stationary structure of the game and of the equilibrium, there are only two deviations to check: from $\sigma^{sc}(t, h(t)) = q^L$, and from $\sigma^{sc}(t, h(t)) = q^H$.

From q^L we have:

$$g_i\left(\gamma_i(q^L), q^L\right) - g_i\left(q^L, q^L\right) \leq \delta\left[g_i\left(q^L, q^L\right) - g_i\left(q^H, q^H\right)\right]$$

From q^H we have:

$$g_i\left(\gamma_i(q^H), q^H\right) - g_i\left(q^H, q^H\right) \leq \delta\left[g_i\left(q^L, q^L\right) - g_i\left(q^H, q^H\right)\right]$$

Specialize to the case where the "carrot" is the monopoly outcome, and the "stick" is the competitive outcome (price equal to marginal cost). That is, $q^L = \frac{a-c}{4b}$ and $q^H = \frac{a-c}{2b}$. Then these expressions can be rewritten as

$$\frac{(a-c)^2}{16b} \leq \delta \frac{(a-c)^2}{8b}$$

$$\frac{(a-c)^2}{64b} \leq \delta \frac{(a-c)^2}{8b}$$

Notice that the second expression is redundant. The first simplifies to $\delta \geq \frac{1}{2}$.

Implications: Since $\frac{1}{2} < \frac{9}{17}$, these strategies allow the firms to sustain the monopoly outcome at lower discount rates than with Nash reversion. Indeed, they can now achieve the monopoly outcome for the same range of discount factors as with the infinitely repeated Bertrand model.

Remark: The stick used here yields zero profits for a single period. One can also use more severe sticks that yield negative profits for a single period. Under some conditions, this allows one to construct punishments that yield zero discounted payoffs. The firms are willing to take losses in the short-term because they expect to earn positive profits in subsequent periods.

6.3.3 Cooperation with cyclical demand

Motivation: There is some evidence indicating that oligopoly prices tend to be countercyclical (oligopolists are more prone to enter price wars when demand is strong). If one thinks in terms of conventional supply and demand curves, this is counterintuitive. Note: the evidence is controversial.

Insight: The ability to sustain cooperation depends generally on the importance of the future relative to the present (we saw this with respect to the role of δ). When the present looms large relative to the future, cooperation is more difficult to sustain. This is what occurs during booms.

Model:

Demand is random. Each period, one of two states, H or L, is realized. The states are equally probable, and realizations are independent across periods. Demand for state i is $Q_i(p)$, with $Q_H(p) > Q_L(p)$ for all p.

N firms acting Bertrand competitors.

Production costs are linear with unit cost c.

Notation:

Let $\pi_k(p)$ represents industry profits in state k with price p:

$$\pi_k(p) \equiv (p - c)Q_k(p)$$

Let π_k^m denote industry monopoly profits in state k:

$$\pi_k^m = \max_p \pi_k(p)$$

Equilibrium analysis:

Conditions for equlibrium:

Consider any stationary, symmetric equilibrium path such that both firms select the price p_H in state H and p_L in state L.

Construct equilibrium strategies using Nash reversion (here, these are the most severe possible subgame perfect punishments since they yield zero profits)

Given the stationary structure of the problem and the usual dynamic programming argument, we need only check to see whether the firms have incentives to make one period deviations in each state.

For state H:

$$\left(\frac{N-1}{N}\right)\pi_H(p_H) \leq \left(\frac{\delta}{1-\delta}\right)\left[\frac{1}{2}\pi_H(p_H) + \frac{1}{2}\pi_L(p_L)\right]\frac{1}{N}$$

For state L:

$$\left(\frac{N-1}{N}\right)\pi_L(p_L) \leq \left(\frac{\delta}{1-\delta}\right)\left[\frac{1}{2}\pi_H(p_H) + \frac{1}{2}\pi_L(p_L)\right]\frac{1}{N}$$

Note that the right-hand sides of these expressions are the same, since the future looks the same irrespective of the current demand state. For any given price, the left-hand side is greater for the high demand state. Therefore, a given price is more difficult to sustain for the high demand state than for the low demand state.

Specialized paramateric assumptions:

Before going further, we will simplify the model by making some parametric assumptions:

(i) $Q = \theta - p$

(ii) $\theta \in \{\theta_L, \theta_H\} = \{1, 2\}$

(iii) $c = 0$

(iv) $N = 2$

Under these assumptions, $\pi_k(p) = p(\theta_k - p)$, $\pi_L^m = \frac{1}{4}$, $p_L^m = \frac{1}{2}$, $\pi_H^m = 1$, and $p_H^m = 1$.

Question: When do we get the full monopoly solution, p_L^m in state L, and p_H^m in state H?

Look back at the constraints. If the constraint is satisfied for monopoly in state H, then it is also satisfied for monopoly in state L. Therefore, we need only check the constraint for state H. Substituting, we have

$$1 \leq \left(\frac{\delta}{1-\delta}\right)\left[\left(\frac{1}{2} \times 1\right) + \left(\frac{1}{2} \times \frac{1}{4}\right)\right]$$

This is equivalent to $\delta \geq \frac{8}{13}$.

Note: Since $\frac{8}{13} > \frac{1}{2}$, it is more difficult to sustain full monopoly here than in the Bertrand model with time-invariant demand.

Question: What happens for lower δ?

We know p_H^m becomes unsustainable for state H. However, the constraint for p_L^m in the low state holds with strict inequality at $\delta = \frac{8}{13}$. Consequently, one would still expect it to hold for slightly smaller δ.

Proceed as follows: Assume that, for some $\delta < \frac{8}{13}$, p_L^m is sustainable. Calculate the highest level of sustainable profits, π_H, for state H. If $\pi_H \geq \pi_L^m$, then the initial assumption is valid, and we have an equilibrium.

To compute the highest level of sustainable profits, π_H, for state H under the aforementioned assumption, we substitute into the equilibrium constraint:

$$\pi_H \leq \left(\frac{\delta}{1-\delta}\right)\left[\frac{1}{2}\pi_H + \left(\frac{1}{2} \times \frac{1}{4}\right)\right]$$

For the highest sustainable level of state H profits, this constraint holds with equality. Rearranging yields

$$\pi_H^\delta = \frac{\delta}{8-12\delta}$$

One can check the following:

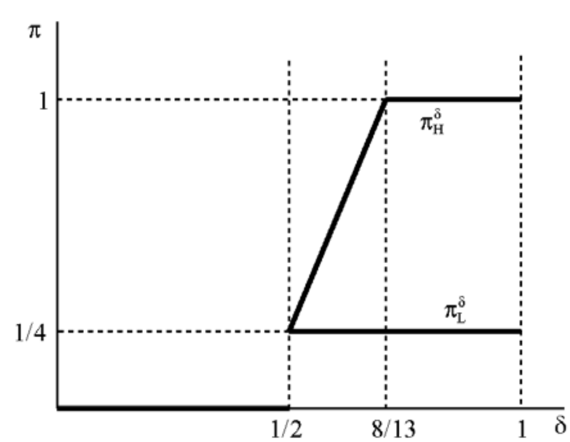
$$\text{For } \delta = \frac{8}{13}, \;\; \pi_H^\delta = 1 = \pi_H^m$$
$$\text{For } \delta = \frac{1}{2}, \; \pi_H^\delta = \frac{1}{4} = \pi_L^m$$

Thus, as long as $\delta \in \left[\frac{1}{2}, \frac{8}{13}\right]$, the assumption that π_L^m is sustainable is valid.

Exercise: Verify that, when $\delta < \frac{1}{2}$, the only $SPNE$ outcome involves repetitions of the static Bertrand outcome (price equal to marginal cost). As in the standard repeated Bertrand model, no cooperation is sustainable for discount factors below $\frac{1}{2}$.

Properties of equilibrium:

(i) π_H^δ (sustainable profits in the high demand state) is increasing in δ. Graphically:



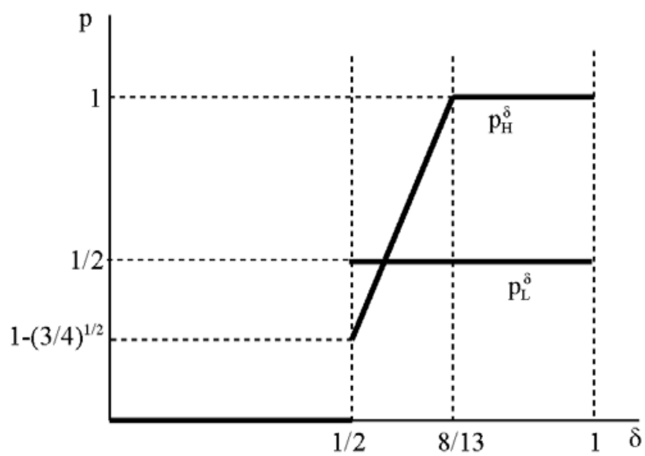
(ii) Comparison of prices in the two states

For $\delta \geq \frac{8}{13}$, $p_H = p_H^m > p_L^m = p_L$. Prices move pro-cyclically (higher in booms).

For $\delta = \frac{1}{2}$, $\pi_H = \pi_L^m$. To achieve the same profits in the high demand state as in the low demand state, prices must be lower in the high demand state. Therefore, prices move counter-cyclically.

To find the best cooperative price in state H for $\delta = \frac{1}{2}$, we set $p_H(2 - p_H) = \pi_L^m = \frac{1}{4}$, which yields $p_H = 1 - \left(\frac{3}{4}\right)^{1/2}$.

Graphically, the most cooperative sustainable prices look like this:



Conclusion: There is a range of discount factors over which the best sustainable price moves countercyclically.

6.3.4 Multimarket contact

Motivation: In certain circles, there is a concern that large, conglomerate enterprises are anticompetitive.

Corwin Edwards: When firms come into contact with each other across many separate markets (geographic or otherwise), opportunistic behavior in any market is likely to be met with retaliation in many markets, and this may blunt the edge of competition.

Is this reasoning correct from the perspective of formal game theory?

(i) It is correct that, with multimarket contact, deviations may lead to more severe punishments involving larger numbers of markets. However,

(ii) Knowing this, if a firm were to deviate from a cooperative agreement, it would deviate in all markets. Consequently, it is not obvious that multimarket contact does anything more than increase the scale of the problem.

It turns out that multimarket contact can facilitate cooperation, but not for the reasons suggested by Edwards.

The central insight:

Notation:

i denotes firm

k denotes market

G_{ik} denotes the net gain to firm i from deviating in market k for the current period, for a particular equilibrium

π_{ik}^{c} denotes the discounted payoff from continuation (next period forward) for firm i in market k, assuming no deviation from the equilibrium in the current period.

π_{ik}^p denotes the discounted "punishment" payoff from continuation (next period forward) for firm i in market k, assuming that i deviates from the equilibrium in the current period.

Equilibrium conditions when markets are separate:

For each i and k,

$$G_{ik} + \delta \pi_{ik}^p \leq \delta \pi_{ik}^c$$

Note that there are $N \times K$ such constraints.

Equilibrium conditions when markets are linked strategically:

For each i,

$$\sum_k G_{ik} + \delta \sum_k \pi_{ik}^p \leq \delta \sum_k \pi_{ik}^c$$

Notice that there are N constraints.

Implication: Multimarket contact pools incentive constraints across markets. This may enlarge the set of outcomes that satisfies the incentive constraints.

For example, the set $\{(x, y) \mid x \leq 4 \text{ and } y \leq 4\}$ is strictly smaller than the set $\{x, y \mid x + y \leq 8\}$.

As it turns out, pooling incentive constraints strictly expands the set of sustainable outcomes, and in particular improves upon the best cooperative outcome, in a number of different circumstances. We will study one of them.

Illustration: Slack enforcement power in one market

Idea: When there is more enforcement power than needed to achieve full cooperation in one market, the extra enforcement power can be used in another market where full cooperation is not achievable. We will consider two examples.

Example #1: Differing numbers of firms in each market.

Suppose firms produce homogeneous goods in each market and compete by naming prices (Bertrand)

Imagine that there are two markets. There are N firms in market 1 and $N+1$ firms in market 2. Moreover,

$$\frac{N-1}{N} < \delta < \frac{N}{N+1}$$

From our analysis of the infinitely repeated Bertrand problem, we know that the monopoly price is sustainable for market 1:

$$\sum_{t=1}^{\infty} \frac{\pi(p_1^m)}{N} \delta^{t-1} > \pi(p_1^m)$$

However, no cooperative price $p_2 > c$ is sustainable for market 2:

$$\sum_{t=1}^{\infty} \frac{\pi(p_2)}{N+1} \delta^{t-1} < \pi(p_2)$$

Thus, if single-market firms operate in both markets, market 1 will be monopolized, while market 2 will be competitive.

Now suppose that N conglomerate firms operate in both markets, and that one single-market firm operates in market 2. Let $1-\alpha$ denote the share of market 2 served by the single-market firm. We will attempt to sustain a cooperative arrangement wherein the N conglomerate firms divide the remaining share (α) equally. The incentive constraint for the single-market firm is:

$$\sum_{t=1}^{\infty} (1-\alpha)\pi(p_2)\delta^{t-1} \geq \pi(p_2)$$

This is equivalent to the requirement that $\alpha \leq \delta$. Thus, the conglomerate firms must cede at least the share $1-\alpha$ to the single-market firm to deter the single-market firm from deviating.

For the conglomerate firms, the incentive constraint becomes

$$\sum_{t=1}^{\infty}\left[\frac{\pi(p_1^m)}{N}+\alpha\frac{\pi(p_2)}{N}\right]\delta^{t-1}\geq\pi(p_1^m)+\pi(p_2)$$

We will attempt to sustain a cooperative arrangement that cedes as little market share to the single-market firm as possible ($\alpha=\delta$). Making this substitution and rearranging, we obtain (after some algebra):

$$\pi(p_2)\leq\pi(p_1^m)\left(\frac{\delta-\frac{N-1}{N}}{\frac{N}{N+1}-\delta}\right)\left(\frac{N}{N+1}\right)$$

Under our assumptions (cooperation is sustainable in market 1 but not in market 2), the RHS of this inequality is strictly positive. Thus, through multimarket contact, one can always sustain $p_2>c$ in market 2 without sacrificing profits in market 1. If δ is sufficiently close to $\frac{N}{N+1}$ (cooperation in market 2 is almost sustainable in isolation), one can achieve monopoly profits in market 2. Note that the conglomerate firms always cede a larger market share to the single-market share to sustain cooperation.

Example #2: Cyclical demand

Consider the model analyzed in the preceding section. Imagine that there are two such markets, and that the same firms operate in both markets. Suppose moreover that the demand shocks in these markets are perfectly negatively correlated (so that there is always a market in state H and a market in stage L). Since the markets are symmetric, this means that there is really only one demand state. Pooling incentive constraints across markets, we obtain:

$$\begin{aligned}&\left(\frac{N-1}{N}\right)[\pi_H(p_H)+\pi_L(p_L)]\\\leq&\left(\frac{\delta}{1-\delta}\right)\left[\frac{1}{2}\pi_H(p_H)+\frac{1}{2}\pi_L(p_L)\right]\frac{2}{N}\end{aligned}$$

After cancellation, one obtains

$$N-1\leq\frac{\delta}{1-\delta}$$

This is equivalent to $\delta \geq \frac{N-1}{N}$, which is exactly the same as for the simple repeated Bertrand model. For example, with $N = 2$, we obtain full cooperation in both states for all $\delta \geq \frac{1}{2}$.

Remark: As the correlation between the demand shocks rises, the gain to multimarket contact declines. When the shocks are perfectly positively correlated, there is no gain. This implies that the potential harm from multimarket contact is greater when the markets are less closely related.

6.3.5 Price wars

Motivation: Price wars appear to occur in practice. However, in the standard model, one only obtains price wars off the equilibrium path. This means they happen with probability zero.

Insight: One can generate price wars on the equilibrium path by considering repeated games in which actions are not perfectly observable. To enforce cooperation, the players must punish outcomes that are correlated with deviations. But sometimes those outcomes occur even without deviations, setting off a price war. In that case, the punishments must be chosen very carefully to assure that the consequences of the occasional war do not outweigh the benefits of cooperation.

Model:

$N = 2$ firms produce a homogeneous good with identical unit costs c

The firms compete by naming prices (Bertrand); indifferent consumers divide equally between the firms.

Firms do not observe each others' price choices, even well after the fact.

Demand in each period is either "high" or "low"

Low demand states occur with probability α. Consumers purchase nothing.

High demand states occur with probability $1 - \alpha$. Consumer purchase $Q(p)$.

Realizations of demand are independent across periods.

The firms cannot directly observe the state of demand, even well after the fact.

A firm only observes its own price and the quantity that it sells.

Let p^m denote the solution to $\max_p (p - c)Q(p) \equiv \pi^m$.

Analysis of equilibrium:

Object: sustain p^m

Problem: if a firm ends up with zero sales, there are two explanations: (i) demand is low, and (ii) its competitor has deviated. It cannot tell the difference.

To sustain p^m, the equilibrium must punish deviations. The only alternative is to enter a punishment phase (price war) any time a firm has zero sales.

Key difference from previous models: the trigger for a price war occurs with strictly positive probability in equilibrium.

Because price wars will actually occur, the firms want the consequences of these wars to be no more severe than absolutely necessary to sustain cooperation. We no longer use *grim strategies* that involve punishing forever.

"Trigger price" strategies:

Charge p^m initially.

As long as firm i has played p^m and made positive sales (or played $p > p^m$) in all previous periods, it continues to play p^m

If, in any period $t-1$, firm i either deviated to $p < p^m$ or made zero sales, the game enters a punishment phase in period t.

In the punishment phase, both firms charge $p = c$ for T periods.

When the punishment phase is over, the strategies reinitialize, treating the first non-punishment period as if it were the first period of the game.

Value functions:

Let V^c denote the expected present value of payoffs from the current period forward when play is not in a punishment phase.

Let V^p denote the expected present value of payoffs from the current period forward in the first period of a punishment phase.

These valuations are related as follows:

$$V^c = (1-\alpha)\left[\frac{\pi^m}{2} + \delta V^c\right] + \alpha\left[0 + \delta V^p\right]$$

$$V^p = \sum_{s=0}^{T-1}(\delta^s \times 0) + \delta^T V^c$$

Substituting in the first expression for V^p using the second expression yields:

$$V^c = (1-\alpha)\left[\frac{\pi^m}{2} + \delta V^c\right] + \alpha\delta^{T+1}V^c$$

Next we solve for V^c:

$$V^c = \frac{(1-\alpha)\frac{\pi^m}{2}}{1 - \delta(1-\alpha) - \alpha\delta^{T+1}}$$

Deviations:

Now we evaluate the profitability of a deviation. Plainly, there is no incentive to deviate during a punishment phase (the prescribed actions constitute the equilibrium for the stage game). We need only check the desirability of deviating outside of punishment phases. The best possible deviation is to slightly undercut p^m. The expected present value of the resulting profits is given by

$$\begin{aligned} V^d &= (1-\alpha)\left(\pi^m + \delta V^p\right) + \alpha\left(0 + \delta V^p\right) \\ &= (1-\alpha)\pi^m + \delta^{T+1} V^c \end{aligned}$$

Deviations are unprofitable as long as $V^c \geq V^d$. This requires

$$V^c \geq (1-\alpha)\pi^m + \delta^{T+1} V^c$$

This is equivalent to:

$$V^c \geq \frac{\pi^m(1-\alpha)}{1-\delta^{T+1}}$$

Now we substitute the expression for V^c derived above. The π^m term cancels – as in the standard repeated Bertrand model, the feasibility of cooperation is all or nothing. Rearranging terms yields the equilibrium condition:

$$2\delta(1-\alpha) + (2\alpha - 1)\delta^{T+1} \geq 1$$

The length of the punishment period:

Note that the equlibrium does not hold for $T = 0$ (the left hand side reduces to δ)

An increase in T reduces the absolute value of the second term on the LHS. This can increase the value of the LHS only if $2\alpha - 1$ is negative. In that case, the value of the LHS remains bounded below $2\delta(1-\alpha)$. Consequently, the equilibrium holds for some $T > 0$ if and only if

(i) $2\alpha - 1 < 0$, and

(ii) $2\delta(1-\alpha) > 1$

Condition (ii) implies condition (i), so we only need to check (ii). When (ii) is satisfied, cooperation is possible. The best cooperative equilibrium involves the least severe punishments consistent with incentive compatibility. This requies us to pick the smallest value of T satisfying the equilibrium condition.

Note that (ii) can be rewritten as

(ii)′ $\delta > \frac{1}{2(1-\alpha)}$

For the special case of $\alpha = 0$, this gives $\delta > \frac{1}{2}$, which is the correct answer for the Bertrand model when there is no observability problem.

Conclusions:

(i) Price wars are observed in equilibrium.

(ii) Price wars are set off by declines in demand. (Note that this contrast with the model of cyclical demand, in which prices fall when demand is high. The key difference concerns observability.)

When an equilibrium price war occurs, everyone knows that no one deviated. It may seem odd to enter a punishment phase under those circumstances. However, if the firms didn't punish this non-deviation, the incentives to comply with the cooperative agreement would vanish.

(iii) Equilibrium price wars are transitory.

(iv) Imperfect observability makes cooperation more difficult (it raises the threshhold value of δ).