CHAPTER 14

Cooperation

1. Through the course of a relationship or partnership, incentives to invest vary: at times, we may be upset over the way the partnership goes, with low expectations about sustained cooperation and little hope that making effort would increase prospects. At other times, we feel good about prospects, ready to invest again in the relationship and worried that decreasing effort might undermine cooperation. We are interested in building a simple model that explains conditions under which cooperation (i.e., high effort level on both sides) can be sustained, at least periodically.

A partnership bears some resemblance to the strategic interaction studied in the previous chapter. When I choose an effort level, I affect whether the other side (my partner) will be satisfied or dissatisfied. The consequence is as in the reputation problem: if my partner follows a dynamic strategy influenced by her satisfaction (reducing effort when dissatisfied), this creates an incentive for me to cooperate (that is, to continue high effort); and doing so may prevent my partner from reducing effort.

There is a difference, however: sustained cooperation requires that *both sides* put in high effort. In particular, if one player's incentives to put in effort are so strong that he prefers to remain cooperative at all times, the incentive for the other side to put in effort is eliminated; both players must thus be provided with incentives to use a dynamic strategy, that is, a strategy that sometimes cooperates and sometimes defects depending on observations.

This chapter explains how this might be done. In essence, there are many reasons why one might want to condition behavior on observations. One is that the benefits of cooperation may vary over time, and that players attempt to use recent experience to determine whether cooperation is currently productive. We formalize this idea in a simple exchange game.[1]

[1] In contrast to the repeated game literature, which often assumes that the exact same game is repeated, we introduce the possibility that the payoff structure is subject to persistent shocks. The presence of these shocks will provide a motive for using dynamic strategies, as these may generate higher welfare than a constant strategy that would cooperate always, irrespective of current conditions.

2. *A gift exchange game.* There are two players who exchange gifts each period. Each has two possible actions, one that corresponds to not making an effort in choosing a gift, and a second corresponding to making a costly effort. Gifts may or may not be perceived as "thoughtful," and a gift is more likely perceived as thoughtful when the other puts in costly effort.

Payoff structure. Actions are $\{C, D\}$ with C representing costly effort. C stands for "cooperation" and D stands for "defection." The cost of effort is γ. The expected payoff that player i receives from the interaction is 0 when the other does not put in effort, and it is x when the other puts in effort.[2] The players' expected payoffs are thus:

	C	N
C	$x-\gamma, x-\gamma$	$-\gamma, x$
N	$x, -\gamma$	$0, 0$

We refer to x as the *benefit of cooperation*.

Process. The benefits of cooperation are assumed to vary, and to model the variations, we assume that x follows a stochastic process. For simplicity, we assume that in each period, with small probability α, there is a new draw of x. We also assume that $x \in \{0, g\}$, with $g > \gamma$, and we designate $Q = \Pr(x = 0)$. This means that effort is either valuable for both players, or unprofitable. One interpretation is that there are times when each player's understanding of the other's needs decreases, and effort becomes useless.[3] Ideally, agents should cooperate in good times (when $x = g$) and defect in bad times (when $x = 0$). Agents, however, do not observe x perfectly.

Observations. Agents do not observe whether times are good or bad, nor the action of the other player. Rather, each receives a private signal $y_i \in Y_i = \{\underline{y}, \overline{y}\}$ which (imperfectly) reflects the other's effort and the benefits x.[4] Specifically, we assume:

$$p = \Pr\{y_i = \overline{y} \mid a_j = C, x_i = g\} \text{ and}$$
$$q = \Pr\{y_i = \overline{y} \mid (a_j, x_i) = (a, x)\} \text{ for all } (a, x) \neq (C, g),$$

[2] Our analysis would carry over to cases where the benefits of cooperation are agent-specific, i.e., equal to x_i for agent i.

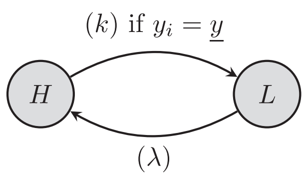
[3] For ease of exposition, we model these changes as perfectly correlated shocks, both in time and value across players. Our analysis would carry over to more general noise structures. In particular it would carry over to cases where benefits are private (i.e., x_i for player i) and where changes in benefits are not necessarily synchronized.

[4] It may be convenient to think of y_i as the realized payoff for player i, in which case one would assume the following relationships: $g = p\overline{y} + (1-p)\underline{y}$ and $0 = q\overline{y} + (1-q)\underline{y}$.

with $p > q$ so that one can refer to $y_i = \underline{y}$ as a “bad” signal and $y_i = \overline{y}$ as a “good” signal. In words, the main difficulty that a player faces is that a bad signal may either reflect bad luck, bad times (low x), or times perceived as bad by the other side (who plays D).

Dynamic strategies. One option for an agent is to ignore signals and follow a constant strategy, playing C at all dates (strategy σ^C) or playing D at all dates (strategy σ^D). Alternatively, an agent may attempt to adjust his effort level to circumstances (ideally whether times are good and whether the other cooperates) and adopt a dynamic strategy whereby his effort level is a function of past experience. We follow the path set out in previous chapters: we restrict attention to a few simple dynamic strategies, and assume that agents adopt the optimal dynamic strategy within a limited set, to be defined next.[5]

To define the dynamic strategies we consider, we assume that each player i may be in one of two possible belief states $\theta_i \in \{H, L\}$, high or low, and that he plays C in belief state H, and plays D in belief state L. Next, we describe how a player’s belief state changes over time. For any given $k \in (0,1)$ and $\lambda \in [0,1]$, we define the strategy $\sigma^{\lambda,k}$ where player i’s belief state changes over time based on past experience according to the following transitions:[6]



In other words, when $\theta_i = H$ in a given period, player i transits to L with probability k when he receives a bad signal, and he remains in H if he receives a good signal. When in state L at the end of a given period, he transits back to H next period with probability λ.[7] We refer to λ as an individual’s *propensity to cooperate*.

The dynamic strategies proposed illustrate a plausible way that agents might use past experience. They echo the family of strategies that our agent used in Chapter 12 to learn about his environment: the agent is uncertain as to when effort is worthwhile; and, when pessimistic (i.e., in state L), he nevertheless *experiments* with cooperation, going back to state H with probability λ. The environment is strategic, but we assume that he considers the same set of strategies as in Chapter 12.

[5] As always, endowing an agent with a richer set of strategies allows him to adjust to circumstances more finely. A simple set will be sufficient for our purpose.

[6] In previous chapters, the probability k was assumed to be 1.

[7] Note that λ induces a dynamic strategy that remains nontrivial in the long run only when $\lambda \in (0,1)$. When $\lambda = 1$ for player i, this player cooperates at all dates; when $\lambda = 0$, he defects at all dates in the long run.

One might argue that the current environment differs from that of Chapter 12 in that signals are received even when he defects, and these signals might be usefully employed to modify how he transits back to state H. We see no objection to following that path, and later in this chapter, we investigate another family of strategies in which player i transits back to state H with probability λ_i, but only does so in events where he receives a good signal (see Section 9). We denote by $\overline{\sigma}^{\lambda_i,k}$ such strategies.

Nevertheless, it is useful to investigate the strategic interaction between less sophisticated agents, not knowing (or not attempting to evaluate) whether the signal observed when in state L should prompt different courses of play.

Strategy restrictions. For the first part of this chapter, we fix particular values for $k \in (0,1)$ and $\lambda \in (0,1)$, define $\widehat{\sigma} \equiv \sigma^{\lambda,k}$, and assume that each player i compares three possible strategies:

$$\sigma_i \in \Sigma_i = \{\sigma^D, \sigma^C, \widehat{\sigma}\}.$$

Next, in Section 6, we continue to fix k but assume that each player i compares all possible strategies $\sigma^{\lambda_i,k}$ as λ_i varies:

$$\Sigma_i = \{\sigma^{\lambda_i,k}\}_{\lambda_i \in [0,1]}.$$

Finally, in Section 9, we continue to fix k and assume that each player i compares all possible strategies $\overline{\sigma}^{\lambda_i,k}$ as λ_i varies:

$$\Sigma_i = \{\overline{\sigma}^{\lambda_i,k}\}_{\lambda_i \in [0,1]}.$$

We shall see how this latter family affects incentives and the scope for cooperation.

Long-run values and equilibrium. When players adopt nontrivial dynamic strategies, they typically alternate between *cooperative phases* in which both players cooperate, and *punishment phases* where at least one player defects. For example, if they both adopt $\sigma^{\lambda,k}$ with λ small, one expects that punishment phases will be of long duration, with players missing many opportunities for profitable cooperation. On the other hand, if λ is high, their propensity to cooperate may be so large that they attempt to cooperate in many events where cooperation is not profitable (when $x = 0$).

Thus, adjustment to the underlying condition x cannot be perfect, even if agents could jointly choose λ_1 and λ_2 to maximize their joint interest. Our aim below is to examine the strategic consequences of letting each player follow the dynamic strategy σ_i that suits best his own interest. Toward this end, we define the long-run payoff $v(\sigma)$ induced by $\sigma = (\sigma_1, \sigma_2)$, and then define equilibrium behavior.

Define the state s as the triplet (θ_1, θ_2, x), where $\theta_i \in \{H, L\}$ is the current belief state of player i and $x \in \{0, g\}$ is the current benefit. Denote by S the set of possible such states. For any pair of strategies $\sigma = (\sigma_1, \sigma_2)$, and any state $s \in S$, one may define the long-run frequency of state s induced by σ.[8] Call $\phi^\sigma(s)$ that long-run frequency. Also denote by $u(s)$ players' expected gain in a period where the state is s. We associate to each σ the long-run value $v(\sigma)$ induced by σ:

$$v(\sigma) = \sum_{s \in S} u(s)\phi^\sigma(s).$$

Given the simple structure of the game, a player, say player 1, only cares about two events: the event A where $\theta_2 = H$ and $x = g$; and the event B where $\theta_1 = H$. Defining ϕ_z^σ as the long-run frequency of event z,[9] we have:

$$v_1(\sigma) = g\phi_A^\sigma - \gamma\phi_B^\sigma.$$

This expression reflects the main tradeoff: a more cooperative strategy by player 1 is more costly (higher ϕ_B^σ), but it may increase the chances of profitable cooperation (higher ϕ_A^σ).

We say that σ is an *equilibrium* if for each player i and each strategy $\sigma_i' \in \Sigma_i$:

$$v_i(\sigma_i', \sigma_{-i}) \leq v_i(\sigma).$$

We are interested in finding the model parameters for which cooperation occurs in equilibrium.

3. *Incentives to cooperate.* A key aspect of dynamic strategies is that they create incentives to cooperate (that is, put in effort), for reasons similar to those in the previous chapter. Assume that player 2 follows the dynamic strategy $\widehat{\sigma}$. Then, through his choice of action, player 1 can influence the belief state (and behavior) of player 2. If player 1 puts in high effort, he makes it more likely that player 2 will remain in the good belief state, and this generates incentives to put in effort if the cost of effort γ is not too high. We make this precise next.

To simplify notation, we let ϕ_z^D (respectively ϕ_z^C and $\widehat{\phi}_z$) denote the long-run probability of the event z when player 1 follows σ^D (respectively σ^C and $\widehat{\sigma}$). We have:

$$v_1(\sigma^D, \widehat{\sigma}) = \phi_A^D g$$
$$v_1(\sigma^C, \widehat{\sigma}) = \phi_A^C g - \gamma$$
$$v_1(\widehat{\sigma}, \widehat{\sigma}) = \widehat{\phi}_A g - \widehat{\phi}_B \gamma.$$

[8] We adopt the convention that if player i adopts the constant strategy σ^C (respectively σ^D) he remains in state H (respectively L).

[9] $\phi_z^\sigma = \sum_{s \in z} \phi^\sigma(s)$.

Because player 2 is responsive to signals (he follows $\widehat{\sigma}$), and because $p > q$, greater effort by player 1 increases the frequency with which player 2 puts in effort. As can be easily verified:

$$\phi_A^C > \widehat{\phi}_A > \phi_A^D.$$

The consequence is that if γ is not too large, player 1 has incentive to cooperate. The condition that σ^C improves over σ^D is:

$$\gamma < (\phi_A^C - \phi_A^D)g,$$

and the condition that $\widehat{\sigma}$ improves over σ^D is:

$$\gamma < \frac{\widehat{\phi}_A - \phi_A^D}{\widehat{\phi}_B} g.$$

In words, incentives to cooperate require that effort induce a sufficiently large increase in the probability of cooperation.[10]

4. *Incentives to use a dynamic strategy.* That player 1 is willing to cooperate is not the end of the story. If the consequence is that player 1 is willing to cooperate all the time (i.e., use σ^C), then player 2 will have no incentive to use the dynamic strategy $\widehat{\sigma}$. An equilibrium involving cooperation requires that *both* sides use the dynamic strategy $\widehat{\sigma}$. Formally, this requires:

$$v_1(\widehat{\sigma}, \widehat{\sigma}) > v_1(\sigma^C, \widehat{\sigma})$$

that is:

$$\widehat{\phi}_A g - \widehat{\phi}_B \gamma > \phi_A^C g - \gamma$$

or equivalently

$$\gamma > \frac{\phi_A^C - \widehat{\phi}_A}{1 - \widehat{\phi}_B} g.$$

Intuitively, under the above condition, a player has incentive to use the dynamic strategy (which implies making no effort, sometimes) rather than the cooperative strategy because this allows him to economize on costs in events where effort is not worthwhile. If the other player is persistently in belief state L, or if the benefit x is persistently low, it is unnecessarily costly to put in effort. By using a dynamic strategy, player 1 may hope to reduce effort,

10 Alternatively, for fixed probabilities $\phi_A^C, \widehat{\phi}_A$ and ϕ_A^D ranked as above, incentives to cooperate require that gains for defecting against a cooperative partner be not too large – a standard insight.

putting forth effort only when it is profitable (i.e., when $(\theta_2, x) = (H, g)$).[11] One difficulty, of course, is that player 1 does not observe (θ_2, x) in the current period, and he may choose not to put in effort in events when it would, in fact, be profitable. The effect is a decrease in cooperation: $\widehat{\phi}_A$ is smaller than ϕ_A^C.

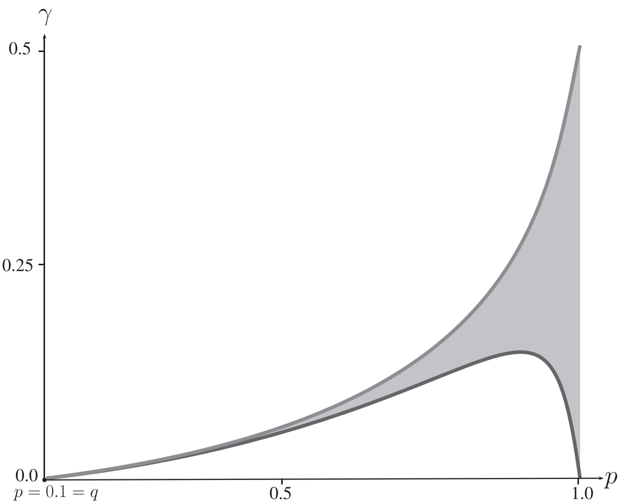
There is thus a tradeoff between reduced cooperation and lower expenses, and this tradeoff tilts behavior in favor of the dynamic strategy when the cost of effort γ is sufficiently high.

5. *Equilibrium.* Combining previous inequalities, the dynamic strategy $\widehat{\sigma}$ improves upon both σ^D and σ^C when

$$\frac{\phi_A^C - \widehat{\phi}_A}{1 - \widehat{\phi}_B} g < \gamma < \frac{\widehat{\phi}_A - \phi_A^D}{\widehat{\phi}_B} g.$$

Checking the compatibility of these inequalities is easy when p is close to 1 and α sufficiently small. Then, when $x = g$, bad signals are rare, so using $\widehat{\sigma}$ rather than σ^C cannot hurt much. In contrast, when $x = 0$, the strategy σ^C is very bad because it involves effort with zero expected payoff in return: the strategy $\widehat{\sigma}$ is better because it reduces the number of periods in which effort is made. Further details are in the Appendix.

The following figure provides the region of parameters (p, γ) where $(\widehat{\sigma}, \widehat{\sigma})$ is an equilibrium, when $q = 0.1$, $\lambda = 0.1$, $k = 0.2$, $Q = 1/2$ and α is very



[11] This incentive to use a nontrivial dynamic strategy potentially exists even in a standard repeated game where x doesn't vary, because one wishes to adjust to the behavior of the other player. Variations in x amplify that incentive, enabling cooperation as an equilibrium phenomenon in our model (see Footnote 12 on next page).

small.[12] The range of values of γ for which $(\widehat{\sigma},\widehat{\sigma})$ is an equilibrium expands when p approaches to 1.

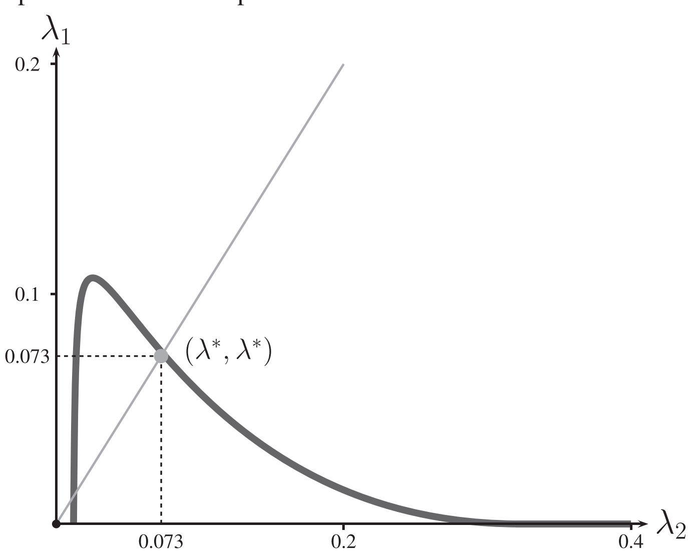
6. *More sophisticated players.* We assumed that players can only compare few strategies. This makes it relatively easy to check for incentives. Our objective below is to understand whether cooperation can still be obtained when players consider more strategies, with the aim of endogenizing players' propensities to cooperate λ_1 and λ_2, that is, the degree to which players experiment with cooperation when in state L. Toward this end, we fix k and assume that players use the rule $\sigma^{\lambda_i,k}$ that maximizes their long-run payoff, among all possible values of $\lambda_i \in [0,1]$.

Computations are similar. Simplifying notation, we now refer to $\phi_z^{\lambda_1,\lambda_2}$ as the long-run frequency of event z when the strategy pair (λ_1,λ_2) is used. We have:

$$v_1(\lambda_1,\lambda_2) = \phi_A^{\lambda_1,\lambda_2} g - \phi_B^{\lambda_1,\lambda_2} \gamma.$$

Player 1 faces the following tradeoff. By increasing λ_1, he increases $\phi_A^{\lambda_1,\lambda_2}$ because this increases the likelihood of cooperation in events where player 2 is in, or about to be in, belief state H. However, he also increases the frequency $\phi_B^{\lambda_1,\lambda_2}$.

The thick curve below shows the best response of player 1 as a function of λ_2, for specific values of the parameters.[13]



[12] The existence of a set of parameters for which cooperation is possible contrasts with Compte and Postlewaite (2015), who show that this type of strategy cannot support equilibrium cooperation. The difference is that, here, we consider a setup in which the benefits of cooperation vary over time, which makes it worthwhile to use $\sigma^{\lambda,k}$ rather than either constant strategy.

[13] We use $p = 0.95, q = 0.1, k = 0.2, \gamma = 0.3$ and $Q = 0.5$.

The thick curve illustrates that when λ_2 is too small, there is no point in trying to cooperate. The probability of reciprocal cooperation is too low, and player 1's best response is to set λ_1 to 0. It also illustrates that when λ_2 is too high, player 1 is better off taking advantage of his cooperative-minded partner, and his best response is again to set λ_1 to 0.

For intermediate values of λ_2, the best response of player 1 is to experiment somewhat with cooperation ($\lambda_1 \in (0,1)$). At $\lambda_2^* \simeq 0.073$, the best response of player 1 is to set $\lambda_1^* = \lambda_2^*$: a symmetric equilibrium. At this equilibrium, the propensities to cooperate are substitutes: the larger λ_2, the smaller player 1's incentives to cooperate.

We have exhibited an equilibrium involving cooperation, but we do not suggest that cooperation could be obtained for all parameter values. When signals are noisier ($p = 0.9$), the best response has a similar shape, but the two curves do not cross. The response of player 1 is never strong enough, and the only equilibrium is that both remain in state L.

7. *Punishment and recoordination.* When players use nontrivial dynamic strategies, behavior alternates between phases of cooperation and phases of defection (in which at least one player defects). The defection phase may be interpreted as a *punishment* phase that ends once players manage to *recoordinate*. We describe this alternation below.

A defection phase begins when a player receives a bad signal and reacts to it by moving to belief state L. He then starts defecting, good signals become less likely, and it becomes likely that the other player will soon move to belief state L as well: both players remain in belief state L for some time. The defection phase is initiated by one player reacting to a bad signal that may have two causes: either bad luck (both players put in effort and yet a bad signal occurs – $p < 1$); or a change in the underlying benefits from cooperation ($x = 0$). In the former case, it is a bad idea to trigger a defection phase, while in the latter case, it is a good idea to stop attempting to sustain cooperation that is no longer worthwhile.

Once both players are in belief state L, the return to cooperation may take time. The dynamic strategy calls for occasionally checking (i.e., with probability λ) whether the time is ripe to cooperate again. The difficulty is that these attempts need not be simultaneous, so players may remain in the defection phase for quite some time if λ is low. When these attempts happen to be simultaneous (or close to simultaneous if k is not too close to 1), players may start a cooperative phase again, possibly a very long phase if p is close to 1 and $x = g$.

In summary, the dynamic strategy plays two roles.

(i) It acts as *a punishment device* that moves behavior to a defection phase when the state is low or a player defects. This is an incentive mechanism similar to that in reputation. Potentially, this incentive mechanism is more powerful for player 1 when the punishment is likely to last long, that is, when λ_2 is low.

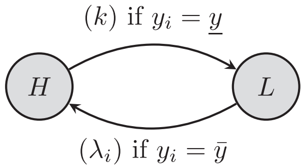
(ii) It acts as a *recoordinating mechanism* that brings behavior back to a cooperative phase when cooperation is profitable. When λ_2 is too low however, recoordination is sufficiently difficult that attempts to move cooperation back on track are too likely to fail – incentives to cooperate vanish.

8. *A static perspective.* Even though the interaction is dynamic, the game can be described as a simple static game in which each player i chooses his propensity to cooperate λ_i. The higher λ_i, the more costly the strategy is (because ϕ_B increases with λ_1), but the more one can expect the other to reciprocate (because ϕ_A increases with λ_1).

More precisely, for low (but not too low) values of λ, propensities are strategic complements, meaning that a higher propensity of one's partner to cooperate generates a higher propensity for oneself. For higher values of λ, propensities become strategic substitutes. At the equilibrium we obtain, propensities are strategic substitutes, and players would be better off if they could jointly commit to a higher level λ_i^{**}. For the parameters that we chose, if players could jointly set the pair (λ_1, λ_2), they would optimally set them at a higher level $(\lambda_1^{**} = \lambda_2^{**} = 0.21)$.

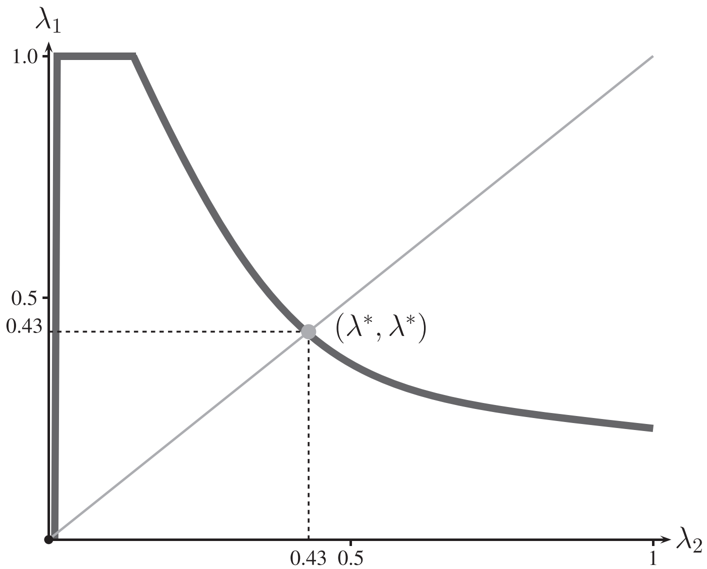
The degree to which players cooperate in equilibrium depends on the exact shape of the functions ϕ_A and ϕ_B. The exact shape of these functions depends on the family of strategies considered. Other families may generate different shapes, but with a similar tradeoff (with higher effort translating into high costs and more reciprocation). Some families of strategies, however, could do a better job at inducing higher equilibrium effort, for example by facilitating recoordination back to cooperation. We investigate this next.

9. *Recoordination made easier.* To illustrate that some families of strategies may be more conducive to cooperation, we investigate a family in which each player i goes back to state H with probability λ_i upon receiving a good signal.



Compared to previous transitions, attempts to recoordinate to cooperation are only made in events where $y_i = \bar{y}$. An essential difference is that once a player, say i, attempts to cooperate, and if the underlying state happens to be good ($x = g$), there is a greater chance that the other player (j) observes the good signal ($y_j = \bar{y}$) when the state happens to be good. In particular, if p and λ_j are large, recoordination is likely to be successful. Thus, one virtue of the transitions above is that recoordination attempts may be infrequent (if $q\lambda$ is small), and yet successful when $x = g$ (if $p\lambda$ is large). This is in contrast with previous transitions where recoordination attempts had to be relatively frequent to be successful.

The thick curve below shows the best response of player 1 as a function of λ_2, for the same parameter values as before.[14]



The figure illustrates as before that when λ_2 is too small, there is no point in trying to cooperate. For high values of λ_2, incentives decrease because each player would prefer that the other incur the cost of attempting to recoordinate. Indeed, the current state might be $x = 0$, and in such events, an attempt to recoordinate is bound to fail.

At $\lambda_2^* \simeq 0.43$, the best response of player 1 is to set $\lambda_1^* = \lambda_2^*$, a symmetric equilibrium. Note that in equilibrium, if both players are in state L, a player attempts to recoordinate with probability $q\lambda_2^* = 0.043$. So these attempts are less frequent than before, but they succeed more frequently. As a result, players' expected values are higher in this equilibrium (0.3, compared to 0.2 before).

Another consequence is that there are many parameter values for which cooperation was impossible before, but become possible under this new family.

Further Comments

Public and private signals. *Much of the literature has been concerned with distinguishing between cases where signals are public and cases where signals are private. This is largely for technical reasons. Equilibrium construction has*

[14] Again, $p = 0.95, q = 0.1, k = 0.2, \gamma = 0.3$ and $Q = 0.5$.

mostly relied on dynamic programming techniques, and these are easier to implement when at any point in time, players are in complete certainty about each other's continuation strategy.[15]

There are several difficulties with this:

(i) Descriptively, it is not very compelling that players would be certain, at any point of time, of each other's current action. Our everyday experience is that we are rarely certain about how others behave.

(ii) A primary role of signals is to enable players to learn about what others are doing. By focusing on equilibria in which each knows others' plans of action, theory neglects this role. Learning is then absent from equilibrium behavior.

(iii) The theory might suggest that since it is technically easier to deal with public signals, it should also be easier for agents to cooperate when signals are public. One lesson from our exercise is that the seeds for cooperation lie in a simple ingredient (the fact that players use dynamic strategies for simple learning reasons), and that the public or private nature of signals is not necessarily a help or a hindrance to players conditioning behavior on their observations.

(iv) Whether signals are public or private, how one reacts to signals is likely to be private in practice, that is, partially influenced by other unmodeled private elements or perceptions.[16] *So the distinction between public and private signals is likely to be irrelevant, or at least less important than what theory suggests.*

Tit for Tat. *Tit for Tat is a strategy in which a player echoes his opponent's latest actions: a Tit-for-Tat player cooperates after seeing cooperation, and defects after seeing a defection. The dynamic strategies that we considered in Section 9 have a Tit-for-Tat flavor (with a hysteresis twist): they produce stochastic reactions that trigger a cooperative mood*[17] *after good signals (with probability λ, the degree to which one is forgiving), and a noncooperative mood after bad signals (with probability k, the degree to which one reacts promptly to bad signals). These strategies are natural generalizations of Tit for Tat in environments in which one cannot observe perfectly the actions of others. In the spirit of Axelrod's tournaments, one could examine which*

[15] With public signals, dynamic programming techniques typically summarize the future by a continuation value vector (see for example Abreu, Pearce and Stacchetti (1990), and Fudenberg, Levine and Maskin (1994) for the case of very patient players). With private signals, dynamic programming techniques are more difficult to implement because future play is uncertain, and potentially depends on each player's belief about signals received by others. Attempts to use dynamic programming techniques in this context have produced "belief free equilibria," the fragility of which we discuss in Compte and Postlewaite (2013).

[16] See also the comments in Chapter 10, page 141.

[17] The mood is potentially long lasting. Behavior is not solely driven by the latest signal observed (hence the hysteresis twist).

of these strategies would be selected by evolution. Our analysis provides a partial answer to that question, endogenizing forgiveness for a given stochastic environment and a given k.[18]

Further research/applications. Through strategy restrictions, one obtains a more parsimonious treatment of repeated relationships, which enables us to represent them as static games with few strategic variables. This makes it possible to establish a connection with a classic static contribution game, in which one's propensity to experiment the cooperative arm can be interpreted as a contribution to the relationship. In that spirit, one may investigate the existence and stability of asymmetric equilibria between a sophisticated player who conditions experimentation on receiving a good signal, and a less sophisticated player who experiments unconditionally. One might also investigate equilibria in which each player has both strategic variables available (i.e., the propensity to experiment unconditionally, and the propensity to experiment after a good signal), and examine the connection with strategic experimentation models in which players share a common interest (see Chapter 12).

Another interpretation of our strategy restrictions is that we constrain each agent to hold a limited number of belief states, and (partially) endogenize belief revisions by focusing on a few stochastic transitions and finding the agent's optimal transition probabilities.[19] We believe this technique could be useful in many contexts, in particular when one wishes to represent agents as alternating between a few states of mind (convinced or unconvinced, suspicious or trustful, attentive or inattentive) as a function of the evidence or the messages received, rather than representing them as holding precise probabilistic beliefs (derived from the knowledge of the model and the application of Bayes rule).

Appendix

We consider the limit case where α is arbitrarily close to 0. Under this assumption, we can compute long-run frequencies as if there were an initial draw of x that did not change. We also examine the case where p is close to 1.

Consider first the event where $x = g$. If player 1 uses either $\widehat{\sigma}$ or σ^C and if p is sufficiently close to 1, both players seldom exit from the belief state (H,H). This means that conditional on $x = g$, they both remain in state (H,H) most of

[18] Our equilibrium is locally stable, as the best response function $\lambda_1(\lambda_2)$ is relatively flat near the equilibrium value that we identified.

[19] Other work along this line includes Wilson (2014).

the time. It follows that

$$\widehat{\phi}_A \simeq \phi_A^C \simeq \Pr(x = g) = 1 - Q.$$

Consider now the event $x = 0$. Whatever actions are being used, each player receives the good signal $\overline{y}$ with probability q in every period. When player 1 uses $\widehat{\sigma}$, he thus returns to H periodically, and in the long run, he is in belief state H with frequency ϕ_0 that solves $\phi_0 = (1-(1-q)k)\phi_0 + \lambda(1-\phi_0)$, or equivalently, $\phi_0 = \frac{\lambda}{\lambda+(1-q)k}$. To compute $\widehat{\phi}_B$, we aggregate both events ($x = 0$ and $x = g$) and obtain

$$\begin{aligned}\widehat{\phi}_B &= \Pr(x = g)\widehat{\phi}_A + \Pr(x = 0)\phi_0 \\ &\simeq (1 - Q) + Q\phi_0\end{aligned}$$

implying that $1 - \widehat{\phi}_B \simeq Q(1 - \phi_0)$.

References

Axelrod, R. (1984). *The Evolution of Cooperation*, New York: Basic Books.

Abreu, D., Pearce, D. and Stacchetti, E. (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58, 1041–1063.

Compte, O. and Postlewaite, A. (2013). Belief free equilibria, mimeo. University of Pennsylvania, 13–020.

Compte, O. and Postlewaite, A. (2015). Plausible cooperation. *Games and Economic Behavior*, 91, 45–59.

Fudenberg, D., Levine, D. and Maskin, E. (1994). The folk theorem with imperfect public information. *Econometrica*, 62, 997–1039.

Wilson, A. (2014). Bounded memory and biases in information processing. *Econometrica*, 82(6), 2257–2294.