LA THEORIE DES JEUX REPETES :

APPLICATION A LA CONCURRENCE OLIGOPOLISTIQUE

PARTIE I

Thierry Pénard

ICI, ENST Bretagne[1]

Avril 1998

Résumé : Les stratégies dans un jeu répété peuvent prendre des formes extrêmement complexes puisqu'elles sont supposées spécifier la conduite des joueurs dans n'importe quelle situation. Les *stratégies simples* développées par Abreu [1986,1988] permettent de lever ces difficultés, en réduisant l'ensemble des règles de décisions d'un jeu à n joueurs à seulement (n+1) règles qui sont indépendantes de l'histoire passée, mis à part l'identité du dernier déviant. Pour chacun des joueurs, il suffit alors de déterminer la pire punition qu'il peut se voir infliger, en termes de gains, pour obtenir une caractérisation de l'ensemble des

---

[1] Ecole Nationale Supérieure des Télécommunications de Bretagne, Département d'Economie, BP 832, 29 285 Brest Cedex
Tel : 02 98 00 14 61, Email : thierry.penard@enst-bretagne.fr

équilibres du jeu répété. L'ensemble de ces n punitions constituent le *code pénal optimal* du jeu.

Ces résultats peuvent être appliqués à la concurrence oligopolistique. On peut montrer, à la suite d'Abreu [1986], que le *code pénal optimal* d'une concurrence en quantité prend une forme extrêmement simple. Ce dernier est construit sur le principe *du bâton et de la carotte*. Il est composé de punitions symétriques qui comportent deux phases : une première phase extrêmement sévère (*le bâton*) qui dure une période et une seconde phase de durée infinie qui conduit les joueurs sur le sentier le plus collusif du jeu (*la carotte*). *Le bâton et la carotte* se révèlent bien supérieurs aux punitions de Cournot-Nash proposées par Friedman [1971]. Si parmi tous les codes pénaux, les firmes privilégient les codes les plus simples mais aussi les plus sévères, on devrait alors observer dans la réalité une utilisation courante de stratégies de représailles ayant la forme de *la carotte et du bâton*.

PARTIE I

LES STRATEGIES DE COOPERATION DANS UN JEU REPETE

Introduction

La coopération est l'enjeu de la plupart des relations économiques, politiques ou sociales. Les firmes ont intérêt à s'entendre sur des prix élevés, les Etats ont intérêt à faire la paix et à développer des échanges commerciaux, l'employeur et le salarié ont intérêt à entretenir une relation de confiance. Mais, les intérêts privés peuvent faire obstacle à la coopération.

La théorie des jeux répétés offre un cadre d'analyse des phénomènes de coopération en dehors de tout contrats. Par exemple, Greif, Milgrom et Weingast [1994] expliquent par la théorie des jeux répétés l'existence de guildes de marchands au moyen âge. Ces guildes avaient pour rôle de contraindre les autorités d'une ville à assurer la sécurité des marchands, en les menaçant d'un embargo et favorisaient ainsi l'essor du commerce.

Le meilleur domaine application de la théorie des jeux répétés est sans aucun doute la concurrence oligopolistique. Les jeux répétés sont très utiles pour modéliser et évaluer les comportements stratégiques qui facilitent la coopération et la collusion entre firmes[2].

[2] On parle de collusion lorsqu'il s'agit d'une coopération illégale : exemple, une entente sur les prix ou sur les parts de marché.

Dans la section 2, nous définissons les notions de base en théorie des jeux. Dans la section 3, nous analysons la coopération dans les jeux infiniment répétés, à l'aide des outils développés par Abreu dans sa thèse puis dans les articles de 1986 et 1988. Ces outils sont séduisants à plus d'un titre. Ils simplifient la représentation stratégique d'un jeu répété et permettent de caractériser l'ensemble des équilibres parfaits, en particulier le meilleur équilibre en termes de gains pour les joueurs.

Section 2. Notions générales en théorie des jeux

Un jeu répété est un jeu ordinaire réitéré plusieurs fois de suite. Par jeu ordinaire, nous entendons un jeu statique (joué une seule fois) dans lequel les joueurs choisissent simultanément leurs stratégies.

Historiquement, la théorie des jeux est une théorie des jeux statiques. Il a fallu attendre les années soixante-dix pour penser les jeux à plusieurs périodes comme des jeux à part entière. Il n'est pas abusif de parler aujourd'hui d'une théorie des jeux répétés tant il est vrai que cette dernière a acquis une véritable autonomie conceptuelle par rapport à la théorie des jeux statiques.

Avant d'entrer dans la théorie des jeux répétés, il nous semble nécessaire d'apporter quelques précisions lexicales et de dresser une typologie des jeux répétés. Un **jeu répété** (en anglais « *repeated game* ») au sens strict est un *jeu répété stationnaire* : c'est le même jeu ordinaire, appelé *jeu constituant*, qui est répété de période en période. Les conditions de jeu ne se modifient pas au cours du temps : même nombre de joueurs, mêmes ensembles de stratégies, mêmes fonctions de gain, même facteur d'actualisation[3].

Dans un premier temps, nous caractérisons le jeu constituant. Ce jeu constituant permet de construire le jeu répété.

[3] Si l'une de ces conditions évolue au cours du temps, nous parlons de superjeu (« *supergame* »). Un **superjeu** est donc une séquence de jeux ordinaires qui peuvent différer d'une période à l'autre.

2.1 Le jeu constituant :

Nous considérons un jeu non-coopératif dans lequel les joueurs doivent sélectionner simultanément une action[4]. Chaque joueur cherche à maximiser son utilité ou ses gains individuels en tenant compte des choix attendus des autres joueurs. Nous plaçons quelques restrictions sur les caractéristiques de ce jeu.

Hypothèse 1 : Le nombre de joueurs est fini et égal à n. Les joueurs sont indicés de 1 à n. L'ensemble des joueurs est noté $N=\{1, 2,.., i,..., n\}$.

Hypothèse 2 : L'ensemble de stratégies S_i du joueur i est compact et convexe.

Précisons qu'un ensemble de dimension finie est compact lorsqu'il est borné et fermé[5]. Il est équivalent de supposer que les joueurs choisissent des stratégies pures dans un ensemble compact de stratégies ou qu'ils choisissent des distributions de probabilités sur un ensemble fini de stratégies pures (stratégies mixtes). L'ensemble des distributions de probabilités est alors compact. Dans tout ce chapitre, nous considérons que le jeu est en stratégies pures et satisfait les hypothèses 1 et 2. L'ensemble de stratégies des n joueurs est noté S et est défini comme le produit cartésien des n ensembles individuels de stratégies ($S = \mathrm{X}_{\{i \in N\}} S_i$). Cet ensemble est aussi compact et convexe. Dans le jeu à n joueurs, un vecteur de stratégies s'écrit $s=(s_1,...,s_i,...,s_n) \in S$. La $i^{ème}$ composante du vecteur s correspond à la stratégie du joueur i. Notons $s_{-i} =(s_1,..., s_{i-1}, s_{i+1,...,} s_n)$ le vecteur de stratégies des (n-1) joueurs différents de i et $S_{-i} = \mathrm{X}_{\{j \neq i\}} S_j$ l'ensemble des stratégies de ces (n-1) joueurs. $(t_i ; s_{-i})=(s_1,..., s_{i-1}, t_{i,} s_{i+1,...,} s_n)$ est le vecteur de stratégies dans lequel le joueur i choisit t_i et les joueurs $j \neq i$ choisissent s_j. Le vecteur de stratégies s peut alors s'écrire $(s_i; s_{-i})$. Les deux hypothèses suivantes concernent les fonctions de gain des joueurs.

---

[4] Nous pouvons très bien envisager le même jeu ordinaire ou statique dans lequel les joueurs choisissent simultanément plusieurs actions. Si les joueurs ont k choix à réaliser, leur espace de stratégies est de dimension k.

[5] Un ensemble A de dimension finie k est borné s'il existe un nombre positif K tel que pour tout vecteur $x=(x_1,...,x_j,...,x_k) \in A$, nous avons $\Sigma_j |x_j| \leq K$. Il est fermé si et seulement si pour toute suite

Hypothèse 3: la fonction de gain du joueur i, $g_i(s) \in \mathbf{R}$, est définie, continue et bornée pour tout $s \in S$ et pour tout $i \in N$.

Notons que $g_i(.)$ est une application de S vers **R**.

Hypothèse 4 : $g_i(t_i\,;s_{-i})$ est quasi-concave en t_i pour tout $s_{-i} \in S_{-i}$ et pour tout $i \in N$.

Le vecteur des fonctions individuelles de gains s'écrit $g=(g_1,...,g_i,...,g_n) \in \mathrm{R}^N$. Le jeu constituant caractérisé par l'ensemble des joueurs *N*, l'ensemble de stratégies *S* et le vecteur de gains *g* et qui satisfait les hypothèses 1 à 4 est alors noté $\Gamma=(N, S, g)$. Dans la théorie des jeux répétés, $\Gamma=(N, S, g)$ est le **jeu constituant** qui est rejoué à chaque période. Il constitue la base ou le « squelette » du jeu répété.

Nous supposons que le jeu Γ=(N, S, g) est en information complète. Chaque joueur connaît l'ensemble de stratégies et les fonctions de gains des autres joueurs. De plus, cette information complète sur N, S et g est une connaissance commune[6] : chaque joueur sait que les autres joueurs connaissent N, S et g, de même chaque joueur sait que les autres joueurs savent qu'il connaît N, S et g, et ainsi de suite...

Le point central de cette section réside dans la définition du concept d'équilibre non-coopératif, appelé communément équilibre de Nash[7].

2.2. L'équilibre de Nash :

---

convergente $\{t^l\} \in A$, alors $\lim_{l \to \infty} t^l \in A$. Enfin, il est convexe si pour tout x et $x' \in A$ et $\lambda \in [0,1]$, $\lambda x + (1-\lambda)x' \in A$.

[6] D'autre part, on suppose que chaque agent est parfaitement rationnel, qu'il a la certitude que les joueurs sont aussi rationnels et que cette certitude est connaissance commune. Enfin, les capacités de calcul des joueurs ne sont pas limitées et les informations nécessaires aux prises de décisions sont collectées et traitées sans coût. Ces hypothèses sur la rationalité des joueurs sont très fortes et font l'objet de nombreuses critiques (Friedman [1992]).

[7] Lorsque l'on parle d'un équilibre sans autre qualificatif, il s'agit toujours d'un équilibre de Nash. Nous verrons par la suite qu'il existe d'autres concepts d'équilibre qui permettent de raffiner l'équilibre de Nash (équilibre parfait, Pareto-Optimal, robuste à la renégociation...).

Un équilibre de Nash est un état dans lequel aucun joueur ne souhaite modifier sa stratégie étant donné les stratégies adoptées par les autres joueurs. Chaque stratégie est une meilleure réponse aux stratégies des autres joueurs.

Définition 1 : Un vecteur de stratégies s^* est un équilibre de Nash du jeu $\Gamma=(N, S, g)$ si et seulement si pour tout $i \in N$ et $t_i \in S_i$, on a :

$$g_i(s^*) \geq g_i(t_i\ ;\ s^*_{-i}) \qquad (1)$$

Cette définition est simple à comprendre. Considérons l'expression de gauche de l'inégalité (1). $g_i(s^*)=g_i(s^*_i\ ; s^*_{-i})$ est le gain du joueur i quand il choisit s^*_i étant donné que les autres se conforment à s^*_{-i}. Le terme de droite $g_i(t_i\ ;\ s^*_{-i})$ correspond au gain du joueur i quand il dévie de s^*_i et sélectionne une stratégie t_i, alors que les n-1 autres joueurs se conforment à s^*_{-i}. Les conditions d'équilibre nous disent que le joueur i ne peut tirer aucun bénéfice d'une déviation de s^*_i.

Il existe de nombreux théorèmes relatifs à l’existence d’un équilibre de Nash. Le théorème de Nash [1951] dit que tout jeu ayant un nombre fini de stratégies pures admet un équilibre en stratégies mixtes. Debreu [1952] et Glicksberg [1952] donnent les conditions d’existence d’un équilibre de Nash en stratégies pures. Ce théorème d’existence a été étendu aux jeux ayant un nombre infini (ou un continuum) de stratégies pures ou ayant des fonctions de gain discontinues (Dasgupta et Maskin [1986]).

Sous les hypothèses 1 à 4 énoncées précédemment (continuité et quasi-concavité des fonctions de gains d’une part, convexité et compacité des ensembles de stratégies d’autre part), l’existence d’un équilibre de Nash dans le jeu constituant $\Gamma=(N, S, g)$ est assurée[8]. Le théorème d’existence d’un équilibre de Nash est le résultat central de la théorie des jeux statiques ; dans la théorie des jeux répétés, l’équivalent de ce théorème est le Folk Théorème qui, sous les mêmes hypothèses, garantit l’existence d’une multitude d’équilibres de Nash.

[8] La démonstration est omise.

2.3 Les stratégies dans un jeu répété :

Nous considérons le jeu répété dans lequel les joueurs font face à chaque période au jeu constituant défini précédemment. La répétition de ce jeu permet aux joueurs de conditionner leurs choix présents et futurs sur les choix passés. Cette interdépendance temporelle peut conduire à des solutions plus coopératives ou plus agressives que celles observées dans le jeu constituant. La répétition élargit l'ensemble des solutions.

Nous définissons $s_i(t)$ comme l'action du joueur i à la date t. $s(t) = (s_1(t),\ldots,s_i(t),\ldots,s_n(t)) \in S$ est alors le profil d'actions sélectionnés par les n joueurs à la date t et $h_t = (s(0), s(1),\ldots, s(t-1))$ l'histoire du jeu à cette même date. L'histoire du jeu correspond à l'ensemble des actions que les joueurs ont choisi entre la période initiale 0 et la période t-1. Nous avons $h_t \in S^t$ où $S^t = \mathrm{X}_{(t\ fois)} S$.

Dans un jeu répété, le profil d'actions sélectionné par les joueurs à chaque période crée un nouveau sous-jeu qui est entièrement défini par l'histoire à cette date.

Définition 2 : Dans un jeu répété, une stratégie pour le joueur i consiste en une séquence de règles de décision, une par période. Cette stratégie est notée $\sigma_i=(\sigma_{i(0)}, \sigma_{i(1)}, \sigma_{i(2)},\ldots, \sigma_{i(t)},\ldots)$ où $\sigma_{i(t)}$ est la règle de décision du joueur i à la période t.

La règle de décisions $\sigma_{i(t)}$ est une application de S^t vers S_i qui spécifie pour chaque histoire possible du jeu h_t une action ou une conduite à tenir à la date t. Si $\Sigma_{i(t)}$ est l'ensemble des règles de décision du joueur i à la date t, l'ensemble des stratégies du joueur i dans le jeu répété est alors[9] $\Sigma_i = S_i \mathrm{x} \Sigma_{i(1)} \mathrm{x} \Sigma_{i(2)} \mathrm{x}\ldots$ Nous notons $\sigma=(\sigma_1,\ldots,\sigma_i,\ldots,\sigma_n)$ le profil de stratégies du jeu répété et $\sigma_{(t)}=(\sigma_{1(t)},\ldots,\sigma_{i(t)},\ldots,\sigma_{n(t)})$ le profil des règles de décisions à la période t. On a $\sigma\in\Sigma$ où Σ est le produit cartésien des ensembles de stratégies noté ($\Sigma = \mathrm{X}_{i=1}^{n} \Sigma_i$).

A ce stade, nous pouvons distinguer plusieurs classes de stratégies selon la place que l'histoire occupe dans les règles de décisions des joueurs.

Les stratégies en boucle ouverte. Le joueur ne tient pas compte de l'histoire du jeu. En fait cela revient à dire que le joueur choisit initialement une séquence d'actions (une par période) et qu'il applique ce programme quel que soit le comportement de ses rivaux[10]. Le comportement d'un tel joueur peut être qualifié de myope. Toute la dimension dynamique du jeu répété est éliminée Une stratégie en boucle ouverte s'écrit $\sigma_i = \{(s_{i(t)})_{t=0,1,...}\}$ ou encore $\sigma_i=(\sigma_{i(0)}, \sigma_{i(1)},..., \sigma_{i(t),....})$ avec $\sigma_{i(t)}[h_t]=\sigma_{i(t)}[h'_t]$ pour toutes les histoires $h_t \neq h'_t \in S^t$.

Les stratégies en feed-back. Le joueur conditionne son choix présent sur les seules actions de la période précédente et non sur toute l'histoire du jeu. Ce type de stratégie revient à supposer une mémoire imparfaite des joueurs. Nous pouvons aussi concevoir des stratégies qui se fonderaient sur les seules actions des deux ou trois périodes précédentes.

Les stratégies en boucle fermée[11]. Les joueurs tiennent compte de l'histoire du jeu. Initialement, chaque joueur adopte une séquence de règles de décision, une règle par période. A la période t, il observe l'histoire du jeu h_t et joue l'action prescrite par sa règle de décision $\sigma_{i(t)}$. Nous pouvons alors avoir $\sigma_{i(t)}[h_t]\neq\sigma_{i(t)}[h'_t]$ pour deux histoires différentes h_t, $h'_t \in S^t$. Les stratégies en boucle fermée sont une représentation générale de l'ensemble des stratégies d'un jeu répété et notamment des stratégies en boucle ouverte ou en feed-back.

Dans un jeu répété, les règles de décision peuvent prendre des formes extrêmement complexes, puisqu'elles sont supposées spécifier la conduite à tenir dans n'importe quel sous-jeu et face à n'importe quelle histoire[12]. Le mérite d'Abreu [1986,1988] est de proposer une représentation des stratégies sous une forme *simple*. L'ensemble des stratégies d'un jeu répété comportant n joueurs est, en effet, réduit à seulement (n+1) règles d'actions qui sont elles

[9] Notons que $\Sigma_{i(0)} = S_i$ car faute d'histoire à la période t=0, le joueur s'engage de manière inconditionnelle sur une action $s_{i(0)} \in S_i$.

[10] On peut démontrer facilement qu'un équilibre en boucle ouverte consiste nécessairement en une séquence d'équilibres de Nash du jeu constituant.

[11] Cette terminologie se rencontre aussi dans les jeux dynamiques qui dépendent d'un état (state-dependant). Une stratégie en boucle fermée tient compte non seulement de l'état actuel mais aussi de l'évolution de cet état dans le passé. En revanche, une stratégie en boucle ouverte ne tient jamais compte de l'état.

mêmes définies par (n+1) sentiers d'actions. La détermination des équilibres est alors plus facile.

Section 3. Stratégies simples et codes pénaux

3.1 Les sentiers d'actions :

Dans un jeu répété, les joueurs choisissent à chaque période leurs actions selon des règles de décisions définies préalablement. La séquence des profils d'actions sélectionnés au cours du jeu détermine un *sentier unique* dans le jeu complet. Ce sentier est une succession de déplacement des joueurs d'un point de décision à un autre point de décision et d'un sous-jeu à un autre sous-jeu.

De manière formelle, un sentier est une séquence infinie de profils d'actions $\{s(t)\}_{t=0}^{\infty}$ où $s(t)$ est le profil d'actions sélectionné par les joueurs à la période t.

Exemple de sentier : Soit $s^*=(s^*_1,..,s^*_n)$ l'équilibre de Nash du jeu constituant $\Gamma(S,g,N)$. $\{s^*, s^*, s^*,...\}$ où $\{s(t)=s^*\ \forall\ t\}$ est un sentier stationnaire dans lequel les n joueurs répètent à chaque période l'équilibre de Nash du jeu constituant.

Un jeu infiniment répété admet une infinité de sentiers. Même dans un jeu répété fini, le nombre de sentiers possibles devient rapidement élevé. Par exemple, dans le jeu du dilemme du prisonnier répété dix fois, il existe plus d'un million de sentiers possibles[13].

Définition 3 : Soit $\sigma = (\sigma(0), \sigma(1), \ldots, \sigma(t), \ldots) \in \Sigma$ un profil de stratégies où $\sigma(t) = (\sigma_1(t), \ldots, \sigma_i(t), \ldots, \sigma_n(t))$ sont les règles de décision des n joueurs à la date t. Le profil de stratégies σ génère un sentier noté $Q[\sigma] \in S^{\infty}$ défini par $Q[\sigma] = \{q[\sigma](t)\}_{t=0}^{\infty}$ tel que :

---

[12] De même, la détermination des équilibres d'un jeu répété n'est pas une tâche facile, surtout si l'on exige la perfection sur tous les sous-jeux.
[13] Il existe exactement 4^{10} sentiers possibles dans ce jeu à deux joueurs qui disposent chacun de deux actions possibles.

$$q[\sigma](0) = s(0) = (s_1(0), \ldots, s_i(0), \ldots, s_n(0))$$

$$q[\sigma](t) = \sigma(t)\big(q[\sigma](0), q[\sigma](1), \ldots, q[\sigma](t-1)\big) \qquad \text{pour tout } t \geq 1.$$

Un profil de stratégies spécifie un profil d'actions initial $s(0) \in S$ que les joueurs choisissent à la période t=0. Puis à la période 1, ils choisissent un profil d'actions $q[\sigma](1)$ conformément à leurs règles de décision $\sigma(1)$ et en se basant sur l'histoire du jeu (c'est à dire sur les choix observés à la date t=0). A la période t, l'histoire du jeu est $h_t = \big(q[\sigma](0), q[\sigma](1), \ldots, q[\sigma](t-1)\big)$ et $q[\sigma](t)$ est le profil d'actions spécifiées par les règles $\sigma(t)$.

La valeur présente des gains que génère le sentier $Q[\sigma]$ est égale pour le joueur i à :

$$G_i(\sigma) = \sum_{t=0}^{\infty} \delta^t g_i(q[\sigma](t))$$

avec δ le facteur d'actualisation commun à tous les joueurs.

Nous pouvons aussi noter cette valeur $G_i(Q[\sigma])$. G_i est la fonction de gains du jeu répété avec actualisation. Le jeu constituant ayant été noté $\Gamma=(N,S,g)$, le jeu infiniment répété avec actualisation est défini par $\Gamma^{\infty}=(N,\Sigma,G,\delta)$ où Σ est l'ensemble de stratégies et $G=(G_1,\ldots,G_i,\ldots,G_n)$ le vecteur de gains.

Définition 4 : Dans le jeu $\Gamma^{\infty}=(N,\Sigma,G,\delta)$, le profil de stratégies σ est un équilibre de Nash si et seulement si pour tout $i \in N$ et tout $\sigma'_i \in \Sigma_i$ on a :

$$G_i(\sigma) \geq G_i(\sigma'_i, \sigma_{-i})$$

Dans un jeu répété, l'équilibre de Nash est un concept insatisfaisant parce qu'il n'est pas assez exigeant sur les profils d'actions prescrits dans les sous-jeux hors équilibre. C'est pourquoi les équilibres de Nash d'un jeu dynamique sont toujours soumis au critère de sous-jeu parfait.

3.2 Crédibilité et équilibre de sous-jeu parfait :

La notion d'équilibre de sous-jeu parfait a été définie par Selten [1975]. Elle permet d'éliminer les équilibres de Nash qui spécifient sur certains sous-jeux hors équilibre des actions non crédibles. Un profil d'actions est crédible si chaque joueur a intérêt individuellement à s'y conformer étant donné que les autres joueurs font de même. La crédibilité dans un jeu répété est étroitement liée aux punitions que les joueurs sont censés mettre en oeuvre lorsque l'un d'eux dévie du sentier d'équilibre. Ces punitions sont crédibles s'il est bien dans l'intérêt des joueurs de les appliquer lorsqu'une déviation est observée.

De manière formelle, un profil de stratégies σ est un équilibre parfait si c'est un équilibre de Nash sur le sentier d'équilibre, mais aussi sur tous les sentiers hors équilibre ou sur tous les sous-jeux possibles. Un sous-jeu étant entièrement défini (ou localisé) par une date t et une histoire h_t, nous appelons $\sigma^{h_t} = (\sigma_i^{h_t};\sigma_{-i}^{h_t})$ le profil de stratégies induit par σ sur le sous-jeu qui commence en t après l'histoire h_t et $\Sigma_i^{h_t}$ l'ensemble des stratégies du joueur i dans ce même sous-jeu.

Définition 5 : le profil de stratégies σ est un équilibre parfait dans le jeu répété Γ (N,Σ,G,δ) si et seulement si :

i) σ est un équilibre de Nash sur l'ensemble du jeu Γ :

$G_i(\sigma) \geq G_i(\sigma'_i, \sigma_{-i})$ pour tout $i \in N$ et tout $\sigma'_i \in \Sigma_i$

ii) σ^{h_t} est un équilibre de Nash pour tout t et toute histoire h_t :

$G_i(\sigma^{h_t}) \geq G_i(\sigma'_i;\sigma_{-i}^{h_t})$ pour tout $i \in N$, tout t=1,2,..., h_t et $\sigma'_i \in \Sigma_i^{h_t}$.

Un profil de stratégies d'équilibre parfait du jeu Γ peut être pensé en termes de sentiers et de règles spécifiant un sentier initial $Q°$, un sentier Q^1 en cas de déviation du sentier Q^0, un sentier Q^2 en cas de déviation du sentier Q^1, ..., un sentier Q^l en cas de déviation du $l^{ième}$ sentier en cours Q^{l-1}. On mesure la complexité que peut prendre un tel profil de stratégies si le sentier prescrit après la $l^{ième}$ déviation dépend de l'histoire entière du jeu, de la date de cette déviation, de son ampleur ou de l'identité du coupable. La recherche d'un équilibre parfait nécessite de prendre en compte toutes les histoires possibles et toutes les séquences de déviations finies ou infinies. Cette tâche dépasse de loin les capacités de calcul du théoricien des jeux et encore plus celles des joueurs en présence.

Le mérite d'Abreu est de proposer des profils de *stratégies simples* qui facilitent la détermination des équilibres parfaits et qui fournissent une représentation assez juste des comportements stratégiques dans un jeu répété.

3.3 Les profils de stratégies simples :

L'idée d'Abreu est de restreindre le nombre de sentiers possibles du jeu à seulement $n+1$ sentiers. Les joueurs se déplacent sur ces sentiers selon les règles explicitées dans la définition suivante.

Définition 6 : Un profil de stratégies simples (*PSS*) est un ensemble de règles de décisions entièrement définies par $n+1$ sentiers (Q^0, Q^1,...,Q^n) et qui spécifient :

1-) de commencer sur le sentier Q^0 à la date t=0.

2) de suivre Q^0 aux dates suivantes si aucun joueur n'a dévié unilatéralement de ce sentier.

3-) de passer sur le sentier Q^j dès que le joueur j dévie unilatéralement du sentier en cours Q^i (i=0,1,..,n).

4-) de rester sur le sentier Q^i si Q^i est le sentier actuellement prescrit et si aucune déviation n'est observée ou si 2 joueurs ou plus ont dévié simultanément[14] de Q^i.

Q^0 est le sentier d'équilibre que les joueurs veulent soutenir initialement et Q^i est le sentier de punition qui s'impose à tous les joueurs lorsque le joueur i a dévié. Les règles de décision d'un *PSS* sont indépendantes de l'histoire passée, mis à part l'identité du dernier déviant. Elles sont fondées simplement sur la comparaison entre les actions prescrites sur le sentier en cours et les actions réellement sélectionnées. Si à la période précédente tous les joueurs se sont conformés aux actions prescrites, les joueurs restent sur le sentier en cours. Si un des joueurs a

[14] Il existe une variante dans laquelle si plusieurs joueurs dévient simultanément, c'est celui qui a le plus petit indice qui est puni. Dans ce cas, le classement des individus de 1 à n n'est plus neutre. Il peut être effectué selon la réputation d'honnêteté, la confiance que l'on accorde à chaque joueur. Ce classement peut remettre en cause la perfection de l'équilibre si certains joueurs dévient en comptant que d'autres classés devant eux feront de même et seront punis à leur place. Une nouvelle dimension stratégique est alors introduite dans laquelle le choix des règles et des punitions n'est plus indifférent.

dévié du sentier en cours, la règle prescrit de démarrer (ou de redémarrer) la punition à l'encontre de ce joueur.

Les règles de décisions proposées par Abreu sont symétriques ($\sigma_{i(t)=}\ \sigma_{j(t)}$ pour tout i≠j) et stationnaires à partir de la date t=1 ($\sigma_{i(t)=}\ \sigma_{i(1)}$ pour tout t). L'ensemble d'information pertinent pour mettre en oeuvre ces règles se limite au sentier en cours et au profil d'actions observé à la période précédente.

De manière formelle, si le sentier Q^i est défini par $Q^i=\{q^i(\tau)\}_{\tau=0}^{\infty}$ pour tout i=0,1,...,n, alors un PSS $\sigma=\{\sigma_{(0)}, \sigma_{(1)}, \sigma_{(2)},..., \sigma_{(t)},...\}$ caractérisé par les n+1 sentiers ($Q^0, Q^1,...,Q^n$) satisfait :

i) en t=0, $\sigma_{(0)}=q^0(0)$.

ii) en t ≥1, si $\sigma_{(t-1)}[h_{t-1}]=q^i(\tau\text{-}1)$ et si $s(t\text{-}1)=q^i(\tau\text{-}1)$ pour i=0,1,...,n, et τ≤t, alors $\sigma_{(t)}[h_{t-1}, q^i(\tau\text{-}1)]= q^i(\tau)$ (règle de décision en l'absence de déviation à la période précédente).

iii) en t ≥1, si $\sigma_{(t-1)}[h_{t-1}]=q^i(\tau\text{-}1)$ et si $\exists\ j, k \in N, j\neq k$ tels que $s_j(t\text{-}1)\neq q^i_j(\tau-1)$ et $s_k(t\text{-}1)\neq q^i_k(\tau-1)$ alors $\sigma_{(t)}[h_{t-1}, s(t\text{-}1)]=q^i(\tau)$ (règle de décision en cas de déviations multiples à la période précédente).

iv) si $\exists$ i tel que $s_i(t\text{-}1)\neq \sigma_{(t-1)}[h_{t-1}]$ et $s_j(t\text{-}1)= \sigma_{(t-1)}[h_{t-1}]$ pour tout j≠i,
alors $\sigma_{(t)}[h_{t-1}, s(t)]= q^i(0)$ (règle de décision en cas de déviation unilatérale à la période précédente).

En ii), il est dit que si à la date t-1, les joueurs sont depuis (τ-1) périodes sur le sentier Q^i (ils ont donc déjà joué $q(0), q(1),..., q(\tau\text{-}2)$) et que chaque joueur se conforme à la $\tau^{ième}$ action de ce sentier, c'est à dire $q(\tau$-1), la règle prescrit alors de jouer à la date t le profil d'actions $q(\tau)$. La même règle prévaut si au moins deux joueurs ont dévié du sentier Q^i à la date t-1. Le PSS défini par les n+1 sentiers ($Q^0, Q^1,...,Q^n$) est noté $\sigma(Q^0, Q^1,...,Q^n)$.

Avec des stratégies simples, le critère de perfection est facile à vérifier. Etant donné le sentier en cours, l'ensemble des histoires à la date t se réduit à ($n+1$) *histoires caractéristiques*. La première histoire correspond à la situation dans laquelle personne n'a dévié ou plus de deux joueurs ont dévié du sentier en cours à la date précédente. La deuxième histoire correspond à une déviation du joueur 1 à la date précédente,..., la $(n+1)^{\text{ième}}$ histoire à une déviation du joueur n à la date précédente.

Le PSS $\sigma(Q^0, Q^1,...,Q^n)$ défini par $(n+1)$ sentiers et $(n+1)$ histoires caractéristiques permet de dériver n autres profils de stratégies simples. Pour tout $i \in N$, soit $\sigma^i(Q^1,...,Q^n)=\sigma(Q^i, Q^1,...,Q^n)$ le profil de stratégies simples induit par une déviation du joueur i à la période précédente. Il est identique au PSS $\sigma(Q^0, Q^1,...,Q^n)$ sauf dans le sentier qu'il prescrit initialement (Q^i au lieu de Q^0). Un PSS $\sigma(Q^0, Q^1,...,Q^n)$ est alors parfait si ce profil et les n profils induits sont des équilibres de Nash. Nous devons donc vérifier que pour tout i et tout $\sigma^i(Q^1,...,Q^n)$, aucun joueur ne peut obtenir un meilleur gain en déviant de sa stratégie. Cette tâche est facilitée par un résultat central en théorie des jeux répétés selon lequel si aucune déviation d'une période n'est profitable alors aucune séquence finie ou infinie de déviations n'est profitable[15].

3.4. Le principe de déviation unique :

Considérons le gain et le coût d'une déviation d'une période. Une telle déviation consiste pour un joueur à ne pas respecter le sentier en cours puis à se conformer à sa punition. Si le sentier en cours est Q^j (j=0,1..,n) et ce depuis τ périodes, le joueur i obtient un gain actualisé égal à $G_i(Q^j;\tau)=\sum_{s=0}^{\infty}\delta^s g_i\left(q^j(\tau+s)\right)$ en ne déviant pas.

En revanche, s'il dévie pendant une période du profil d'actions $q^j(\tau)$, puis se conforme dès la période suivante à sa stratégie d'équilibre son gain actualisé est au plus égal à $\max_{\{t_i\}} g_i\left(t_i;q_{-i}^j(\tau)\right)+\delta G_i(Q^i)$. En effet après la déviation, le jeu s'engage sur le sentier Q^i, destiné à punir le joueur i. Cette punition a une valeur présente de[16] $G_i(Q_i)=\sum_{s=0}^{\infty}\delta^s g_i\left(q^i(s)\right)$ pour le joueur i. Posons $g_i^d(q)=\max_{\{t_i\}} g_i(t_i;q_{-i})$ le gain maximum que peut obtenir le joueur i lorsque les autres joueurs choisissent q_{-i}. Etant donné le sentier Q^j en cours depuis τ périodes, aucune déviation d'une période n'est profitable si pour tout i on a :

$$G_i(Q^j;\tau)\geq g_i^d(q^j(\tau))+\delta G_i(Q^i)$$

Après réorganisation, cette inégalité peut s'écrire sous la forme suivante :

[15] En anglais, c'est le « one-stage deviation principle » que l'on peut traduire par le principe de déviation instantanée ou de déviation unique.

[16] Par abus de notation, nous notons $G_i(Q^i;\tau=0)=G_i(Q^i)$ qui est le gain actualisé du sentier complet Q^i.

$$g_i^d(q^j(\tau)) - g_i(q^j(\tau)) \leq \delta\left[G_i(Q^j;\tau+1) - G_i(Q^i)\right] \qquad (2)$$

avec $G_i(Q^j;\tau+1) = \sum_{s=0}^{\infty}\delta^s g_i\left(q^j(\tau+s+1)\right)$

Le terme de gauche dans l'inégalité (2) est le gain net d'une déviation d'une période et le terme de droite le coût d'opportunité de cette déviation. Si ce coût est supérieur au gain attendu de la déviation pour tous les joueurs et pour tous les sentiers, alors aucune déviation d'une période n'est profitable.

Proposition 1 : *Le profil de stratégies simples $\sigma(Q^0, Q^1,...,Q^N)$ est un équilibre de sous-jeu parfait si et seulement si pour tout $i \in N$, $j=0,1,..,n$ et $\tau=0,1,...$:*

$$g_i^d(q^j(\tau)) - g_i(q^j(\tau)) \leq \delta\left[G_i(Q^j;\tau+1) - G_i(Q^i)\right] \qquad (3)$$

Une condition nécessaire et suffisante[17] **pour avoir un équilibre parfait en *stratégies simples* est que les déviations d'une période soient non profitables** pour les n joueurs le long des n+1 sentiers possibles.

La démonstration de la proposition 1 consiste à mettre en évidence d'une part que si une séquence de déviations infinies est profitable, alors une séquence finie le sera et d'autre part que si aucune déviation d'une période n'est profitable, aucune séquence finie de déviations ne le sera par induction à rebours car la dernière déviation d'une séquence finie est équivalente à une déviation d'une période. Logiquement, la profitabilité d'une séquence finie ou infinie de déviations se ramène à la profitabilité d'une déviation à une période. Ce résultat est toutefois conditionnel à l'utilisation du critère d'actualisation des gains et à l'hypothèse 3 (fonction de gain bornée)[18].

Preuve : (nécessité $\Rightarrow$) Si σ est un profil d'équilibre parfait, alors l'inégalité (3) doit être vérifiée, par définition. En effet, si ce n'était pas le cas, il existerait une date t et une histoire h_t telles que l'un des joueurs aurait une stratégie $\tilde{\sigma}_i$ différente de $\sigma_i^{h_t}$ uniquement

[17] Cette condition représentée par l'inégalité (3) est appelée « condition de non déviation ».

[18] De manière plus générale, ce résultat est vrai si le jeu est continu à l'infini, c'est à dire si pour tout joueur i et toute histoire infinie h_∞ avec h_T sa restriction aux T premières périodes, la fonction de gain du jeu répété satisfait : $\lim_{T\to\infty} \sup|G(h_\infty) - G(h_T)| = 0$.

à la date t qui serait une meilleure réponse face à $\sigma_{-i}^{h_t}$. Ceci contredit la propriété de perfection du profil σ.

(Suffisance $\Leftarrow$) Si l'inégalité (3) est satisfaite, aucune déviation d'une période n'est profitable. Il s'agit de montrer aussi qu'aucune séquence de déviations menée par le joueur i ne peut accroître son gain.

Supposons qu'une séquence infinie de déviations soit profitable. Il existe alors une stratégie $\tilde{\sigma}_i$, une date t et une histoire h_t telles que :

$$G_i(\tilde{\sigma}_i, \sigma_{-i}^{h_t}) - G_i(\sigma^{h_t}) = \varepsilon > 0$$

Comme la fonction de gain g_i est bornée par hypothèse et que les joueurs actualisent leurs gains, il existe un T fini tel que la différence de gains reçus dans le sous-jeu commençant en T et actualisés à la date t, soit inférieure à $\varepsilon/2$. La stratégie $\hat{\sigma}_i$ qui est équivalente à $\tilde{\sigma}_i$ jusqu'à la date T, puis équivalente à la stratégie σ_i après T doit alors être supérieure en termes de gains à $\sigma_i^{h_t}$ sur le sous-jeu commençant après h_t d'au moins $\varepsilon/2$. En d'autres termes, si une séquence infinie de déviations est profitable, il existe toujours une séquence finie qui l'est aussi.

Supposons que les dates auxquelles se font les déviations pour la stratégie $\hat{\sigma}_i$ sont t_1, $t_2,\ldots,t_s,\ldots,t_l$ avec $t_{s-1}<t_s$. Considérons la dernière déviation en t_l. Comme les déviations d'une période ne sont pas profitables, cette déviation peut être remplacée par un comportement de non déviation et le gain du joueur i peut soit augmenter, soit rester le même, il ne peut pas être strictement inférieur. Le même raisonnement s'applique à l'avant dernière déviation, puis par induction à rebours à toutes les déviations. En conclusion, si l'inégalité (3) est vérifiée, aucune séquence finie de déviations ne peut être profitable, ce qui élimine aussi la possibilité d'une séquence infinie de déviation qui le serait. Ceci garantit bien la perfection du profil de stratégies σ.

Cette proposition est fondamentale en théorie des jeux répétés. Sans ce résultat, toute caractérisation des équilibres parfaits serait quasi impossible.

Nous définissons $\Sigma^p \subset \Sigma$ comme l'ensemble des profils de stratégies d'équilibre parfait du jeu infiniment répété $\Gamma^\infty = (N,\Sigma,G,\delta)$.

$M^p = \left\{ Q[\sigma] / \sigma \in \Sigma^p \right\}$ est alors l'ensemble des sentiers générés par des profils de stratégies d'équilibre parfait. $M^p \subset S^\infty$. Si $Q \in M^p$, Q est appelé un sentier d'équilibre parfait.

Lemme 2 : *Si* $\Gamma^\infty = (N,\Sigma,G,\delta)$ *satisfait les hypothèses 1 à 4 alors* Σ^p *est non vide.*

Preuve : Sous les hypothèses 1 à 4, il existe au moins un équilibre de Nash noté s^* dans le jeu constituant $\Gamma=(N,S,G)$. $Q^*=\{s^*, s^*, ...,s^*,...\}$ est le sentier stationnaire associé à s^*. Comme aucune déviation n'est profitable sur le sentier Q^*, il existe au moins un profil de stratégies simples d'équilibre parfait $\sigma(Q^0,Q^1,...,Q^n)$ tel que $Q^i=Q^*$ pour tout i=0,1,...,n. Ce profil consiste à répéter indéfiniment l'équilibre de Nash du jeu constituant.

A facteur d'actualisation donné, la caractérisation de l'ensemble des équilibres parfaits Σ^p ou de l'ensemble des sentiers d'équilibres parfaits M^p passe par la recherche des punitions les plus sévères possibles. En effet, une fois connues les pires punitions pour chacun des joueurs, on peut déterminer l'ensemble des sentiers d'équilibres Q^0 que les joueurs peuvent suivre sans être incités à dévier.

3.5 Les codes pénaux :

La dissuasion des déviations passe par des règles de punitions qui forment le *code pénal* du jeu. A la suite d'Abreu, nous définissons les notions de code pénal optimal (*CPO*) et code pénal simple (*CPS*).

Définition 7 : Un code pénal simple ($Q^1,...,Q^n$) est un n-vecteur de profils de stratégies simples[19] $\{\sigma^1(Q^1,...,Q^n),..,\sigma^n(Q^1,...,Q^n)\}$ où le profil simple σ^i décrit la conduite à tenir pour punir le joueur i.

[19] Un code pénal simple est défini par n sentiers ou punitions ($Q^1,...,Q^n$), tandis qu'un profil de stratégies simples est défini par n+1 sentiers ($Q^0,Q^1,...,Q^n$).

Définition 8 : Un code pénal simple ($Q^1,...,Q^n$) est parfait si et seulement si $\sigma^i(Q^1,...,Q^n)$ est un équilibre de Nash pour tout i=1,...,n.

Définition 9 : Un code pénal optimal est un n-vecteur de profils de stratégies ($\underline{\sigma}^1,..,\underline{\sigma}^n$) tel que :

$$G_i[\underline{\sigma}^i]=\text{Min}\{G_i[\sigma] \;/\; \sigma\in \Sigma^p\}\text{pour tout i=1,..n}$$

Le profil $\underline{\sigma}^i$ est le pire équilibre parfait en termes de gains pour le joueur i. Il prescrit la punition la plus sévère que peut se voir infliger le joueur i de manière crédible. Notons $\underline{G}_i=G_i[\underline{\sigma}^i]$. A la suite d'une déviation, aucun joueur ne peut être tenu à un gain actualisé inférieur à $\underline{G}_i$.

Abreu démontre alors un résultat central par rapport à l'objectif de caractérisation de l'ensemble des équilibres parfaits : un code pénal optimal peut toujours se ramener à un code pénal simple.

Proposition 3 : *Dans le jeu $\Gamma^\infty=(N,\Sigma,G,\delta)$, un code pénal optimal simple existe toujours.*

Preuve : Au préalable, nous démontrons le lemme suivant.

Lemme 4 : *Si $\sigma \in \Sigma^p$ et $Q[\sigma] \in M^p$ est le sentier généré par σ alors pour tout $i\in N$ et pour tout t :*

$$g_i^d(q[\sigma](t)) - g_i(q[\sigma](t)) \leq \delta\left[G_i(Q[\sigma];t+1) - \underline{G}_i\right] \qquad (4)$$

La démonstration de ce lemme est simple. Le joueur i ne peut être tenu à un gain actualisé inférieur à $\underline{G}_i$, car c'est le plus petit gain qui peut être infligé de manière crédible au joueur i. Supposons que l'inégalité (4) ne soit pas vérifiée pour un i, s_i et t. Alors il existerait un sous-jeu dans lequel le joueur i pourrait obtenir un gain plus grand en déviant de σ. Dans ce cas, σ ne serait pas un équilibre de sous-jeu parfait.

Revenons maintenant à la proposition 3. Pour tout i=1,...,n, soit $\{\sigma^{il}\}_{l=1}^{\infty} \in \Sigma^p$ une suite de profils de stratégies pour lesquels $G_i(\sigma^{il})$ converge vers $\underline{G}_i$. De plus, les sentiers générés par ces profils de stratégies sont notés $Q[\sigma^{il}]=Q^{il}$. Nous avons $Q^{il} \in M^p$.

Comme Σ est un ensemble compact et G_i est une fonction continue, la suite de sentiers Q^{il} converge vers le sentier $\underline{Q}^i$ défini par $G_i(\underline{Q}^i)=\underline{G}_i$.

Montrons que le profil de stratégies simples $\sigma^i=(\underline{Q}^1,...,\underline{Q}^n)$ défini par les n sentiers limites est un équilibre parfait.

Si σ^i n'est pas un profil d'équilibre parfait, alors il existe un joueur j, un sentier $k\in\{1,...,n\}$ et un τ tels que :

$$g_j^d\,(q^k(\tau)) - g_j(q^k(\tau)) > \delta[G_j(Q^k,\tau+1) - \underline{G}_j]$$

Par continuité de G_j et pour l grand, nous avons :

$$g_j^d\,(q^{kl}(\tau)) - g_j(q^{kl}(\tau)) > \delta[G_j(Q^{kl},\tau+1) - \underline{G}_j]$$

D'après le lemme 4, ceci contredit $\sigma^{kl} \in \Sigma^p$. σ^i est bien un équilibre parfait pour tout $i\in N$.✍

La proposition 3 signifie que si l'on recherche un code optimal, on peut sans risque d'erreur limiter les investigations aux seuls codes pénaux simples. La proposition 5 soutient même que tout code pénal optimal peut être transformé en un code pénal simple sans perte d'efficacité.

Proposition 5 : *Soient $(\underline{\sigma}^1,..,\underline{\sigma}^n)$ un code pénal optimal du jeu et $\underline{Q}^i=Q(\underline{\sigma}^i)$ le sentier généré par $\underline{\sigma}^i$, alors $(\underline{Q}^1,..,\underline{Q}^n)$ définit un code pénal optimal simple.*

Preuve : Comme $\underline{\sigma}^i \in \Sigma^p$ alors $\underline{Q}^i=Q(\underline{\sigma}^i)\in M^P$ et d'après le lemme 4, nous avons les inégalités suivantes :

$$g_j^d\,(\underline{q}^i\,(\tau)) - g_j\,(\underline{q}^i\,(\tau)) \le \delta\Big[G_j\,(\underline{Q}^i\,;\tau+1) - \underline{G}_j\Big]$$

pour tout $j\in N$, $i=1,..,n$, et τ.

D'après la proposition 1, nous pouvons conclure que $\sigma^i(\underline{Q}^1,..,\underline{Q}^n)$ est un équilibre parfait pour tout i. De plus, comme $G_i(\underline{Q}^i)=\underline{G}_i$, $(\underline{Q}^1,..,\underline{Q}^n)$ est un bien un code pénal optimal simple. ✍

Ce code pénal est simple car l'ensemble des conduites à tenir face à n'importe quelle situation et histoire passées est entièrement défini par *n* sentiers. Il est optimal car en cas de déviation, le joueur est puni le plus sévèrement possible.

Proposition 6 : *Supposons qu'un code pénal optimal* ($\underline{\sigma}^1,..,\underline{\sigma}^n$) *existe. Alors* $Q^0 \in M^p$ *si et seulement si pour tout* $i \in N$ *et* τ *:*

$$g_i^d(q^0(\tau)) - g_i(q^0(\tau)) \leq \delta\left[G_i(Q^0;\tau+1) - \underline{G}_i\right] \qquad (5)$$

Preuve : La nécessité découle du lemme 4.

(Suffisance $\Leftarrow$) Soit ($\underline{Q}^1,..,\underline{Q}^n$) le code pénal optimal simple dérivé de ($\underline{\sigma}^1,..,\underline{\sigma}^n$). Le lemme 4 appliqué à chaque $\underline{Q}^i \in M^p$ conduit aux n inégalités suivantes :

$$g_j^d(\underline{q}^i(\tau)) - g_j(\underline{q}^i(\tau)) \leq \delta\left[G_j(\underline{Q}^i;\tau+1) - \underline{G}_j\right]$$

pour tout $j \in N$, $i=1,..,n$, et ι.

Ces n inégalités plus l'inégalité (5) de la proposition 6 nous permettent d'appliquer la proposition 1 et de conclure que $\sigma(Q^0, \underline{Q}^1,..,\underline{Q}^n)$ est un équilibre parfait ; c'est à dire que $Q^0 \in M^p$. ✎

Cette proposition nous donne une condition simple pour caractériser l'ensemble des sentiers qui peuvent être soutenus par des stratégies d'équilibres parfaits. Ce résultat est très utile lorsque les jeux répétés sont appliqués à la concurrence oligopolistique. En effet, **il suffit de déterminer le code pénal optimal du jeu pour évaluer le degré de collusion maximale entre les firmes.** Les deux précédentes propositions peuvent être résumées dans la proposition suivante.

Proposition 7 : *Supposons qu'un code pénal optimal* ($\underline{\sigma}^1,..,\underline{\sigma}^n$) *existe. Alors, on peut construire un code pénal optimal simple* ($\underline{Q}^1,..,\underline{Q}^n$) *avec* $\underline{Q}^i = Q(\underline{\sigma}^i)$ *tel que* $Q^0 \in M^p$ *si et seulement si le PSS* $\sigma(Q^0, \underline{Q}^1,..,\underline{Q}^n)$ *est un équilibre parfait.*

L'originalité d'Abreu vient certainement de la place qu'il accorde aux punitions et codes pénaux. En définissant une punition comme un sentier prescrit à l'ensemble des joueurs lorsque l'un des joueurs a dévié du sentier en cours et un code pénal comme l'ensemble des punitions du jeu répété, il offre une représentation assez juste des stratégies de coopération qui peuvent apparaître dans une relation répetée. Le résultat central est que la coopération sera d'autant plus efficace que les joueurs disposent de punitions sévères.

Les travaux d'Abreu trouvent un prolongement tout naturel dans l'analyse des comportements stratégiques des firmes. Sur les marchés oligopolistiques, les firmes sont appelées à se rencontrer régulièrement. En s'accordant sur un *code pénal*, elles peuvent soutenir un *cartel tacite*. Dans la seconde partie, nous présentons quelques codes pénaux qui peuvent garantir une *collusion* en prix ou en quantités et comparons leur efficacité respective.