CHAPTER 19

Unraveling

1. Sophistication allows one to precisely tailor behavior to circumstances. In strategic interactions, circumstances may include not only the particular payoff structure an agent faces, but also the behavior of other agents. For some interactions, sophistication improves players *ability to coordinate*, generating multiplicity of possible outcomes.[1] For other interactions, sophistication gives rise to competitive pressures or *unraveling effects*.

The two-person Nash demand game provides a simple illustration of the multiple ways in which players may coordinate in a strategic interaction. There is a pie of given size, S, to be shared, and each makes a demand. If demands are compatible, each gets his demand. If incompatible, the pie is lost. In this game, *any* sharing (x_1, x_2) of the pie (such that $x_1 + x_2 = S$) is an equilibrium outcome.

But players' ability to adjust to one another does not necessarily generate multiple outcomes. The centipede game illustrates the competitive pressures that sophistication creates. In this game, the size of the pie increases over time, and the first to exit gets a larger share of the pie, while the other gets the remainder. Thus there is a joint incentive to wait to enlarge the size of the pie along with a private incentive to exit first, driving both players to exit immediately.

2. In both examples, one may feel that adjustment to the other's behavior is too easily achieved, especially since equilibrium analysis is silent about how equilibrium outcomes come about. A more realistic description of the interaction may thus rely on limiting each player's ability to adjust to others' behavior. This can be done by assuming that agents tremble in making decisions, or by introducing uncertainty on the exact size of the pie or asymmetries in information. As is well known,[2] such imperfections may drastically reduce coordination possibilities in the Nash demand game, and the model then typically generates a single prediction.

In the context of the centipede game, one expects that reducing sophistication will reduce private incentives to exit before the other – smoothing

[1] This is the topic of Chapter 20.
[2] See Nash (1953) and Carlson (1991).

competitive pressures, at least to some extent. One purpose of this chapter is to affirm that intuition. Another purpose is to propose a simple model of limited sophistication, based on direct strategy restrictions.

3. *A simple centipede game.* We consider a pie growing linearly over time. The pie stops growing as soon as a player exits, at which point the game terminates and the pie is shared. The player who exits first gets a larger share of the pie.[3]

Formally, the pie at date t has size St, and we denote by a the share of the agent who exits first ($a > 1/2$).[4] The other player gets the remainder (a share of size $1-a$). We also denote by $v_i(t_i, t_j)$ player i's payoff when the exit dates chosen are t_i and t_j. To avoid technical difficulties we assume discounting; there is a discount factor, but we take it to be arbitrarily close to 1.[5] Over the range of relevant dates, and letting $\rho = (2a-1)/a$, we thus have:

$$\begin{aligned} v_i(t_i,t_j) &= aS[\min(t_i,t_j) - \Delta(t_i,t_j)] \text{ with} \\ \Delta(t_i,t_j) &= 0 \text{ if } t_i < t_j \text{ and} \\ &= \rho t_j \text{ if } t_j < t_i. \end{aligned}$$

The term $\min(t_i, t_j)$ reflects the joint incentive to set high exit dates. The term $\Delta(t_i, t_j)$ reflects the penalty from exiting last. The penalty increases over time, and that penalty is smaller when a is close to $1/2$. When players are allowed to choose exit dates with arbitrary precision, this game has a unique equilibrium in which both exit immediately.[6]

4. *Strategy restrictions and payoffs.* To limit players' ability to adjust to one another, we assume that even when a player *targets* a particular exit date λ_i, his actual exit date t_i is stochastic: t_i is a function $\mathbf{t}(\lambda_i, \tau_i)$ of the target λ_i and a random component τ_i, where each τ_i is drawn independently from a distribution with density h, and cumulative distribution H. In the first part of this chapter, we assume that the noise term is additive:

$$t_i = \mathbf{t}(\lambda_i, \tau_i) = \lambda_i + \tau_i,$$

and in later computations, we assume that each τ_i is exponentially distributed with the same parameter $1/\mu$:

$$H(\tau_i) = 1 - \exp -\tau_i/\mu \text{ for } \tau_i \geq 0, H(\tau_i) = 0 \text{ otherwise.}$$

[3] This is a continuous time version of the centipede game. The centipede game was originally introduced by Rosenthal (1981). See Brunnermeier and Morgan (2010) for a continuous time version. This game is also called a preemption game: there are two shares available, and when you exit first you preempt the other players and get the larger share.

[4] We assume equal sharing when both exit at the same date, but under the assumptions to be made, this event will have probability 0.

[5] This ensures that payoffs remain bounded, and that for any discount factor δ, there is an upper bound on the termination date.

[6] The form of penalty is not important. We briefly discuss other payoff structures in Section 9.

The parameter μ reflects noise in decision making. A plausible interpretation of this noise is that a player can be in two possible states of mind: one in which he is inclined not to exit (i.e., prior to λ_i) and a second in which he thinks that he should exit (i.e., after λ_i), but doesn't exit with certainty: he exits with probability dt/μ per period of time dt. With this interpretation in mind, the additive noise τ_i seems natural.

For each player i, a target date λ_i generates a distribution over exit dates t_i. We denote by $V_i(x, \lambda)$ the expected payoff obtained by player i when he targets date $\lambda + x$ while the other targets date λ:

$$V_i(x, \lambda) \equiv E_{\tau_i, \tau_j} v_i(\lambda + x + \tau_i, \lambda + \tau_j)$$

5. *Equilibrium.* We are looking for a symmetric equilibrium, that is, a target date λ^* such that neither player wishes to select a different target date, or equivalently such that $V_i(x, \lambda^*)$ is maximized at $x = 0$. Payoffs being proportional to aS, we normalize aS to 1. We have:

$$V_i(x, \lambda) = \lambda + E \min(x + \tau_i, \tau_j) - \rho \int (\lambda + \tau_j) h(\tau_j)(1 - H(\tau_j - x)) d\tau_j.$$

In equilibrium, the target date λ^* satisfies $\frac{\partial V_i(0, \lambda^*)}{\partial x} = 0$, implying that:[7]

$$\rho \left(\lambda^* + E[\tau \mid \tau_1 = \tau_2]\right) = E\left[\frac{1 - H}{h} \mid \tau_1 = \tau_2\right]. \tag{19.1}$$

The right-hand side is the benefit from a marginal increase in the target date, which exploits the fact that the other player's exit date is dispersed, hence possibly higher (when the distribution is concentrated, h is large and this term is small). The left-hand side is the marginal cost of increasing the target date.

For an exponential distribution with parameter $1/\mu$ we get:[8]

$$\lambda^* = \frac{2 - \rho}{2\rho} \mu.$$

The expression reflects the two forces at work. For ρ close to 0, the private incentive to exit before the other player is weak, because there is not much to gain from early exit. When μ is large, these incentives are even weaker, because getting the larger share with meaningful probability requires a substantial (and costly) decrease in the target date.

6. *Noisy instruments and incentives to delay.* As discussed in Chapter 4, one can think of each agent being endowed with a set of instruments that enables

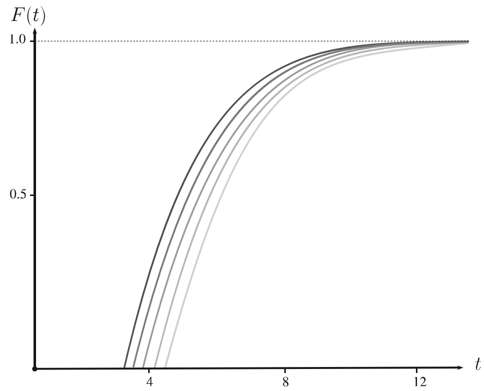
[7] Differentiating the above expression yields $\int h(\tau)(1 - H(\tau)) d\tau = \rho \int (\lambda + \tau)(h(\tau))^2 d\tau$, hence the desired conclusion. Note that this is a necessary condition for equilibrium. For the distributions we consider, the local condition is sufficient.

[8] For the exponential distribution, $\frac{1-H}{h} = \mu$ and $E[\tau \mid \tau_1 = \tau_2] = \mu/2$.

him to parse the environment and adjust to it. From an ex ante perspective, the pair (λ, μ) defines a mixed strategy $\sigma^{\lambda,\mu}$ over possible exit dates, and we have just examined the consequence of restricting each player's strategy to the family

$$\Sigma^{\mu} = \{\sigma^{\lambda,\mu}\}_{\lambda}.$$

Each strategy $\sigma^{\lambda,\mu}$ can be represented as a cumulative distribution $F^{\lambda,\mu}$ over exit dates. The figure below shows the functions $F^{\lambda,\mu}(t)$ for $\mu = 2$ and for various targets λ.



When $\mu = 0$, $\{F^{\lambda,\mu}\}_{\lambda}$ is a family of step functions. When each player has this family available, unraveling is extreme and leads to immediate exit. As μ gets larger, the functions $F^{\lambda,\mu}(.)$ get flatter, and unraveling weakens.

Intuitively, randomness in the *other's action* generates incentives to delay exit, and this effect is sufficient to weaken unraveling. To see why, assume that only player 2's exit date is subject to noise ($\mu_1 = 0$ and $\mu_2 = \mu > 0$). When player 2 targets λ, the optimal exit date t^* for player 1 satisfies:[9]

$$t^* = \frac{1 - H(t^* - \lambda)}{\rho \ \ h(t^* - \lambda)}.$$

For exponential distributions, $t^* = \mu/\rho$ independently of the target date λ.[10]

Another consequence of noise is randomness in one's *own action*: a player should take this into account in setting his target, as on average, his exit date occurs $E\tau = \mu$ after the target. Anticipating this delay, the target should be

9 A marginal increase of dt yields a gain dt in events where $\lambda + \tau_2 > t$, and a loss $\rho th(\tau)dt$ for τ such that $\lambda + \tau = t$, hence the first-order condition $1 - H(t - \tau) = \rho th(t - \lambda)$.

10 Similarly, in response to a deterministic exit date t, the optimal target is λ^* such that $\rho t = H(t - \lambda^*)/h(t - \lambda^*)$. Combined with the expression for t^*, we obtain that $H(t^* - \lambda^*) = 1/2$. When the distribution is exponential, we obtain $t^* - \lambda^* = \mu Ln2$, hence $\lambda^* = \mu(1/\rho - Ln2) > 0$.

reduced. Expression (19.1) reflects this component, and shows that in the symmetric equilibrium the target is reduced in equilibrium by $E[\tau \mid \tau_1 = \tau_2] = \mu/2$, hence only half of the expected individual delay μ.

7. *Multiplicative noise.* We provided a motivation for additive noise, but other sources of noise seem equally plausible. The previous analysis suggests that the parameter $\eta \equiv \frac{\mu(2-\rho)}{2\rho}$ should drive the extent of delay, and the agent might wish to adjust his target based on his perception of η. In that case, multiplicative noise about perceptions of η seems plausible, and the relevant model might be

$$t_i = \mathbf{t}(\lambda_i, \eta_i, \tau_i) = \lambda_i \eta_i + \tau_i$$

where the coefficients η_i are drawn independently from a lognormal distribution ($\log \eta_i \sim \mathcal{N}(0, \sigma^2)$). We denote by H^m the cumulative distribution of η_i and by h^m its density.

Multiplicative noise on the exit date may qualitatively alter the outcome. In the additive noise case, the incentive to delay eventually vanishes for large exit dates, because the penalty $\Delta(t_i, t_j)$ becomes large compared to the magnitude of the noise, explaining why target dates have the same order of magnitude as μ.

In contrast, with multiplicative noise, the incentive to delay exit does not vanish for large exit dates, because the randomness in the other's exit date is larger for large exit dates. We illustrate this by ignoring the additive noise (i.e., $\tau_i \equiv 0$), and by computing the expected payoff $V_1^m(x,\lambda)$ for player 1 when his target is $x\lambda$, while the other's target is λ. Player 1 gets $\lambda \min(x\eta_1, \eta_2)$ and he suffers a penalty of $\rho\lambda\eta_2$ when $x\eta_1 > \eta_2$. Thus we have:

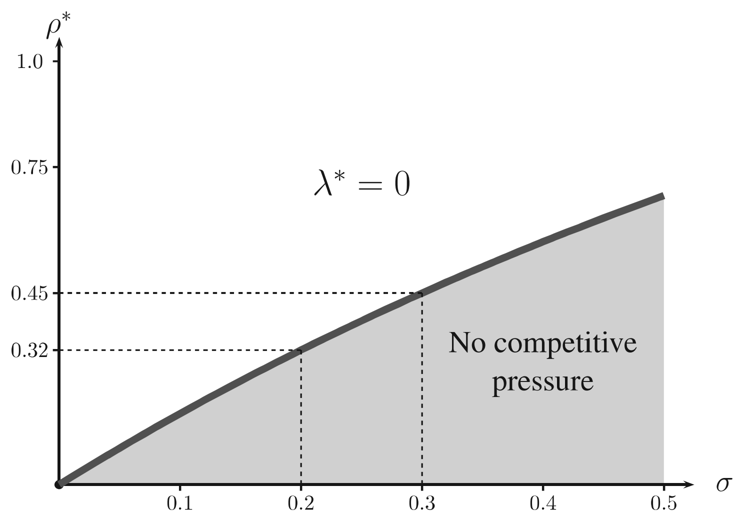
$$V_1^m(x,\lambda) = \lambda[E \min(x\eta_1, \eta_2) - \rho \int \eta(1 - H^m(\eta/x))h^m(\eta)d\eta].$$

Consequently, incentives are independent of λ, and it is immediate that when ρ is below some threshold ρ^*,[11] $\frac{\partial V_1^m(1,\lambda)}{\partial x} > 0$ for all λ, demonstrating that players wish to increase their target date indefinitely – competitive pressures completely vanish.[12] The next figure shows the locus of ρ^* for lognormal noise, as a function of the standard deviation σ.[13]

[11] Differentiating the above expression, one obtains $\rho^* = \int \eta h(\eta)(1 - H(\eta))d\eta / \int \eta^2 (h(\eta))^2 d\eta$.

[12] Of course, with a fixed discount factor, this incentive to set the target date above the other would vanish for sufficiently large targets, and we would find an equilibrium target λ^*. This target would decrease as players become more impatient.

[13] Our analysis ignores the possibility that both choose the target $\lambda_i = 0$. In the absence of additive noise, this target would be implemented without noise, and could constitute an equilibrium. Additive noise, however, would preclude an equilibrium with immediate exit. We omit this issue in the rest of this chapter.



8. *End game effects.* The structure of noise is important and may hinge on how the game is framed. If the game exogenously terminates at date T, and if players are perfectly aware of that termination date, it seems plausible that the noise in exit dates vanishes as players approach the terminal date. An expression for exit dates consistent with this constraint is then:

$$t_i = \mathbf{t}(\lambda_i, \eta_i, T) = \frac{\eta_i \lambda_i T}{1 + \lambda_i \eta_i} \tag{19.2}$$

where η_i is lognormally distributed (i.e., $\log \eta_i \sim \mathcal{N}(0, \sigma^2)$) and $\lambda_i \in (0, \infty)$.

In the absence of noise ($\eta_i = 1$), the exit date would be a fraction

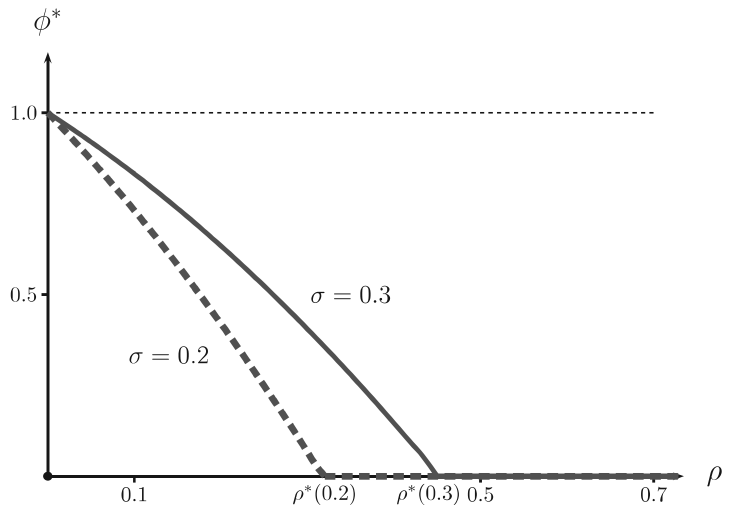
$$\phi_i \equiv \lambda_i/(1 + \lambda_i)$$

of the total game duration T. The actual exit date is a random fraction of the total duration T, obtained by assuming that λ_i is implemented with noise.

Proceeding as before, we compute the equilibrium target λ^*, or equivalently, the equilibrium fraction $\phi^* = \lambda^*/(1 + \lambda^*)$. That fraction captures the competitive pressure – when $\phi^* = 0$, these pressures are extreme, and when $\phi^* = 1$, they are completely absent. The figure on the next page shows how ϕ^* varies as a function ρ, for $\sigma = 0.2$ and $\sigma = 0.3$.

We see from the figure that as ρ decreases, the competitive pressure starts weakening at $\rho^*(\sigma)$ (as defined in the previous section). Because of the presence of an end date, the competitive pressures no longer fully vanish as soon as ρ falls below $\rho^*(\sigma)$, but only do so gradually as ρ decreases below $\rho^*(\sigma)$.

9. *Other payoff structures.* In the centipede game defined in Section 3 above, the loss from late exit (as compared to early exit) grows linearly over time. One might imagine interactions (see below) for which this loss remains constant



over the course of the interaction: the player who exits first gets a bonus G, while the other suffers a penalty L. Formally, letting $\Delta = G + L$, this yields the following payoff structure:

$$v_1(t_1, t_2) = \begin{cases} G + aS\min(t_1, t_2) & \text{if } t_1 < t_2 \\ G + aS\min(t_1, t_2) - \Delta & \text{if } t_2 < t_1 \end{cases}$$

We consider the noise structure defined above (see (19.2)) with a terminal date T large compared to Δ. For any target fraction φ, we now define the parameter $\gamma = T(1 - \phi)$, which represents the length of the end game.

As the target fraction ϕ decreases from 1, the dispersion in exit dates increases, and for γ large enough, the dispersion is large enough to provide players with incentives to delay exit. In equilibrium, exit dates are thus pushed toward the terminal date T in equilibrium, until some date $T - \gamma^*$, which we calculate next.

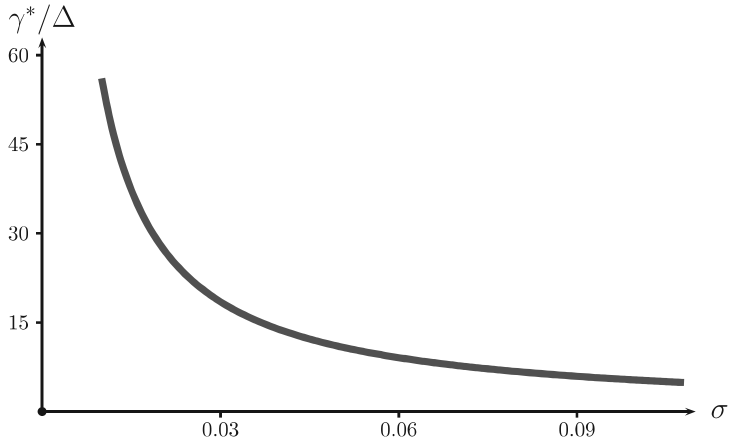
To quantify γ^* simply, we use the approximation $1 - \phi \simeq 1/\lambda$ and obtain:

$$t_i \simeq T - \gamma_i/\eta_i.$$

γ_i is a strategic variable for player i that represents the length of the end game, *as targeted by player i*. The term γ_i/η_i corresponds to a noisy implementation of that target. As before, we define player 1's expected gain $V_1(x, \gamma)$ when he targets $x\gamma$ and the other targets γ:

$$V_1(x, \gamma) = T - \gamma E \max(x/\eta_1, 1/\eta_2) + \Delta \Pr(x/\eta_1 > 1/\eta_2) - L.$$

This yields a simple expression for the equilibrium length γ^* of the end game. This length γ^* depends only on Δ and the distribution of the noise. It is independent of the length of the interaction, and grows linearly with Δ. The figure below illustrates how γ^*/Δ varies with the standard deviation σ.[14]



From the figure, we see that even very small noise can seriously suppress unraveling.

10. *Application to finite duration repeated games.* Consider a prisoner's dilemma game played in continuous time until a terminal date T. The payoff structure below describes the flow payoff to the players as a function of the actions currently played (either *C*ooperate or *D*efect):

	C	D
C	$1,1$	$-L,G$
D	$G,-L$	$0,0$

with $G > 1$ and $L > 0$. We assume that when a player defects, his defection is detected with certainty after a lapse of time of d, leading to mutual defection for the rest of the interaction. We are interested in the date t_i at which a player initiates defection. As in the centipede game, both players have a joint interest in choosing late defection dates, but each has a private incentive to defect before the other.

Formally, we look for an equilibrium in which players targets a defection date. The payoff that player 1 obtains as a function of defection dates t_1 and t_2 is:

$$v_1(t_1,t_2) = \begin{cases} \min(t_1,t_2) + G\min(t_2 - t_1, d) & \text{if } t_1 < t_2 \\ \min(t_1,t_2) - L\min(t_1 - t_2, d) & \text{if } t_2 < t_1 \end{cases}$$

[14] We have $\gamma^*/\Delta = \int \eta(h(\eta))^2 d\eta / \int (1/\eta)(h(\eta))(1 - H(\eta))d\eta$.

The temptation to defect before the other is thus bounded by $(G+L)d$, which makes this game akin to the centipede game examined in the previous section: when the target is implemented with noise, then, even with tiny noise, unraveling is limited. Additionally, the length of the end game (during which players defect) is independent of the duration (T) of the game, and it increases with $(G+L)d$. In other words, that length increases with the stakes (how much one gains from exiting first, how much one loses when the other exits first) and the duration d needed to detect the deviation. The exact characterization is left for future research.

11. *Comparison with a standard Bayesian model.* An alternative modeling strategy might add noise to model parameters, say the share a, with agents observing a noisy signal a_i of a, and assume that each chooses optimally an exit date as a function of the signal a_i received.[15] Our proposed modeling strategy differs in that we do not try to provide a rational justification for each possible exit date that the agent picks. We leave some aspects of decision making unmodeled, arguing only that the exit dates may be subject to shocks. Nevertheless, exit dates are not assumed to be arbitrary. They are driven by realized welfare comparisons on the agent's part, through the optimal choice of the target λ_i.

Said differently, Bayesian models try to endogenize the structure of mistakes that agents make, taking as given the structure of their signals, and assuming that agents can exploit arbitrarily finely these signals, as though they knew precisely the process that generated the signals. In contrast, we make no attempt to endogenize the structure of mistakes, and instead focus on the implications that plausible noise structures have on strategic behavior. We do so for several reasons. We are skeptical of what can fruitfully be endogenized or rationalized as far as mistakes are concerned: the source of mistakes presumably lies in what agents observe or perceive, and we seldom have access to these perceptions. Furthermore, the agents are not likely to have access to the details of the process that generates their perceptions. Also, mistakes need not always be rationalized within the model: one may construct parsimonious models that do not adhere strictly to the Bayesian methodology and, nevertheless, explore strategic consequences of noisy perceptions.

There is another advantage of our approach. One message from this chapter is that the form of the noise may largely shape a model's prediction. One may, of course, rely on perception or psychometric studies to uncover which are the most relevant. But by considering various forms of noise, one may discover which forms lead to predictions that are consistent with our intuitions (or with experimental data). In Section 9, the prediction is that the equilibrium length of the end game is an increasing function of the stake Δ. With additive

[15] Brunnermeier and Morgan (2010), for example, examine a similar game, assuming that the date at which the cake starts growing above 0 is random, and that agents observe noisy signals of that starting date.

noise, we would obtain a threshold Δ^* above which the game unravels back to the start of the game, and below which players stay in for most of the game. We find the prediction obtained from the multiplicative noise structure appealing.

Further Comments

The centipede game was originally proposed by Rosenthal (1981). It provides a simple illustration of the strength of backward induction. The finitely repeated prisoner's dilemma and the chain store game (Selten (1978)) are two other examples of games in which backward induction plays a key role in generating somewhat implausible predictions.

In these three games, noise in decision making reduces the strength of backward induction, and generates predictions more in accord with experimental evidence and intuition. For example in the finitely repeated prisoner's dilemma, one expects that subjects do not defect at all dates, and that their propensity to remain cooperative depends on the payoff structure considered (the benefits from mutual cooperation, the gain from defecting, and the loss from being cooperative against an opponent who defects), in the same way that we expect large exit dates when a is close to $1/2$ *or* ρ *close to* 0.[16]

Rosenthal (1981) explores the fragility of backward induction arguments to the introduction of noise in decision making. In justifying such departures from "full rationality," he writes:

> *"Although a player may not consciously intend to randomize [...] it may be that for some reason his intention is not realized [...] If players with preceding moves realize that this may be the case, they would be wise to consider it when making their choices."*

Rosenthal thus calls for introducing noisy decisions, but decisions still driven by self-interest. Implementing these two considerations simultaneously is a challenge. How should one model mistakes? How should one model the force toward self-interest? If one sees a move that seems unexpected, should this be viewed as evidence that further unexpected moves are to be expected, or should this be viewed as a one-time mistake, inconsequential to figure the benefits of each alternative available?

These considerations lie at the heart of the reputation literature,[17] *which proposed (i) a special structure on mistakes, assuming that with some small probability behavior is determined exogenously (say, an agent never exits) and (ii) a special way of modeling the force toward self-interest: if the agent's behavior is not determined exogenously, then he must be "fully rational" – that is, exploiting to a maximal degree the structure of the model.*

Rosenthal follows a different path. He proposed that the decision is subject to a random component, drawn independently at each date, but driven by the

[16] See Embrey, Fréchette and Yuksel (2016) for experimental evidence.
[17] See Chapter 13 and Milgrom and Roberts (1982) for example.

utility differences between the alternatives available at that date, where the utility is computed taking into account the mistakes that the agent and others might make in the future.[18]

Our modeling strategy falls in neither category. We do not examine incentives at each date, calculating what each believes based on hypothetical presumptions about what could have led to the other not exiting so far; nor do we presume the existence of various types of players, some of whom are crazy or altruistic, while others are perfectly rational; nor do we tailor the magnitude of mistakes made in each round to the correct value comparisons that would be performed in that round, as proposed by Rosenthal. Our family of noisy decision rules directly embodies the two concerns above, i.e., the need to introduce some randomness in decision making and to introduce self-interest.

Finally, this chapter (as well as the other chapters in which we consider noisy strategies) is related to the literature that introduces noise to capture evolutionary pressure (Foster and Young (1990)), or incomplete learning (Roth and Erev (1995)).[19] *One interpretation of this chapter is that equilibrium in noisy strategies is a shortcut to predict behavior in interactions where learning remains incomplete.*

Further research. One could examine various extensions, varying the number of players involved, and relaxing the assumption that the termination date is known with certainty. One could, for example, define the exit function $\mathbf{t}(\lambda_i, \eta_i, T_i)$ where T_i is a noisy estimate of the terminal date, and analyze the effect of all players missing the true termination date (either the pie is lost, or shared equally). The analysis could provide insight about the effect of the number of players on the duration of the end game in the prisoner's dilemma. It might also provide insight about the Dutch auction and the effect of setting a public reserve price: with several objects sold, this game has the structure of a centipede game,[20] and the reserve price, if public, plays the role of a termination date.[21]

[18] McKelvey and Palfrey (1995,1998) elaborate on this idea to define a quantal response equilibrium. In dynamic games, the logic of quantal response applied at each node of the game (rather than ex ante over all dynamic strategies) would be to introduce *independent* noise at each date.

[19] See also Gale et al. (1993).

[20] The Dutch auction is a descending-price auction in which the bidder who exits first gets the object. If there are K objects sold to N bidders, each of whom has the same valuation v for the object, and if, as soon as one bidder exits, others follow immediately, the first mover gets one object with certainty, while the others get it with probability $\rho = (K-1)/(N-1)$. This game has the structure of a centipede game (the pie increases linearly once the price reaches v).

[21] If the reserve is not public, players may only have a noisy estimate of that termination date, and, therefore, may risk losing the pie altogether. Other possible extensions include: (i) examining the incentive to exit after the first exit (in the spirit of Bulow and Klemperer (1994)); and (ii) dealing with the case that exit by others is learned with a random delay.

References

Bulow, J. and Klemperer, P. (1994). Rational frenzies and crashes. *Journal of Political Economy*, 102(1), 1–23.

Brunnermeier, M. K. and Morgan, J. (2010). Clock games: theory and experiments. *Games and Economic Behavior*, 68, 532–550.

Carlsson, H. (1991). A bargaining model where parties make errors. *Econometrica*, 59(5), 1487–1496.

Embrey, M., Fréchette, G. and Yuksel, S. (2016). Cooperation in the finitely repeated prisoner's dilemma, working paper.

Foster, D. and Young, H. P. (1990). Stochastic evolutionary game dynamics. *Theoretical Population Biology*, 38(2), 219–232.

Gale, J., Binmore, K. G., and Samuelson, J. (1995). Learning to be imperfect: the ultimatum game. *Games and Economic Behavior*, 8, 56–90.

McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria in normal form games. *Games and Economic Behavior*, 7, 6–38.

McKelvey, R. D. and Palfrey, T. (1998). Quantal response equilibria for extensive form games. *Experimental Economics*, 1, 9–41.

Milgrom, P. and Roberts, J. (1982). Predation, reputation, and entry deterrence. *Journal of Economic Theory*, 27, 280–312.

Nash, J. (1953). Two person cooperative games. *Econometrica*, 21, 128–140.

Rosenthal, R. (1981). Games of perfect information, predatory pricing, and the chain store paradox. *Journal of Economic Theory*, 25, 92–100.

Roth, A. and Erev, I. (1995). Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8(1), 164–212.

Selten, R. (1978). The chain-store paradox. *Theory and Decision*, 9, 127–159.