Chapter 6: Repeated Continuous Games*

6.1 The Cournot Duopoly

The oldest mathematical analysis of continuous games is due to Augustin Cournot (1833). Cournot considered the problem of two well owners who decide every morning how much water to pump out of their well to sell to the local villagers. The price at which the water will sell during the day depends on the total supply. The more water there is, the lower the price. To formalize the problem let us introduce a "demand function" f that to any total supply q associates the market-clearing price p by the formula $p = f(q)$.[1] Function f must be downward-sloping to reflect the fact that less supply will command a higher price. A very simple example is

$$f(q) = \begin{cases} a - bq & \text{if } 0 \leq q \leq a/b \\ 0 & \text{if } \quad q > a/b \end{cases} \tag{1}$$

with $a > 0$ and $b > 0$. Let us further assume that each well-owner $i \in \{1, 2\}$ has no variable costs of production. i's objective is therefore to choose $q_i \in [0, a/b]$ such that his revenue is maximum. In other words, i wishes to maximize

$$U_i(q_i, q_j) = q_i\big(a - b(q_i + q_j)\big) \tag{2}$$

where q_j denotes the other producer's water supply. One easily obtains the first order condition (as long as $q_i + q_j \leq a/b$)

$$\frac{\partial U_i}{\partial q_i} = a - 2bq_i - bq_j = 0$$

and therefore the revenue maximizing supply for i

$$q_i = \frac{1}{2}\left(\frac{a}{b} - q_j\right) = \varphi_i(q_j) \tag{3}$$

Function φ_i is called i's "reaction function." It describes i's best reply to any anticipated choice by i's counterpart. Of course, the situation is symmetrical so that equilibrium is reached when is is also the case that $q_j = \varphi_j(q_i)$ or when

---

[1]Technically, f is the "inverse demand function" since total demand is a function of price.

$$q_i = q_j = \frac{a}{3b} \qquad (4)$$

At that equilibrium point, each side enjoys revenue $U_i = \frac{a^2}{9b}$. Moreover, one can easily show that the equilibrium is dynamically stable when the two sides behave according to their respective reaction functions (see exercise 1). The point given by (4) is called the Nash-Cournot equilibrium. One can easily generalize the above argument to the asymmetric case with variable costs.

The Nash-Cournot equilibrium appears as a spontaneous outcome of the game when each side repeatedly adjusts its output in reaction to its counterpart's last choice. But what if the two sides join forces, decide on a joint output level and simply divide up the proceeds? To understand the issues, let $q = q_i + q_j$ and let us maximize the joint revenue

$$U(q) = q(a - bq)$$

to obtain the maximizand $q = \frac{a}{2b}$ and maximum joint revenue $\frac{a^2}{4b}$. So, if the two sides agree to each produce $q_i = \frac{a}{4b}$ they will each enjoy revenue $U_i = \frac{a^2}{8b}$ that is strictly greater than $\frac{a^2}{9b}$ $-$ what the Nash-Cournot equilibrium yields. By agreeing to decrease supply from the Nash-Cournot level $q_i = \frac{a}{3b}$ to the collusive level $\frac{a}{4b}$ the two sides can enjoy higher revenues.

But there is a catch: if one side $-$ say $j -$ sticks to the collusive point $q_j = \frac{a}{4b}$ the other side i could revisit its optimization problem and observe that

$$U_i(q_i, \tfrac{a}{4b}) = q_i\left(\tfrac{3a}{4} - bq_i\right)$$

is maximized when $q_i = \frac{3a}{8b}$ for a revenue $U_i = \frac{9a^2}{64b} > \frac{a^2}{8b}$ that is strictly greater than the collusive one. Moreover, the other side sees its revenue go down to $U_j = \frac{3a^2}{32b} < \frac{a^2}{8b}$. So expecting the other to stick to the collusive point, each side is tempted to defect from the arrangement and to exploit the other.

However, if the game is repeated through time and future payoffs are discounted, there is a well known solution to the problem: the Grim trigger. Each side commits to stick to the agreed upon collusive point under the threat that any deviation by either side will trigger a reversion to the Nash-Cournot equilibrium forever. If the two sides have concern for their future revenues, which is undoubtedly a prerequisite to even contemplate the arrangement, and if they discount future revenues by factor ω $(0 < \omega < 1)$ each can figure that sticking to the collusive point will bring an expected stream of revenue

$$E_i^c = \frac{a^2}{8b}(1 + \omega + \omega^2 + ...) = \frac{a^2}{(1-\omega)8b}$$

If, instead, i decides to defect and extract maximum payoff $\frac{9a}{64b}$ on the first turn followed by Nash-Cournot forever after, the revenue stream becomes

$$E_i^d = \frac{9a^2}{64b} + \frac{a^2}{9b}(\omega + \omega^2 + ...) = \frac{9a^2}{64b} + \frac{\omega a^2}{(1-\omega)9b}$$

One can easily verify that $E_i^c > E_i^d$ whenever $\omega > \frac{9}{17}$. In other words, if the two sides have enough concern for tomorrow it is rational to stick to the collusive plan.

But this a rasor-thin construct. The slightest deviation, even in minute amounts from the Nash-Cournot point jeopardizes the plan forever. However, there are more subtle plans that yield a satisfactory result. Consider the case where the two sides agree to collude and to revert to the Nash-Cournot point for only a limited number of turns n in case of defection. The expected payoff of defecting then becomes

$$E_i^l = \frac{9a^2}{64b} + \frac{a^2}{9b}(\omega + \omega^2 + ... + \omega^n) + \frac{a^2}{8b}(\omega^{n+1} + \omega^{n+2} + ...)$$

For instance, two turns $(n = 2)$ of retaliation are enough to ensure the deterrence relationship $E_i^l < E_i^c$, provided $\omega + \omega^2 > \frac{9}{8}$, an inequality that is easily satisfied (for instance with $\omega \geq \frac{3}{4}$). A lower discount factor would require a longer period n of punishment. It is well known that limited time trigger schemes such as this one provide subgame perfect equilibria of the discounted repeated game.

Collusion is usually considered immoral or illegal. A market is supposed to provide a healthy environment for competion that should benefit society as a whole, not just a colluding few of its members. Fortunately, the behavior prescribed by a limited trigger scheme would be sufficiently conspicuous that legal authorities would easily catch the perpetrators and bring them to justice. But what if the scheme is even more subtle, perhaps implementing retaliations in a smooth progressive fashion that would evade detection and still result in some degree of collusion? The question was first raised in the 1960's in the following form: can one design a reaction-function equilibrium akin to the Cournot behavior in that it responds smoothly to deviations, that is subgame perfect, and that results in dynamically stable state of collusion? As we demonstrate in the next sections, the answer is yes[2].

6.2 A General Formulation

We now consider a game $\mathcal{G}$ with N players $i \in \{1, 2, .., \mathrm{N}\}$ for $\mathrm{N} \geq 2$. Each player i has a decision space $\mathcal{A}_i$ that is a compact (closed and bounded) and convex subset of Euclidean space and a continuous payoff function $U_i : \mathcal{A} = \underset{j}{\times} \mathcal{A}_j \to \mathbb{R}$ that is quasi-concave in $x_i \in \mathcal{A}_i$.[3] The Cournot duopoly is a typical example. One has the standard[4]

Theorem 1: The game $\mathcal{G}$ has a Nash equilibrium.

In general, a discounted repeated game, also called a "supergame", is defined from a constituent game such as $\mathcal{G}$ that is repeated at regular intervals of time. The game may start at turn $t = 1$ and repeat potentially forever. A subgame of the repeated game is clearly the same repeated game starting at some turn t. In a sequential rationality approach one will attempt to maximize *at any turn* t the discounted stream of present and future payoffs

$$E_i(X^t, .., X^{t+s}, ..) = \sum_{s=0}^{\infty} \omega_i^s U_i(X^{t+s}) \qquad (5)$$

---

[2]This first proved in J.P. Langlois and J. Sachs, "Existence and Local Stability of Pareto Superior Reaction Function Equilibria in Discounted Supergames", *Journal of Mathematical Economics*, Vol. 22 #3, 1993, pp. 199-222.
[3]$U_i(x_1, .., x_i, .., x_\mathrm{N})$ is quasi-concave in x_i if the set of $\{x_i : \pi_i(x_1, .., x_i, .., x_\mathrm{N}) \geq c\}$ is convex for any value of c. A concave function is quasi-concave but so are functions with bell-shaped curves.
[4]A proof can be found in James W. Friedman, *Oligopoly and the Theory of Games*, 1977.

where $X^{t+s} = (x_1^{t+s}, .., x_j^{t+s}, ..x_{\mathrm{N}}^{t+s}) \in \mathcal{A}$ is the set of players' choices at turn $(t+s)$ and $\omega_i \in (0,1)$ is i's discount factor. A history of play $h^t = \{X^0, X^1, .., X^{t-1}\}$ at time t is the set of all prior developments of the game from its beginning. $\mathcal{H}$ will denote the set of all possible histories of any length. A strategy for a player, say i, in the repeated game is a map $\varphi_i : \mathcal{H} \rightarrow \mathcal{A}_i$ which, to any history at turn t, associates i's decision $x_i^t = \varphi(h^t)$.[5] A strategy profile $\Phi = (\varphi_1, .., \varphi_i, .., \varphi_{\mathrm{N}})$ is a set of one strategy per player. A subgame perfect equilibrium (SPE) is then a strategy profile such that each i maximizes (5) regardless of the prior developments.

In practice one often focuses on so-called "Markov strategies" that react to an interpretation of history in terms of a few rules. For instance, the Grim trigger distinguishes only two kinds of histories: (1) those that have only seen bilateral cooperation; and (2) all others. In the first case it stipulates further cooperation and in the second case defection. Any Markov strategy therefore implictly uses a partition $\mathcal{S} = \{S_k\}$ of $\mathcal{H}$ that can be assumed common to all players.[6] A Markov strategy for i is then a map $\psi_i : \mathcal{S} \rightarrow \mathcal{A}_i$ that to any state associates i's decision $x_i = \psi_i(S_k)$ regardless of the time t at which that state is reached. A Markov strategy profile $\Psi = (\psi_1, .., \psi_i, .., \psi_{\mathrm{N}})$ then defines a response $X = \Psi(S_k)$ for any state S_k and the rules that define the partition $\mathcal{S}$ translate into a map $\mu : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ that defines the new state $S_l = \mu(S_k, X)$. It therefore defines expected payoff streams for each player i at state S_k by

$$E_i(S_k) = U_i(\Psi(S_k)) + \omega_i E_i(\mu(S_k, \Psi(S_k))) \quad (6)$$

A Markov perfect equilibrium (MPE) is a Markov strategy profile such that for any player i at any state S_k the move $x_i = \psi_i(S_k)$ maximizes

$$\gamma_i(x_i, \Psi_{-i}(S_k)) = U_i(x_i, \Psi_{-i}(S_k)) + \omega_i E_i(\mu(S_k, (x_i, \Psi_{-i}(S_k)))) \quad (7)$$

We have

[5]One usually defines a strategy as a complete contingency plan for the entire game. Then, after some history h^t the plan reduces to a so-called "induced strategy" which what we describe here directly.
[6]Otherwise one simply uses the subpartition obtained by intersecting all states in the players' respective partitions.

Theorem 2: Any MPE is a SPE.

Proof (outline): Assume that Ψ is a MPE and that i uses an arbitrary (not necessarily Markov) strategy φ_i instead of ψ_i. Let φ_i^n be the strategy that for any turn t coincides with φ_i for the first n turns and reverts to ψ_i thereafter. Further let X^{t+s} denote play according to (φ_i^n, Ψ_{-i}), and h^{t+s} the corresponding history at $(t+s)$, following h^t. Each h^{t+s} belongs to some state that we can denote S^{t+s}. Since Φ reduces to the MPE Ψ from turn $(t+n)$ on we must have

$$\begin{aligned} E_i(X^{t+n},..,X^{t+n+s},..) &= U_i(X^{t+n-1}) + \omega_i E_i(\mu(S^{t+n}, X^{t+n})) \\ &= \gamma_i(x_i^{t+n}, \Psi_{-i}(S^{t+n})) \leq \gamma_i(\psi_i(S^{t+n}), \Psi_{-i}(S^{t+n})) = E_i(S^{t+n}) \end{aligned}$$

But then, at time $s = n-1$, since $S^{t+n} = \mu(S^{t+n-1}, X^{t+n-1})$:

$$\begin{aligned} E_i(X^{t+n-1},..,X^{t+n+s},..) &= U_i(X^{t+n-1}) + \omega_i E_i(X^{t+n},..,X^{t+n+s},..) \\ &\leq U_i(X^{t+n-1}) + \omega_i E_i(S^{t+n}) = \gamma_i(x_i^{t+n-1}, \Psi_{-i}(S^{t+n-1})) \\ &\leq E_i(S^{t+n-1}) \end{aligned}$$

Continuing backward one obtains

$$E_i(X^t,..,X^{t+s},..) \leq E_i(S^t)$$

Therefore, the switch from ψ_i to φ_i^n does not improve i's payoff stream in response to Ψ_{-1} after any h^t. Moreover, since the argument is independent of n, the deviation to φ_i does not improve i's lot either[7] □

Theorem 4: The repeated play of a Nash equilibrium of $\mathcal{G}$ forms a MPE in the discounted repeated game of constituent game $\mathcal{G}$.

Proof: This is a Markov strategy profile with a single state of the world and the Nash equilibrium play x_i clearly maximizes (7) □

6.3 The Decomposition Theorem

---

[7] One simply writes that $E_i(h^t|\Phi) = \lim_{n\to\infty} E_i(h^t|\Phi^n) \leq E_i(h^t|\Psi)$ where $E_i(h^t|\Phi)$ denotes (5) when play is according to Φ after h^t and where $\Phi^n = (\varphi_i^n, \Psi_{-1})$.

Theorem 2 (Decomposition): A strategy profile $\Phi = (\varphi_1, .., \varphi_i, .., \varphi_{\mathrm{N}})$ forms a SPE of the discounted repeated game if and only if, for every player i, there exists two functions $\pi_i : \mathcal{H} \times \mathcal{A} \to \mathbb{R}_+$ and $g_i : \mathcal{H} \times \mathcal{A}_{-i} \to \mathbb{R}$ such that for any turn t any prior history h^t and any decisions $X^t = (x_1^t, .., x_i^t, .., x_{\mathrm{N}}^t) \in \mathcal{A}$ the following equalities holds

$$g_i(h^t, X_{-i}^t) - \pi_i(h^t, X^t) = U_i(X^t) + \omega_i g_i\big(h^{t+1}, \Phi_{-i}(h^t, X^t)\big) \qquad (8)$$

and $\quad \pi_i(h^t, \Phi(h^t)) = 0$

where $h^{t+1} = \{h^t, X^t\}$ is the history at turn $(t+1)$.[8]

The Decomposition Theorem provides relatively straightforward techniques to construct smooth SPE reaction function equilibria that sustain collusive outcomes. One simple way to do so is to arbitrarily set $\pi_i \equiv 0$ and to limit $g_i(h^t, X_{-i}^t)$ to $g_i(X_{-i}^t)$. In the two-player case (8) further reduces to

$$g_i(x_j^t) = U_i(x_i^t, x_j^t) + \omega_i g_i\big(\varphi_j(x_i^t, x_j^t)\big) \qquad (9)$$

The function g_i, if it is invertible, determines φ_j. To illustrate, we begin with a simple "continuous prisoner's dilemma" defined by

$$U_i(x_i, x_j) = x_i - 2x_j$$

If $x_i \in [0,1]$ is a "level of defection" then $x_i = x_j = 0$ is collusion while the Nash equilibrium of this constituent game $x_i = x_j = 1$ is full defection. To obtain a suitable function g_i one may introduce the following pledge by j: "whenever $x_i = x_j = z$ I will reduce my level of defection z by the factor $0 < \lambda < 1$ so that $\varphi_j(z,z) = \lambda z$. This actually determines entirely the function g_i since

$$g_i(z) = U_i(z,z) + \omega_i g_i(\lambda z) = -z + \omega_i\big(-\lambda z + \omega_i g_i(\lambda^2 z)\big) = ...$$

$$= -z\big(1 + \lambda\omega_i + (\lambda\omega_i)^2 + ...\big) = -\tfrac{z}{1-\lambda\omega_i}$$

It follows that the entire strategy φ_j is the solution of

$$-\tfrac{x_j}{1-\lambda\omega_i} = x_i - 2x_j - \tfrac{\omega_i}{1-\lambda\omega_i}\varphi_j(x_i, x_j)$$

---

[8]A proof of this theorem first appeared in C. Langlois and J.P. Langlois, "Rationality in International Relations: A Game-Theoretic and Empirical Study of the U.S.-China Case," *World Politics*, Vol. 48, #3, April 1996: 358-90.

or $\varphi_j(x_i, x_j) = \frac{1}{\omega_i}\big((1-\lambda\omega_i)x_i + (2\lambda\omega_i - 1)x_j\big)$

If $\frac{1}{2\omega_i} \leq \lambda$, then $\varphi_j(x_i, x_j) \in [0,1]$ and cooperation is dynamically stable under such strategies. For example, $\lambda = \frac{1}{2\omega_i}$ yields $\varphi_j(x_i, x_j) = \frac{x_i}{2\omega_i}$. The duopoly can be discussed similarly (see exercise 2).

Strategies that satisfy (6) have been called "countervailing" since any gain that i could achieve at time t by choice of x_i^t will generate a response φ_j by j that will make i's discounted stream of payoffs equal to $g_i(x_j^t)$, a quantity set by j alone at time t. But then, any choice x_i^t by i is a maximizand, not just the choice $x_i^t = \varphi_j(x_i^{t-1}, x_j^{t-1})$, in response to yesterday's developments. Such an equilibrium can be criticised as not being "strict" and therefore not holding each side to its end of the bargain. In a strict SPE, each side's optimal response to the past must be unique. The criticism can be answered by

Theorem 3: Assume that a countervailing SPE is such that for each i the function $g_i(x_j)$ is differentiable and strictly monotonous. Then that SPE is the uniform limit of a sequence of strict SPE.

Proof: Consider $\pi_i(x_i^t, h^t) = \epsilon|x_i^t - \varphi_j(h^t)|$ with $\epsilon > 0$ and the equation

$$g_i(x_j^t) - \pi_i(x_i^t, h^t) = U_i(x_i^t, x_j^t) + \omega_i g_i\big(\varphi_j(h^{t+1})\big)$$

For ϵ small enough, it has a unique solution φ_j and it is easy to verify that it converges to the countervailing one as $\epsilon \to 0$. Moreover, any choice $x_i^t \neq \varphi_j(h^t)$ yields $\pi_i > 0$ and is therefore not optimal.

Exercises

6.1. Consider an arbitrary sequence of choices (q_i^n, q_j^n) with $n \in \mathbb{N}$ such that

$q_i^{n+1} = \varphi_i(q_j^n)$ and symmetrically for j.

(a) Using formula (3) Verify that $\left|q_i^{n+1} - \frac{a}{3b}\right| = \left|\varphi_i(q_j) - \frac{a}{3b}\right| = \frac{1}{2}\left|q_j^n - \frac{a}{3b}\right|$.

(b) Let $\delta_n = \left|q_i^n - \frac{a}{3b}\right| + \left|q_j^n - \frac{a}{3b}\right|$. Verify that $\delta_{n+1} = \frac{1}{2}\delta_n$ and therefore that $\lim_{n\to\infty} \delta_n = 0$.

(c) Conclude that $\lim_{n\to\infty} (q_i^n, q_j^n) = (\frac{a}{3b}, \frac{a}{3b})$.

6.2. In this exercise you need access to a computer algebra system such as Mathematica or MathLab. Consider the duopoly defined by (2) with $a = 120$ and $b = 1$. For both sides use the discount factor $\omega_i = 0.99$ in the repeated game.

(a) Assume that each side (here j) pledges to play according to $\varphi_j(z, z) = 3.2 + 0.9z$ whenever $q_i = q_j = z$. Verify that the point $q_i = q_j = 32$ is invariant under such pledges.

(b) Verify that equation (6) can be satisfied by a function $g_i(q_j) = \alpha_i + \beta_i q_j + \gamma_i q_j^2$ for appropriate values of α_i, β_i, and γ_i (do not calculate them yet).

(c) Calculate the values α_i, β_i, and γ_i so that the equation

$$g_i(z) = U_i(z, z) + \omega_i g_i(\varphi_j(z, z))$$

holds for all $z \in [0, 120]$. Hint: differentiate twice to obtain γ_i, once to obtain β_i, and finally obtain α_i. You may use a computer algebra system for these steps.

(d) Solve equation (6) for $\varphi_j(q_i, q_j)$ such that $\varphi_j(32, 32) = 32$ (i.e., discard one of the two possible solutions.

(e) Verify that the solution φ_j obtained in (d) maps $\mathcal{S} = [31, 33]^2$ into itself. Hint: Verify that $\frac{\partial \varphi_j}{\partial q_i} > 0$ and $\frac{\partial \varphi_j}{\partial q_i} > 0$ in $\mathcal{S}$ and therefore that the minimum and maximum of φ_j are obtained at $(31, 31)$ and $(33, 33)$ respectively.

(f) Adding the condition that $\varphi_j(q_i, q_j) = 40$ for all $(q_i, q_j) \notin \mathcal{S}$, argue that $\Phi = (\varphi_i, \varphi_j)$ is a SPE (it can also be shown that the point $(32, 32)$ is dynamically stable in $\mathcal{S}$).