Endogenous Barriers to Learning

Olivier Compte[*][†]

March 1st 2023

Abstract

Motivated by the idea that lack of experience is a source of errors but that experience should reduce them, we model agents' behavior using a stochastic choice model, *leaving endogenous the accuracy of their choice*. In some games, increased accuracy is conducive to unstable best-response dynamics. We define the barrier to learning as the minimum level of noise which keeps the best-response dynamic stable. Using logit Quantal Response, this defines a *limitQR equilibrium*. We apply the concept to centipede, travelers' dilemma, and 11-20 money-request games and to first-price and all-pay auctions, and discuss the role of strategy restrictions in reducing or amplifying barriers to learning.

Keywords: Learning, Bounded Rationality, Stochastic choice

JEL Classification Codes: C72, D83, D90

[*]Affiliation: Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris, and Ecole des Ponts Paris Tech olivier.compte@gmail.com

[†]These ideas where partly developped working on blotto games with Alessandra Casella. I thank her, as well as Evan Friedman and Philippe Jehiel for their comments.

Contents

1. Introduction

Playing a Nash equilibrium requires that each player comes to play a best response to others' behavior. One original motivation for studying QRE rather than exact equilibria is that learning to play a best response may be hard to accomplish, or require enough experience. As McKelvey and Palfrey [1995] put it in their seminal work, "*as a player gains experience playing a particular game and makes repeated observations about the actual payoffs received from various action choices, he/she can be expected to make more precise estimates of the expected payoffs from different strategies.*"

This paper pursues this line of thought analyzing games in which there are natural obstacles to learning from experience. For one, we observe by example that increased accuracy in payoff evaluations may undermine the *stability of best-response dynamics* (see below), thus potentially fueling strategic uncertainty and hurting in return the accuracy of payoff estimates. Stability of the best-response dynamic will be viewed as a requirement which, if not met, puts an endogenous limit on the accuracy of payoff evaluations, justifying the notion of *limit QR* Equilibrium we introduce. We illustrate this notion with several well-studied games: the centipede game (Rosenthal [1981]), the traveller's dilemma (Basu [1994]) and the 11-20 money request game (Arad and Rubinstein [2012]).

Second, we consider games for which the richness of the strategy set is an obstacle to learning. Examples of such games include Bayesian games with a continuum of types and actions, like auctions.[1] In these games, from an *ex ante* perspective, the strategy set is a space of functions and performance evaluations of each feasible strategy is clearly out of reach. A possible path consists in exogenously restricting the set of strategies that players actively consider or explore, i.e., strategies for which performance-related data is gathered.[2] Given the restriction, performance evaluations are easier, but, depending on the strategy restriction assumed, the characteristics of the environment, or the auction format considered, the stability of best-response dynamics may or may not be an issue, with different consequences regarding our behavioral predictions. We illustrate this with the analysis of first-price and all-pay auctions.

Stochastic best response dynamics. The central assumption of our paper is in defining what constitutes a barrier to learning. Given a set of alternatives A, we start from a stochastic choice model h^β mapping each distribution f over consequences to a random choice, i.e., a probability distribution over the alternatives $h^\beta(f) \in \Delta(A)$, with β parameterizing how close this is from best responding. Next, for a given game g, each distribution over action profiles p induces a distribution

[1] Other games with rich strategy space would include repeated games, not covered here.

[2] This is in the spirit of Simon [1955] consideration set. This is also the path followed by Compte [2001], in auctions and, more generally, Compte and Postlewaite [2018], to model moderately sophisticated agents. In the context of games analyzed through the quantal response lense, the focus on ex ante performance along with a family of a priori defined ("cutoff")-strategies has been explored by Carrillo and Palfrey [2009]. To our knowledge, other papers in this literature adopt an interim perspective, defining quantal response *conditional on each type*, or agent-QRE (McKelvey and Palfrey [1998]). See also Goeree, Holt, and Palfrey [2002].

$f_i^g(p)$ over consequences that each play i is confronted to, from which we can construct for each player a *stochastic best-response* $\phi_i^\beta(p) = h^\beta(f_i^g(p))$, hence eventually, a function ϕ_β that maps the set of mixed-actions profiles to itself. As $\beta \nearrow \infty$, ϕ_β approaches the standard best-response over pure strategies.

Having defined the stochastic best-response function ϕ_β, we consider its long-run properties. If the iterations of ϕ_β converge to some p_β^* and if ϕ_β is locally stable at p_β^*, we consider that there is scope for improving the precision of the player's response (i.e., scope for increasing β) up to the point where, for some higher β, iterations of ϕ_β ceases to be stable (and thus fail to converge to a locally stable and stationary distribution p_β^*). We define the limit point β^* of this fictitious process as our *barrier to learning*, and we are interested in characterizing the limit distribution $p_{\beta^*}^*$.

Quantal response and Limit QRE. This barrier to learning can be defined for any stochastic choice model one likes, either belief-based or performance-based, for example depending on what consequences are gathered by players (actions of others or payoffs). We focus here on performance-based stochastic choice. Furthermore, for the sake of illustration, we shall focus on the (logit) Quantal Response model where stochastic choice depends on *expected value differences* between alternatives. By construction, when iterations of ϕ_β converge, the distribution obtained is a QR Equilibrium. We shall call *limitQRE* the distribution $p_{\beta^*}^*$ obtained at the stability frontier β^*.

Non-gradual vs. gradual adaptation. A controversial aspect of our hypothesis is that we examine iterations of the stochastic best response function hence the properties of the sequence of mixed strategies

$$p^{(n)} = \phi_\beta(p^{(n-1)}) = \phi_\beta^{(n)}(p^{(0)}), \tag{1}$$

rather than the smooth evolution that would obtain under a smooth/continuous time version of the best-response dynamic, or any replicator-like dynamic taking $\phi_\beta(p)$ as an input to reinforce the prevalence of the better fitted (mixed-)actions. These continuous-time processes over mixed strategies have nicer convergence-to-Nash/QRE properties, in particular for the smooth stochastic choice function ϕ_β that we shall consider.[3] Why then study (1)?

We see two reasons for doing so. First, the smooth evolution requires some synchronicity in the speeds of adjustments on each side (which of course is a fine hypothesis for symmetric games with a single population): synchronicity would get us closer to (1).

Second, and more importantly, we see it as a methodological tool to discriminate between games and assess the degree to which learning is difficult. If we assume a learning process that always converge to Nash in the long-run, such a discrimination cannot be made, except possibly, as in Erev and Roth (1995,1998), through the comparison of speed-of-learning across. We see the convergence of $\phi_\beta^{(n)}$ as a simple all-purpose vehicle for studying barriers to learning, vehicle that we apply across

[3] See for example Hopkins [1999, 2002] and Turocy [2005].

different well-known games.[4]

Nevertheless, the method proposed could allow for a smoother adaptive rule, in the spirit of a population that gradually adapts to its environment, and one could define, for a fixed α that one finds more appropriate, the sequence:

$$p^{(n)} = \phi_{\alpha,\beta}(p^{(n-1)}) \equiv (1-\alpha)p^{(n-1)} + \alpha\phi_\beta(p^{(n-1)}) \tag{2}$$

and check the stability and convergence of $\phi^{(n)}_{\alpha,\beta}$ instead, as β rises. The drawback of course is that if α is set very low, convergence to Nash will obtain in many games, and our ability to discriminate between games on that stability-of-learning dimension will be reduced. We briefly explore this path for the centipede game setting $\alpha = 0.1$.

Other sources of noise and games in targets. Most of the paper is concerned with determining endogenous barriers to learning, hence as a result, *endogenous noise* in how players play the game. Given our modelling of stochastic choice – which is performance-based, the shape of the noise is structured by expected value differences between alternatives, at $p^*_{\beta^*}$.

A complementary perspective is that there may be exogenous sources of noise in the game, which are not necessarily driven by expected value considerations. In the spirit of Carlsson [1991] and the global game approach (Carlsson and van Damme [1993]), an alternative way to introduce errors consists in assuming that when a player targets the alternative κ, she implements it with noise, also selecting "nearby" alternatives with positive probability.[5,6] This defines a *game over targets*, to which we can apply our limitQRE notion as well.

When studying *limitQRE of the game over targets*, the distribution over alternatives obtained is *partially driven* by the exogenous trembles (given the targets chosen) and *partially driven* by the value-estimation errors associated with the expected performance of targets. We shall discuss how trembles modify (though not always) the convergence of $\phi^{(\infty)}_\beta$, thus raising β^*, though not necessarily modifying substantially the induced distribution over alternatives.

Related Literature.

The question of convergence to equilibrium is central in economics: when should we expect players to play a Nash equilibrium and if convergence occurs, which equilibrium is selected. The examples we analyze have a unique equilibrium, so this paper is concerned with the convergence issue, rather the selection issue.

[4] This echoes Erev and Roth [1998] (page 887), who propose to study "*learning in games using simple general models*" and see "*how particular games and economic environments influence the dynamic of learning*".

[5] We shall come back to what "nearby" can mean.

[6] Trembling dates back to Nash [1953]. They are generally used as a selection device, with *vanishing trembles.* Here we are interested in "trembles" that need not be of small magnitude.

As already mentioned, since the seminal work of Maynard Smith [1972], convergence to Nash has been examined through the lens of slow-moving processes,[7] recognizing that fast adaptation would generally create instability of the learning dynamics. The general take from the literature is that for perturbed variations of replicator dynamics (that incorporate enough experimentation) or stochastic fictitious play, convergence is not an issue. We examine a drastically different hypothesis where, at the current precision β, behavior adjusts *fully* (and not partially) to the current behavior.[8]

The centipede game, the traveler's dilemma and the 11-20 money request game have led to extensive theory and experimental work. Let us mention McKelvey and Palfrey [1992], Nagel and Tang [1998], Capra, Goeree, Gomez, and Holt [1999] and **?**. At the heart of the explanation for departures from Nash equilibrium are frictions coming from evolutionary pressures, imperfect learning or various forms of stochastic choice, adequately calibrated for each variant to explain the data observed. Given that our endogenous limit to learning in these games embodies significant behavioral noise, the predictions we obtain will necessarily echo the intuition gathered from these papers. Our contribution is in endogenizing a *minimal level* of noise, simply characterized by our limit β^*, that must show up in these games from *learning-stability considerations only.* Obviously, independently of stability considerations, other relevant sources of noise arising from limited data processing, computation errors, exogenous trembling, payoff uncertainty or misspecified beliefs maybe relevant in explaining the data. However, our analysis of the traveler's dilemma and the 11-20 money request game, where we combine exogenous trembles and endogenized quantal response, suggests some robustness of our analysis: while these exogenous sources actually help learning, the resulting combined noise remains relatively stable across these variations examined.

We also study first-price and all-pay auctions with continuum of types and actions. Anderson, Goeree, and Holt [2002] examines a family of games with a continuum of actions (including a continuous version of the traveler's dilemma). Auctions have been examined through the QRE lens in Anderson, Goeree, and Holt [1998], Goeree et al. [2002] and Camerer, Nunnari, and Palfrey [2016]. Anderson et al. [1998] studies an all-pay auction with no types. Goeree et al. [2002] considers a first-price auction with a small number of types and they study the agent-based version of QRE where stochastic play is defined independently for each type. Camerer et al. [2016] also study an agent-based version of QRE. In contrast, we define a grid of a priori given (linear) strategies and next derive the limitQRE of the induced game. We note the sharp contrast between first price and all-pay. We also note that instability is an issue for the all-pay when values are not dispersed enough, and that exogenous trembling does not improve the stability, unlike the other games we examined.

[7] Or processes that become slow over time as in fictitious play.

[8] We shall nevertheless illustrate the case of partial adaptation (in the centipede).

2. Definitions

Quantal response functions and equilibrium. We specialize the general best-response map ϕ_β presented in the introduction to stochastic choice models h_β that are based on *payoffs*, and actually, following the Quantal Response route, expected payoffs.[9]

Formally, consider first a decision problem with K alternatives $k \in A$. Call u^k the expected payoff associated with k, $\overline{u}$ the maximum payoff and $u = (u^k)_{k \in A}$ the vector of expected payoffs. With fully accurate estimates of consequences, the decision maker would chose an alternative achieving $\overline{u}$. With less precise estimates or lesser experience, other actions may be played with positive probability. Quantal Response assumes that the frequency p^k with which k is played depends on the vector of payoffs u, according to some *quantal response function* parameterized by a precision parameter β

$$p = h^\beta(u)$$

In simulations to come, we will further restrict attention to the classic *logit formulation* where the best performing strategy gets most weight and each p^k is proportional to $\exp -\beta(\overline{u} - u^k)$. We explore other functional forms capturing a notion of satisficing in the Appendix.

Turning to games where each chooses $k_i \in A_i$, we let $p_{-i} \in \Delta(A_{-i})$ denote the distribution over strategies used by other players, $v_i(p_{-i})$ the vector of expected payoffs associated with each alternative $k_i \in A_i$. Finally, for any p, we let $v(p) = (v_i(p_{-i}))_i$ denote the profile of payoff vectors induced by p. A game is characterized by its value function v.

A *Quantal Response Equilibrium* is then defined as the solution of the system

$$p_i = h^\beta(v_i(p_{-i})) \text{ for all } i$$

or, in vectorial form, as a solution of

$$p = \phi_\beta(p) \tag{3}$$

where $\phi_\beta = (\phi_i^\beta)_i$ and $\phi_i^\beta(p) = h^\beta(v_i(p_{-i}))$.

As explained in the Introduction, one interpretation of quantal response equilibria is that they correspond to a static short-cut for modelling players with limited experience: out of some unmodelled learning process, behavior has converged to p^*, and the justification for p^* comes from the fact that experience remaining limited (or learning remaining incomplete), the assessment of which strategy is the best performing one is subject to errors. β is an accuracy parameter meant

[9] More general versions could be proposed, based on the distribution of payoff differences between alternatives, with errors stemming from limits on sample size, of expected size β. This could also be done in the spirit of Osborne and Rubinstein [1998], with each alternative sampled the same number of times, or as in Danenberg and Spiegler [2022], with frequently used actions possibly creating a sampling bias.

These versions would be better at capturing how the variability of consequences may impair the performance comparison between alternatives. We could then expect the comparison of two sure alternatives to be done without errors, even if payoffs are nearby.

to characterize experience, while h_β characterizes the relationship between expected payoffs and stochastic choice for each player.

Beyond this motivation, the general view is also that, given convergence, absent exogenous limitations on experience accumulation, and *so long as play remains stationary*, the performance evaluations of each action should become more accurate, i.e., β should rise, and behavior should eventually approximate Nash Equilibrium play. This paper is about qualifying that view, suggesting that, in some games, there may be endogenous limits to learning.

Best response dynamics, convergence and limit QRE.

We are interested in the ability of players to come to play a Nash or QR equilibrium from experience. Many learning processes could qualify in this endeavor. But we are not so much interested in designing a dynamic process that "works", but whether convergence is somewhat robust. To this effect, we shall focus on a specific dynamic process that generally does not have good convergence properties when β is large, (stochastic) *best-response dynamics*, which means that starting from $p^{(0)}$, we consider iterations[10]

$$p^{(n)} = \phi_\beta(p^{(n-1)}) = \phi_\beta^{(n)}(p^{(0)}).$$

If the process converges, then the limit behavior p^*_β solves (3) and is a QR equilibrium.

Our hypothesis is that *so long as play remains stationary*, there are evolutionary pressures towards more precise performance evaluations, i.e. a raise of β. We thus envision a global dynamic process where, so long as convergence obtains, β gradually rises, until the point β^* where a further rise by a small increment would undermine convergence. Such a rise would likely make play non-stationary, hurting the precision of performance evaluations (so in effect diminishing β).

We see β^* as a endogenous barrier to precision and refer to the limit behavior $p^* \equiv p^*_{\beta^*}$ obtained under β^* as a limit QRE.

Intuitively, estimation errors (β) affect convergence of the best response dynamics. When errors are large, choice probabilities p are barely responsive to *any* variations in v.[11] In such cases, one expects the best response function to be globally stable (i.e. converging from any initial $p^{(0)}$). In contrast, when evaluations are close to accurate, choice probabilities can become very sensitive to (*some*) small variations in v, possibly making the best response dynamic *locally unstable*.[12]

Summarizing this discussion formally, we define (mostly for computational reasons in simulations) a small increment $\nu > 0$, and consider the joint process which starts from $\beta_0 = 0$ and the

[10] Alternatively, one could consider asyncronized best response dynamics where players are selected in random order to generate a change $p_i^{(n)} = \phi_i^\beta(p^{(n-1)})$, keeping $p_j^{(n)} = p_j^{(n-1)}$.

[11] For the logit response with parameter $\beta = 0$, the response is to play all strategies with same probability, independently of v.

[12] The process is locally stable at p if convergence to p obtains from any p' such that $|p' - p| \leq \varepsilon$ for some $\varepsilon > 0$.

uniform distribution over alternative p^0 (for which by $p^0 = \phi^{(\infty)}_{\beta_0}(p^0)$), and then sets, by induction on n,

$$\beta_n = \beta_{n-1} + \nu \text{ and } p^n = \phi^{(\infty)}_{\beta_n}(p^{n-1})$$

defined so long as $\phi^{(\infty)}_{\beta_n}$ *is well-defined and* p^n *locally stable.* We shall look at the largest β^* obtained under this process.[13]

Endogenous versus exogenous barriers to learning.

We interpret β^* as an endogenous barrier to learning, limiting the accuracy of value estimates, and inducing behavioral errors shaped by expected value differences between strategies.

In many economic environments, it is conceivable that there are other sources of errors are relevant, if not first order. Departures from the exact best response may stem from attempts by players to experiment, or may reflect the fact when the environment varies, as there may be discrepancies between current environment and the one that generated the data from which players learn. Or they may reflect some misperception in the environment.

We wish to understand whether and how these exogenous barriers to learning compete with the endogenous barrier we propose.

To this end, in the spirit of Nash [1953], Carlsson [1991] and the global game approach (Carlsson and van Damme [1993]), we shall introduce exogenous perturbations, assuming trembling to "nearby" strategies.

Formally we define a *game over targets* $\kappa_i \in K_i$. We assume that when a player targets the alternative κ_i, she actually implements it with noise, selecting *"nearby" alternatives* k_i with probability $\pi^q_{\kappa_i}(k_i)$, where q is a parameter characterizing dispersion ($q = 0$ means no errors).[14]

The shape of the distributions $\pi^q_{\kappa_i}$ and the notion of "*nearby strategies*" may depend on the application considered. It may require a natural topologically structure on the strategy space, or a structure that reflects a player's strategic thinking about the problem. For example, to illustrate, in games where there is a natural topology on the set of alternatives, we shall use:

$$\pi^q_{\kappa_i}(k_i) = q^{|k_i - \kappa_i|}\pi^q_{\kappa_i}(\kappa_i) \text{ with } \sum_{k_i \in K_i} \pi^q_{\kappa_i}(k_i) = 1 \quad (4)$$

where $|k_i - \kappa_i|$ is the distance between the target and the alternative. Whichever interpretation one favors for the source of noise, these trembles are exogenous and not structured by payoff differences between alternatives, and we shall see them as *exogenous barriers to learning.*

[13] A continuous version would consists in defining a path $p(\beta)$ having the property that each $p(\beta)$ is locally stable under $\phi^{(\infty)}_\beta$ and $p(\beta) = \phi^{(\infty)}_\beta(p(\beta'))$ for all $\beta' \in (\beta - \varepsilon, \beta)$ for some $\varepsilon > 0$.

Also note that we start for low β, so the origin $p^{(0)}$ is irrelevant. Furthermore, for all the game we consider here, $\phi^{(\infty)}_\beta$ will be globally convergent for $\beta < \beta^*$.

[14] The global game approach is generally interested in the selection induced by vanishing noise. We shall be interested here in cases where q is not small.

Given an *exogenous* noise structure characterized by q and a profile of targets $\kappa = (\kappa_i)_i$, one can compute the distribution π^q_κ over profiles k and the expected values

$$V_i^q(\kappa) = E[v_i(k)|\kappa, q]$$

This defines, *for each* q, a *game over targets*.[15] When $q = 0$, we are back to the original game. When $q > 0$, one can study the *limitQR equilibria of the game over targets.*

Whether we consider the original game defined over alternatives, or the perturbed game defined over targets, eventually, both approaches generate a distribution over alternatives which we are interested in comparing.

Methodologically, when studying limitQRE of the original game, we are endogenizing simultaneously the best performing *alternative* and the *distribution* over errors in selecting *alternatives* (based on expected value differences between alternatives).

When studying limitQRE of the perturbed game, we are endogenizing the best performing *target*, as well as the distribution over errors in selecting *targets* (based on expected value differences between targets), keeping exogenous the distributions over alternatives conditional on the target selected.

In other words, in the later case, the distribution over alternatives obtained is partially driven by the exogenous trembles (given the targets chosen) and partially driven by the value-estimation errors associated with expected performance of targets. We shall see that to some extent, introducing exogenous errors does not alter the induced limitQRE distribution over alternatives.

3. Applications

3.1. A centipede game

We consider the two-player game where each player i chooses an exit time $t_i \in \{1, ..., 100\}$, assuming that the joint benefit is $\underline{t} = \min(t_1, t_2)$ and the player who exits first gets a share $a > 1/2$ of the joint benefit, the other player getting the rest, assuming equal sharing when $t_1 = t_2$:

$$\begin{aligned} v_i(t_i, t_j) &= a \min(t_i, t_j) \text{ if } t_i < t_j \\ &= (1-a) \min(t_i, t_j) \text{ if } t_i > t_j \\ &= \frac{1}{2} \min(t_i, t_j) \text{ if } t_i = t_j \end{aligned}$$

Each player in this game thus has an incentive to slightly undercut his opponent (except possibly for small values of t).

[15] Note that the value $V_i^q(\kappa)$ is computed ex ante, taking into account the tremble that i herself is subject to. Player i is implicitly gathering information about the effect of selecting a particular target, and this evaluation thus takes into account own trembling.

In this game, there is an equilibrium force towards low values of t, and this force is more pronounced for larger values of a, and with sufficiently large β, the Quantal Response equilibrium would naturally concentrate most weight on low values of t. Nevertheless, there is an endogenous barrier β^* which limits this unravelling. Figure 1 below shows the limit QRE for different values of a, where the curves plot the equilibrium distributions over strategies:

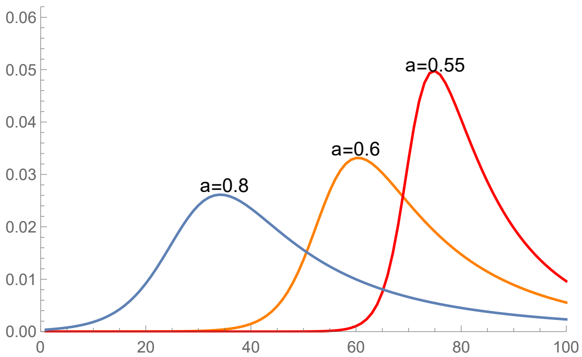

Figure 1: Limit QRE distributions

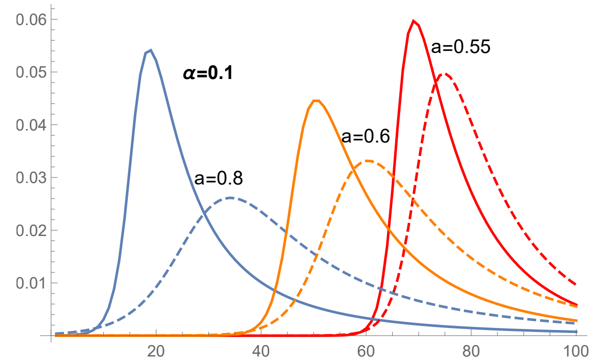

Figure 2: Gradual adj. $\alpha = 0.1$

As one expects, a higher value of a generates more competitive pressures. As a matter of fact, one can examine the quantal response equilibria *for a fixed* β, and observe that competitive pressures are stronger for higher a, leading to more unravelling. And competitive pressure are also higher for higher β's, as higher β gets one close to equilibrium. The stability frontier β^*_a however is decreasing in a,[16] which implies that the limitQRE outcomes across a's are close to another than what would be predicted by an analysis of QRE with fixed β.

Intuitively, the reason why β cannot rise to levels that would foster unravelling and low $t's$ is that whenever β starts being high enough to be conducive to such low t's, payoff differences across strategies become small, so many strategies including some with high $t's$ must get weight, fueling a rise in the incentives to choose high t's which in turn destabilize the best response dynamics.

Formally, for any given β (and p^0) and at any step of the best response dynamics, one can compute $p^{(n)} = \phi^{(n)}_\beta(p^0)$ as well as the best performing alternative $t^{(n)} = \arg\max v_i(t_i, p^{(n-1)})$. For small enough values of β, $p^{(n)}$ converges to p^*_β and so $t^{(n)}$ also converges to some t^*_β. For large enough β, however, convergence stops and $t^{(n)}$ cycles.

To illustrate, *set* $a = 0.7$. Convergence stops at $\beta^*_{0.7} = 0.3$ and the limit distribution has a mode at $t^*_{\beta^*} = 42$. Then for a large range of values of β above $\beta^*_{0.7}$, $t^{(n)}$ cycles.[17] Cycles have low amplitude when β is close to β^*, and the amplitude increases as β increases. For $\beta = 0.5$, the sequence is:

$$58, 52, 45, 39, 34, 29, 25, 58, 52...$$

For $\beta = 10$, the sequence takes almost all values between 60 down to 3 before jumping back to

[16] β^* respectively equal to 0.23, 0.47 and 0.79 (for $a = 0.8, 0.6$ and 0.55).

[17] Technically, there is a tiny range of $\beta > \beta^*_{0.7}$ where $p^{(n)}$ is no longer stable and yet $t^{(n)}$ still does not cycle.

60. And it takes $\overline{\beta} = 30$ to get again full unravelling (to $t^*_{\overline{\beta}} = 3$). In other words, over the range $\beta \in (0.3, 30)$, cycling prevails, suggesting a thick barrier to unravelling.

Gradual adaptation. Finally, we report how the limitQRE would be altered under a more gradual version of the best-response dynamics, defined in introduction (see (2)). We report above (see Figure 2) the distributions obtained for $\alpha = 0.1$ (the dashed curves corresponds to $\alpha = 1$). The gradual adjustement shifts all distributions to lower demands and raise all β^*'s, preserving the natural comparative statics and the ranking of β^*_a. In the Appendix, we also report how the distribution over claims would be modified under an alternative to logit.

As is well-known, the centipede game (Rosenthal [1981]) was designed precisely to illustrate how noise could greatly undermine the unravelling forces suggested by subgame perfect equilibrium analysis. The prevalence of cycling for large β is reminiscent of work that has examined the centipede game through an evolutionary lens (Cressman and Schlag [1998] and in particular Ponti [2000]). Our limitQRE analysis selects an (a-dependent) upperbound β^*_a on precision above which cycling prevails.

3.2. The Traveler's dilemma

We next consider a version of the traveler's dilemma, where two players report a claim $t_i \in \{180, .., 300\}$. When reports are t_i and t_j, they each get $\underline{t} = \min(t_1, t_2)$, but the player making a strictly lower report gets a bonus G, while the other gets a penalty L. Formally:

$$\begin{aligned} v_i(t_i, t_j) &= \min(t_i, t_j) + G \text{ if } t_i < t_j \\ &= \min(t_i, t_j) - L \text{ if } t_i > t_j \\ &= \min(t_i, t_j) \text{ if } t_i = t_j \end{aligned}$$

This game thus has a structure similar to the centipede game defined above: the pie grows linearly with $\underline{t}$, but the reward (respectively the penalty) for claiming a lower (respectively a higher) amount is constant, rather than linear and rising with $\underline{t}$.

Figure 3 reports limitQRE distributions for $L = G = 5$ (red), 10 (orange) and 15 (blue) and 50 (green) with lower rewards and penalties generating less unravelling. The thick orange distribution also corresponds to the distribution obtained when $G = 13$ and $L = 7$, illustrating that only the sum $G + L$ matters.[18]

Let us call $t^*_{G,L}$ the optimal claim at the limitQRE and $d^*_{G,L} = 300 - t^*_{G,L}$ the distance to the maximal feasible claim. This game has a finitely repeated prisoner's dilemma interpretation,[19]

[18] This is because, omitting the rare event where $t_1 = t_2$, $v(t_1) \simeq t_1 \Pr(t_2 > t_1) + G +$ $\Pr(t_2 < t_1)(E(t_2|t_2 < t_1) - (G + L))$, so the comparison between two alternative claims only depend on $G + L$.

[19] Assume 1 is the benefit from joint cooperation, g the extra gain from defecting while the other cooperates and ℓ the loss from cooperating while the other defects. If it takes a lapse of time δ before being detected, $G = \delta g$ and

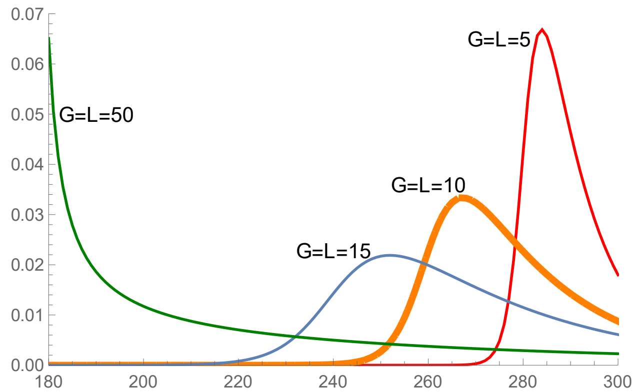

Figure 3: The Traveler's Dilemma

where the distance d can be interpreted as the length of the end game (during which defections occur).

For moderate values of L and G, low claims are essentially not played, and neither the lowerbound 180 or the upperbound 300 play a role in determining the limitQRE distance $d^*_{G,L}$: when the upperbound is changed by T, one may just shift the distributions over claims by T and incentives are not modified (to the extent that the shifted distributions do not hit the lower bound).

The comparative statics with respect to $\Delta \equiv G + L$ are analogous to those of the centipede game, with higher β (for a fixed Δ) and higher Δ (for a fixed β) both generating more unraveling. We also observe that in the range of bonus and penalties considered, β^* *decreases with* Δ, thus moderating the effect of Δ on the extent of unravelling, compared to what a standard QRE analysis with fixed β would deliver.

Intuitively, raising Δ increases incentives to lower claims, but it also raises *payoff differences* between alternative claims. This implies more concentration in claims, and this increased concentration destabilizes the best response dynamics, so β has to go down to stabilize it.[20] This also explains why, at limitQRE, the dispersion of claims increases when Δ rises.

Compte and Postlewaite [2018] (see Chapter 19) examines the target version of this game, where the targets d_i are implemented with *multiplicative noise*, observing that higher variance fosters cooperation, and that higher stakes Δ increase unravelling – and dispersion. But the increased dispersion there was obtained as a by-product of the multiplicative noise assumption. LimitQRE permits to simultaneously endogenize the magnitude and shape of the noise, given the logit response assumption, with similar qualitative predictions.

It is obviously not novel that noise – whether coming from imperfect information, stubbornness, or payoff or game-length uncertainty – helps foster cooperation (Kreps, Milgrom, Roberts, and

$L = d\ell$ approximately corresponds to the cumulated gain from taking advantage of the other (while not detected) and the loss of being taken advantage of (while not detecting it).

[20] For example, for $\beta^*_{5,5}$ obtained when $G = L = 5$, the best response dynamics obtained under $\beta^*_{5,5}$ but applied to $G = L = 9$ cycles from 283 down to 180 before jumping back up to 283.

Wilson [1982]). The contribution here is to suggest, based on difficulties in learning to play the game, an endogenous limit to precision – hence a limit to unraveling, given the logit response assumed.

Another implication of the analysis, worth noting in view of experimental work, is that these limits β^* vary depending on the game considered and also for different variations of the same game. So, with our interpretation of the logit parameter in mind, there may not be reasons to expect that the agent's accuracy of choices (measured by β) be similar across games or variations of the same game.

Adding exogenous noise. We now examine the game over targets for different level of noise. Formally, for any target τ, we define a distribution π^q_τ where $\pi^q_\tau(t)$ is the probability of selecting claim t when claim τ is targeted, and we assume:

$$\pi^q_\tau(\tau + m) = q^{|m|}\pi^q_\tau(\tau) \text{ for any } m \neq 0 \text{ such that } \tau + m \in \{180, .., 300\} \quad (5)$$

where $q > 0$ thus characterizes dispersion. In other words, when targeting τ, players tremble to "nearby" strategies applying a geometrically decaying factor as a function of the distance $|m|$ to the target. We think of these trembles as capturing exogenous reasons for the agent failing to choose the optimal claim.

Figure 4 plots three distributions $q = 0.6, 0.8$ and 0.9 for target $\tau = 260$. The dashed curve corresponds to the limitQR equilibrium previously obtained (thus in the game without trembles), when $L = G = 10$. As q increases from 0.6 to 0.9, the magnitude of the exogenous errors thus increases significantly.

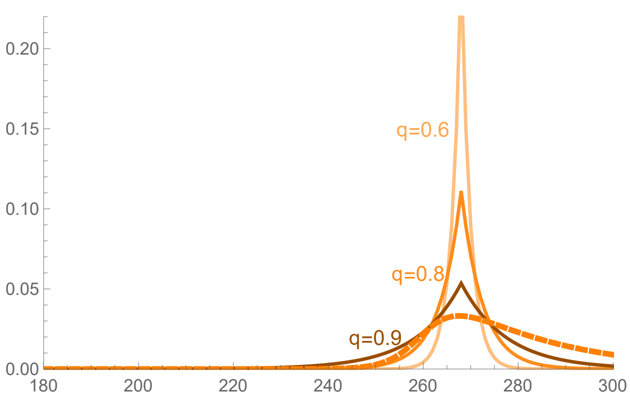

Figure 4: Cond. distributions over claims given target $\tau = 269$.

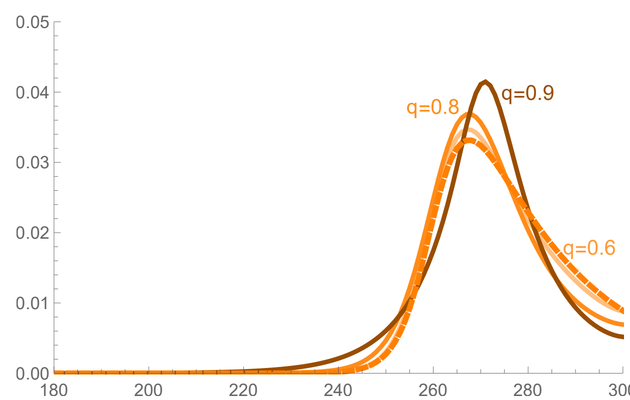

Figure 5: Induced distributions over claims in limitQRE

Each q defines a *game with trembles*, and we look for the limitQR equilibrium of these games. The limitQRE gives us a *distribution over targets*, and we then compute the *induced distribution over claims*, for different values of q. These distributions over claims thus combines exogenous uncertainty over claims given the target chosen, and endogenous uncertainty over targets. Figure 5 summarizes our finding (the dashed curves correspond to the limitQRE when $q = 0$ (no noise)).

Intuitively, the exogenous uncertainty improves the stability of the dynamic system, and the limit β^* rises with q. Nevertheless, when combining both sources of uncertainty, the effect is remarkably small, except for $q = 0.9$: in the latter case, endogenous uncertainty plays little role, targets become concentrated on few claims, and the limitQR equilibrium is mostly driven by exogenous noise.

3.3. The 11-20 Money Request game

We review the game proposed by Arad and Rubinstein [2012]. In its basic version, each player $i \in \{1, 2\}$ chooses a claim $t_i \in \{11, ..., 20\}$ and obtains it. In addition, player i obtains a bonus equal to 20 whenever $t_i = t_j - 1$, the payoffs are thus summarized by

$$\begin{aligned} v_i(t_i, t_j) &= t_i + 20 \text{ if } t_i = t_j - 1 \\ &= t_i \text{ otherwise} \end{aligned}$$

Like in previous games, each player thus has an incentive to undercut the other player. But there are two notable differences. Undercutting does not reduce the payoff that the other side can secure, as each can secure 20 by choosing $t_i = 20$. So standard equilibrium analysis cannot give rise to full unravelling to the lowest feasible claim and the Nash equilibrium must be in mixed strategy. Second, reducing one's claim (which is costly) delivers a bonus *only in the event where one undercuts by just one unit*. Since behavior is likely to incorporate some randomness, undercutting is potentially risky, and even more so for low claim levels (as one forgoes sure gains), unless these low levels claims are played with higher probability.

This explains why, in the Nash equilibrium of this game, among the claims t chosen with positive probability $p(t)$, lower claims are chosen with higher probability in equilibrium (i.e., $p(t)$ is downward sloping). As Arad and Rubinstein [2012] shows, the equilibrium outcome conflicts both with data and intuition (see Figure 6, Nash distribution plotted in dashed red, data in red), and the authors suggests that some form of level-k thinking may be at work.

We examine this game through the limit QRE lens. We also examine the game over targets for different level of noise. Formally, for a given $q > 0$ measuring noise and $\tau \in \{11, ..20\}$, we define the stochastic strategy π_τ^q where $\pi_\tau^q(t)$ is the probability of selecting claim t when τ is selected, and we assume:

$$\pi_\tau^q(\tau + m) = q^{|m|}\pi_\tau^q(\tau) \text{ for any } m \neq 0 \text{ such that } \tau + m \in \{11, ..20\} \tag{6}$$

For a given q, each pair of targets $\tau = (\tau_i, \tau_j)$ induces an expected gain $V_i^q(\tau)$ and a game over targets.

For each value of q, the limit QRE of this game is a distribution over targets, from which one obtains the distribution over the claims induced, given the trembles. We report below these limitQRE distribution over claims for different values of q (darker curves for higher values of q where q's are multiples of 0.1 from 0 to 0.4).

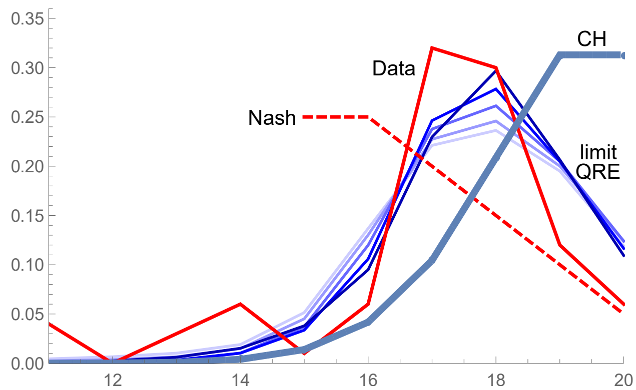

Figure 6: The 11-20 game

Note that while the exogenous trembles improve the stability of the best response functions ϕ_β, yielding significantly more concentrated distributions over targets as q rises, the consequence for the distribution over claims is mild. In all cases, the range 17-19 gets a weight approximately equal to 70%.

A similar robustness obtains when one adds a small fraction of stubborn/naive types that do not optimize and play $t = 20$.[21] The presence of this naive type induces (other) players to reduce the weight on the $\tau = 20$. Overall, the distributions look alike, except for a small shift towards higher claims (see Appendix).

Cognitive hierarchies.

By considering targets rather than actual claims, we effectively consider a game where players choose a strategy within a family of stochastic strategies, parameterized by a target $\tau \in \{11, 20\}$, where each "error" in selecting the actual claim has a probability that depends only on the distance $|t - \tau|$: they are homogenous in τ, only exploiting the natural topology on the strategy set, and not exploiting some other plausible causes of errors: in the spirit of the cognitive hierarchies proposed by Camerer, Ho, and Chong [2004], one could define a family for which the error probability also depends on τ, with higher cognitive levels (i.e., lower τ) generation move dispersion.

Specifically, let us associate to each claim t a cognitive level $k_t = 20 - t$. The target $\tau = 20$ is assumed to be implemented with no noise, but for any target $\tau < 20$, which corresponds to a *target cognitive level* $\kappa_\tau = 20 - \tau$, we assume that the player implements a cognitive level k (hence

[21] Let α be the probability of type $t = 20$. For a given α, the payoffs associated with τ become $V_i^\alpha(\tau) = (1 - \alpha)V_i(\tau) + \alpha E[v_i(t_i, 20)|\tau_i]$, which defines a new game.

a claim $t = 20 - k$) that follows a (truncated) Poisson distribution with parameter κ_τ, that is[22]

$$\frac{\pi_{\kappa_\tau}(k+1)}{\pi_{\kappa_\tau}(k)} = \frac{\kappa_\tau}{k+1} \text{ or equivalently, } \frac{\pi_\tau(t-1)}{\pi_\tau(t)} = \frac{\kappa_\tau}{k_{t-1}} \text{ and } \sum_t \pi_\tau(t) = 1 \tag{7}$$

So if $\tau = 19$ (and $\kappa = 1$), the induced distribution over claims has an expected cognitive level equal to $\kappa = 1$ with weights on 20 to 17 equal to 58%, 29%, 9%, 2% respectively. Endowed with a family of targets $\tau \in \{11, ..., 20\}$, one finds that the limit QRE is also the Nash equilibrium of game over targets cognitive levels, and it is concentrated on $\kappa^* = 2$ (or $\tau^* = 18$). Figure 6 plots (thick dashed blue curve–"CH") the induced distribution over claims.[23]

Intuitively, the larger dispersion in claims conditional on lower targets τ makes lower targets more risky, hence less attractive, explaining the distribution shifts towards on *higher claim levels*, compared to the previous case where trembles where homogenous across targets.

Discussion. Level-k thinking may certainly be at work in structuring how people play or think about the game and the strategies that they find relevant. Here, the cognitive levels are proxies for the strategies that player use (according to the relationship $k = 20 - t$). Level-k thinking may also be important in structuring the distribution over errors, with for example more dispersion for higher levels as in (7). We have examined the consequences of players targeting a cognitive level (or a claim) with more or less precision (i.e., different values of q). Our analysis of limit QRE suggests that, whether noise is endogenous (with β^* reflecting a frontier to experience accumulation) or only partially endogenous (i.e., looking at the game over targets), with enough experience playing the game, there should be forces that drive a large fraction of claims (about 70%) on the 17-19 range, with a mode on 18 under homogenous errors, and even higher mode for level-dependent errors. In both cases, expected claims are higher than that reported in Arad and Rubinstein [2012].

Said differently, level-k thinking and an exogenous bound on sophistication may be relevant in explaining how players with little experience play this game. We suggest that in this game, even with experience, unravelling towards lower claims is unlikely, suggesting that exogenous bounds on sophistication may not be a binding constraint.

3.4. Auctions

We consider an auction where each participant has a private value v_i for the object, drawn independently from the same distribution f. In simulations to come, we further assume that f is a lognormal distribution with variance σ.

[22] Note that this differs from Camerer et al. [2004] in the sense that players are not trying to assess the cognitive level of others: they target a cognitive level κ, but they implement it with noise, and *noise is larger for higher cognitive levels.*

[23] As a matter of fact, since the parameter of the Poisson distribution need not be an integer, players could be endowed with a finer grid of targets, say $\tau \in T = \{20 - 0.1j, j \in \{0, ...90\}\}$ and one can then solve for a Nash equilibrium and limitQRE in this finer grid. One finds that both coincide, with $\kappa^* = 1.9$ (and $\tau^* = 18.1$).

We think of players attempting to learn how to bid in this environment. The usual strategy space is a space of functions where all functions $b_i(v_i)$ should in principle be compared. In the spirit of Compte [2001] and Compte and Postlewaite [2018], we assume that players are more limited in their ability to gather experience and only compare a limited (here *finite*) number of strategies. Specifically, we consider a finite family or grid Γ of linear shading strategies[24]

$$b_i^{\lambda_i}(v_i) = \lambda_i v_i \text{ with } \lambda_i \in \Gamma = \{k_i/K, k_i \in \{0, ..., K\}\}.$$

For any profile of strategies $\lambda = (\lambda_i, \lambda_{-i})$, each value vector realization $v = (v_i)_i$ induces a vector of bids $b^\lambda(v) = (b_i^{\lambda_i}(v_i))_i$, from which we compute, *given the auction format considered*, the gain $u_i(v_i, b^\lambda(v))$. One can therefore compute the (ex ante) expected gains, where the expectation is taken over value vector realizations:

$$U_i(\lambda) = E_v u_i(v_i, b^\lambda(v))$$

Given $U_i(\lambda)$ and the grid Γ assumed, we adopt the limit QRE methodology to solve this game, and examine two auction formats, first-price and all-pay auctions. If β can grow without bound, this will imply that the limitQRE coincides with a Nash equilibrium of the constrained game (i.e., the game played on the grid). We wish to illustrate when and why such a convergence may or may not occur. We also wish to illustrate how the approach can be used to draw (at least qualitative) comparisons between formats.

3.4.1. First-Price Auctions.

We choose $K = 20$ and check for limit QRE. We also focus on two players ($n = 2$). Our first observation is that limit QRE coincides with the Nash outcome, with $\beta^* \nearrow \infty$ and the limit QRE distribution concentrated on single shading factor λ^*_σ. The best response dynamics is thus well-behaved for all values of β. As a matter of fact, the game defined without the grid restriction (or with an arbitrarily fine grid) also has a Nash equilibrium λ^{**}_σ, and the shading factor λ^*_σ obtained under the grid restrictions approximates λ^{**}_σ.[25] For the sake of comparison, we plot below the equilibrium bid and the standard Bayesian solution of this game (which is obtained by not putting restrictions on the strategy space). The Bayesian solution is not linear and the linear solution is *not* an approximation of it: the restriction induces less competition on most value realizations (the distribution is the dashed line). Figures 7 and 8 report the distribution over values for $\sigma = 0.5$, the Bayesian solution of the first price auction, and the equilibrium bid λ^*_σ under our restriction, for $\sigma = 0.2$ and $\sigma = 0.5$.[26]

Intuitively, under the restriction, players cannot adjust shading to the realized value: if they

[24] For simplicity, we consider identical grids across players, but this assumption is not necessary.

[25] Actual equilibrium shading in the game with arbitrarily fine grids are respectively $\lambda^{**}_{0.1} = 0.84$, $\lambda^{**}_{0.2} = 0.715$, $\lambda^{**}_{0.3} = 0.61$, $\lambda^{**}_{0.4} = 0.525$, and $\lambda^{**}_{0.5} = 0.454$.

[26] For $\sigma = 0.2$, the distribution over values has been scaled down (by a factor 0.7) so that it fits in the frame.

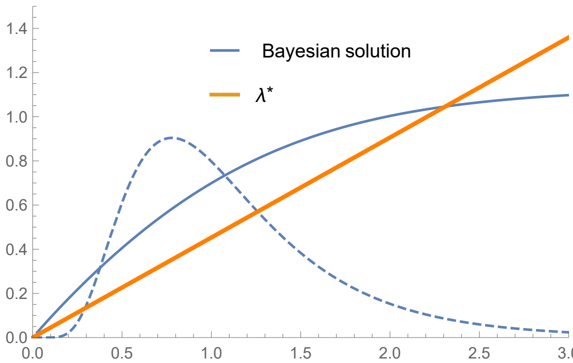

Figure 7: $\sigma = 0.5$

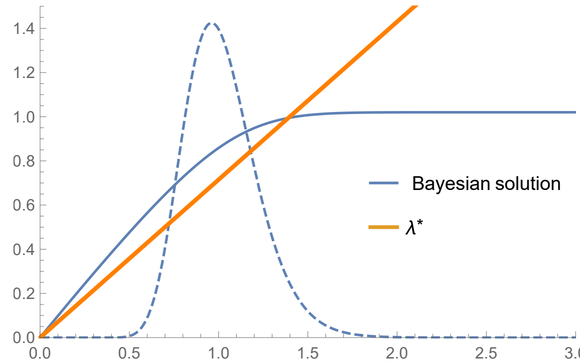

Figure 8: $\sigma = 0.2$

could and if they were able to infer from a, say, high, realization that they are likely to be the highest value bidder by a substantial margin, they would shade more: thus, for sufficiently high draws, there is less shading under the restriction. The opposite is true for middle range values.

The case of dispersion uncertainty

We wish to illustrate here that when the dispersion σ is uncertain, convergence of limit QRE is no longer ensured, thus limiting the usual efficiency properties of the first-price auction, in spite of the symmetric environment considered.

Assume that σ may take two values $\underline{\sigma}$ or $\overline{\sigma}$ with equal probability. Then $\underline{\sigma}$ calls for fierce competition (hence high λ likely generating low profits), while $\overline{\sigma}$ calls for weaker competition. When $\underline{\sigma}$ and $\overline{\sigma}$ are not too far apart, then there exists a smooth resolution of these conflicting incentives: there exists a pure Nash equilibrium of the constrained game, and the limitQRE continues to converge to it.

When $\underline{\sigma}$ is sufficiently low ($\underline{\sigma}$ and $\overline{\sigma}$ sufficiently far apart) however, such a smooth resolution does not exist.[27] Then β^* remains bounded, and limit QRE settles short of the (mixed) Nash.[28] Figure 9 gives the distribution over λ's induced by limit QRE for $(\underline{\sigma}, \overline{\sigma}) = (0.05, 0.5)$ (continuous curve):

Weights are no longer concentrated on a single λ but mostly spread over $\{0.6, .., 0.8\}$, these strategies cumulating 86% of the weights.

Competition in targets. To conclude on the analysis of first-price, we analyze the game where players choose target shading factors. In the spirit of the 11-20 money request game, we assume that rather than choosing $k \in \{0, .., K\}$, players each select a target $\kappa \in \{0, .., K\}$, with each κ inducing a probability distribution over actual shadings λ_k, i.e. a probability distribution $\{\pi^q_\kappa(k)\}_k$,

[27] The reason for non-existence of a pure strategy equilibrium is analogous to those arising in price competition with differentiated products (Caplin and Nalebuff [1986]), where some conditions on the demand function are needed to ensure existence of a pure strategy equilibrium.

[28] In the QRE equilibrium, weights on strategies are unequal, reflecting the fact that expected values differ across the strategy played.

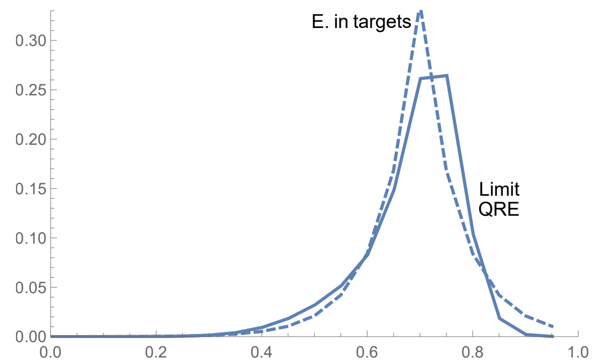

Figure 9: Limit QRE and Equilibrium in targets (under dispersion uncertainty)

where dispersion is parameterized as in (6) by q. As for the 11-20 money request game, for any given dispersion q, one can compute the limitQRE *of the game defined over targets*. With enough dispersion $q > q^*(= 0.17)$, the game over targets has a Nash equilibrium in pure strategies, κ_q^*, and the limitQRE of that game converges to it.

In Figure 9, we plot (dashed curve) the distribution over λ's induced by κ_q^* for $q = 0.5$. This illustrates the connection between limit QRE without trembles and the equilibrium in targets where trembles are high enough. These are two different (and possibly complementary) ways to introduce frictions in learning how to play a game, with, for this game, similar consequences on the distribution over λ's, for an appropriately chosen q.

3.4.2. All-pay Auctions.

With 2 players, limitQRE converges to a pure λ^* when $\sigma \geq 0.5$. This λ^* is also a Nash equilibrium of the game with restrictions. In contrast to the first-price auction however, convergence of limitQRE fails when dispersion of values is too low ($\sigma < 0.5$).[29] We plot in Figure 10 the limitQRE for $\sigma = 0.4$.

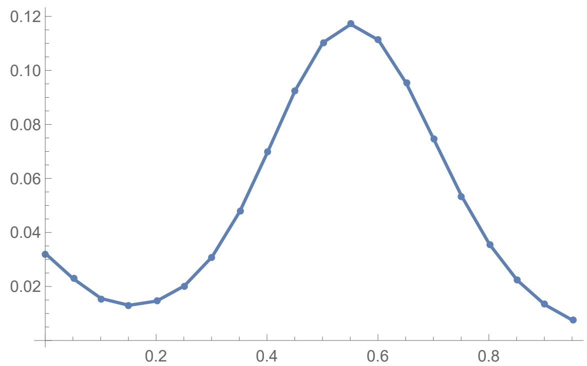

Figure 10: $\sigma = 0.2$

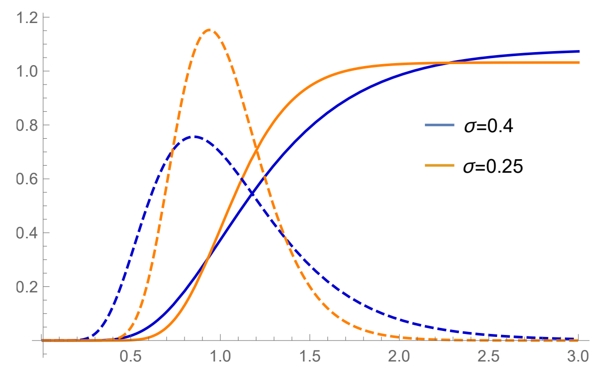

Figure 11: $\sigma = 0.2$

[29] With more than 2 players, even more dispersion in values would be needed to obtain convergence to a pure shading λ^*.

A notable feature of the limitQRE distribution is that it has **two modes**, with very low bids surfacing as reasonably good strategies. The reason is the following.

When σ is not too high, each player has an incentive to overbid the other for a large range of shading factors: upward competitive pressures are strong.[30] But when λ becomes too high, $\lambda = 0$ becomes a better option. In other words, with restrictions to linear strategies and $\sigma < 0.5$, the game has the structure of a Rock-Paper-Scissors game (a non-zero sum one with more than three options): if one looks at the best response dynamics over pure strategies (k_1, k_2 best response to k_1 etc...) when $\sigma = 0.4$, the sequence obtained eventually cycles over 8 strategies:

$$k = 0, 1, 3, 6, 9, 11, 12, 13, 0, 1...$$

So there cannot be pure strategy equilibria in this game.

It is instructive to compare with the Bayesian solution (where no strategy restrictions are made). The Bayesian solution would be in pure strategies,[31] and the equilibrium bid function b^{eq}_{σ} would be S-shaped, with more steepness at the inflection point for lower σ. (See Figure 11– which plots for b^{eq}_{σ} for $\sigma = 0.4$ (blue curve) and $\sigma = 0.25$ (orange curve) – the dashed curves give the shape of the distributions over values).[32] Under the Bayesian solution, players draw correct inference from getting a low draw of v, i.e., their chances of winning are slim hence bidding is low: players are not spending unnecessary resources on auctions they will most certainly lose.[33]

With linear strategies, such savings are not possible. So long as σ is large, this is not so much of an issue, because competitive pressures are not too strong and there are only two players: $\lambda = 0$ is not best response. With smaller σ, this becomes an issue, because competitive pressures make λ rise (up to 13/20 when $\sigma = 0.4$), at which point $\lambda = 0$ is a better strategy, which fuels the cycle.

Comparing with the 11-20 money request game. It is also instructive to compare with the 11-20 request game, which exhibits a similar (cycling) best response dynamic over pure strategies. One key difference between the two games appears when one examines the game over targets, where, as before, players target κ_i and tremble over nearby strategies, with q characterizing the dispersion of the trembles. In the money request game, there is a level of dispersion q that restores a pure strategy equilibrium in targets. *This is not the case in the all-pay auction.*

Intuitively, to get a Nash equilibrium in targets, one has to ensure that players do not have too much incentive to overbid. This can be done by increasing the dispersion of errors. The

[30] When σ is very small, slight overbidding is sufficient to win with very high probability. For larger σ, the gap has to be large enough to significantly modify the winning probability, and for large enough σ, this bid increase is too costly relative to the increase chance of winning.

[31] By revenue equivalence, equilibrium bids of the (unconstrained) Bayesian game would be such that the utility $u(v)$ of type v coincides with that obtained in a second price auction, so $u(v) = vF(v) - b(v) = F(v)(v - E[v_2|v_2 < v])$

[32] These dashed curves are scaled down on the vertical axis (by a factor 0.7) in order to fit in the same figure.

[33] But of course, this requires that the agent understand which v's are low. A possible inability to perform that task is what motivates the linear shading rules.

issue however is that the level of dispersion needed leads to negative expected profits. So if no participation is an option, or if the agent does not tremble when choosing the lowest target $\kappa = 0$, this cannot be an equilibrium.

These observations are consistent with the two-mode distribution of Figure 10, and they suggest a bidding behavior in all-pay auctions that differs substantially from first-price (not surprisingly), where players either behave as strong competitor (high κ or k, with the risk of facing another stronger competitor), or behave safe, aiming for an *easy win* (hoping that others' too will behave safe), and not participating. It also suggests that, with more potential participants, and since incentives to overbid are then stronger, the expected number of strong competitors will generally not exceed 2 (as it cannot be profitable to have more than 2 strong bidders). So unlike in first price where efficiency and seller's revenue will increase with competition, this is unlikely to be the case with all-pay.

Including more sophisticated strategies. When bidding in an all-pay auction, having some idea of how one's valuation compares to others' is key, so that one can afford to bid safely when one has a low valuation. Linear strategies induce lower bidding for low values, but the reduction is not strong enough to induce a pure strategy equilibrium (conducive to efficient allocations) when the dispersion is below 0.5.

To conclude this section, we wish to check whether the inclusion of $S-$shaped strategies which are better suited for all-pay would favor their selection under limitQRE. Specifically, let us define b^{eq}_{σ} as the *Bayesian solution* of the game where dispersion is σ. Figure 11 reports these $S-$shaped bid functions for $\sigma = 0.4$ and $\sigma = 0.25$). We consider strategy sets $\overline{\Gamma} = \Gamma \cup \{b^{eq}_{\sigma}\}$ that include the linear strategies from our previous grid, plus an additional $S-$shaped strategy b^{eq}_{σ}. We make the following observations:

(i) If $\sigma = 0.4$ and $\overline{\Gamma} = \Gamma \cup \{b^{eq}_{0.4}\}$, the limitQRE converges to $b^{eq}_{0.4}$.

(ii) if $\sigma = 0.3$ and $\overline{\Gamma} = \Gamma \cup \{b^{eq}_{0.3}\}$, the limitQRE puts a weight only equal to 9% on $b^{eq}_{0.3}$.

(iii) If $\sigma = 0.4$ and $\overline{\Gamma} = \Gamma \cup \{b^{eq}_{0.25}\}$, the limitQRE puts weight 29% on $b^{eq}_{0.25}$ and the induced distribution on Γ has a single mode.

(iv) If $\sigma = 0.3$ and $\overline{\Gamma} = \Gamma^{coarse} \cup \{b^{eq}_{0.3}\}$ where $\Gamma^{coarse} \in \{0, 0.3, 0.5, 0.7\}$, the limitQRE converges to $b^{eq}_{0.3}$.

We also report in Figures 12 and 13 the (unconditional) distributions induced on Γ for case (ii) and (iii), respectively. These observations show that indeed, the inclusion of the Bayesian solution may allow convergence to it (i). However convergence is not always guaranteed (by (ii)), and the inclusion of a strategy that is only "close to" the equilibrium, may not be sufficient to reach convergence (by (iii)). (iv) however indicates that a coarsening of the strategy set Γ may help the selection of the Bayesian solution.

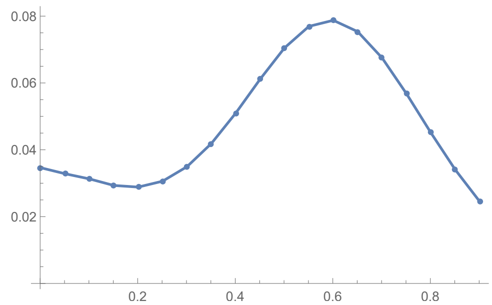

Figure 12: case(ii)

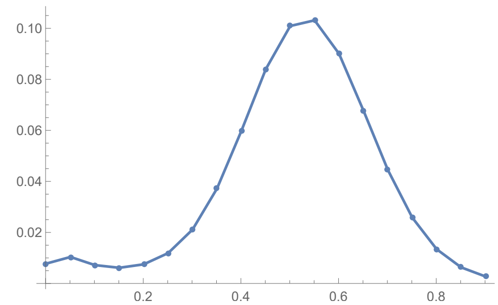

Figure 13: case(iii)

Let us give some intuition for these observations. Call b_S the $S-$shaped bid strategy and $\widetilde{b}_{\beta^*}$ the mixture induced on Γ at the limitQRE. In both cases (ii) and (iii), at the limitQRE, b_S is a best performing strategy, both against itself and against $\widetilde{b}_{\beta^*}$.[34] So if play was restricted to b_S and $\widetilde{b}_{\beta^*}$, one would expect that evolution eventually selects b_S, as more accurate estimates become available. This is not what happens however.

As β increases even slightly above β^*, the best response dynamics starts cycling over various distributions $\widetilde{b}$ and along the cycle, b_S *ceases to be the best performing strategy.* With Γ consisting of multiples of 0.1, and $\beta = 1.2\beta^*$, one gets the following cycle of length 14 among the best performing strategies:

$$0.4, 0.6, 0.7, 0, b_S, 0.5, 0.6, b_S, b_S, b_S, 0.5, 0.6, 0, b_S, 0.4...$$

with b_S occurring a fraction 5/14 of the time.

Here is below another symmetric game with a unique pure strategy equilibrium, for which the limitQRE does not converge to it:

	1	2	3	4
1	1,1	0,2	1,0	0.9,0
2	2,0	0,0	0,2	0,1
3	0,0	2,0	0,0	0,2
4	0,0.9	1,0	2,0	1,1

The restriction to the first three strategies in a game with a unique mixed strategy equilibrium, where the best response cycles $1- > 2- > 3- > 1$. In the presence of the fourth strategy, the game has a unique Nash, playing 4,[35] but the limitQRE stops short of it (at $\beta^* \simeq 2.5$) at distribution $\widetilde{\sigma} = (0.3, 0.21, 0.13, 0.36)$. As estimates get more precise, strategy 4 picks up more weight, but this improves the performance of strategy 1, creating a cycle of best performing strategies which, at $\beta = 3$ is as follows: $4, 4, 1, 2, 3, 4, 1, 2, 3, 4, 4, 1, 2, 4, 4, 1...$

[34] In case (ii), b_S is a best response against itself because it is an equilibrium. One can see from the weight distribution (which puts most weight on b_S) that b_S is also the best strategy against the limit QRE distribution. It turns out that is also better against the conditional distribution $\widetilde{b}_{\beta^*}$.

[35] Mixing over strategies 2 and 3 would require $p_1 = p_2$, but then the 4th strategy dominates others.

4. Conclusion

The issue of convergence to equilibrium is central in economics. The experimental data is a reminder that not all games are conducive to equilibrium play and the analysis of dynamic learning and evolutionary models have been useful in understanding observed departures from equilibrium.

In this spirit, our purpose has been to design a simple method for discriminating between games and making behavioral predictions that incorporate the difficulties that players might face in playing them.

What makes learning difficult? Our premise has been that non-stationary behavior is likely to hurt the ability of player's to find best responses to their environment, hence generate randomness into how people behave. Randomness however is good news: noise makes players less responsive to variations in others' behavior (or the environment), and this has the virtue of stabilizing long-run behavior.

Our method consists in finding the tipping point where player's understanding of their environment is low enough to induce stationary behavior, and where any further precision gathered on it would induce non-stationarities.

We believe that incorporating barriers to learning into the analysis of games is important because it highlights an issue often felt unexplored: the players' "consideration set", i.e., the set of strategies that players do compare. Simon [1955] argued that a key aspect of decision making is the subset of actions that agents actually consider (out of those a priori available), and that this subset depends on the extent of exploration.

Our analysis of auctions highlights the importance of this "consideration set". With a restriction to linear-shading rules, the analysis of competition under first-price is not qualitatively altered (though uncertainty about the dispersion of valuation may be an issue). The competition under all-pay however produces a strategic interaction akin to the Rock-Paper-Scissors game, for which learning how to play is a challenge.

From the agent's perspective, it is not obvious that faced with such a challenge, evolution will be conducive to the exploration of more sophisticated strategies, such as $S-$shaped strategies, as this may require a difficult search for the better locus of the inflexion point. Furthermore, even if reasonably adequate $S-$shaped strategies are added to the consideration set, their selection is not guaranteed, and may depend on the richness of the initial set (a finer grid hurting selection in our auction example).

In summary, we suggest that it is worth worrying about barriers to learning, in particular in games with a rich set of a priori feasible strategies, and study how strategy restrictions may either mitigate or amplify these difficulties.

References

Simon P. Anderson, Jacob K. Goeree, and Charles A. Holt. Rent seeking with bounded rationality: An analysis of the all-pay auction. *Journal of Political Economy*, 106(4):828–853, 1998.

Simon P. Anderson, Jacob K. Goeree, and Charles A. Holt. The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, 69(1):21–47, 2002.

Ayala Arad and Ariel Rubinstein. The 11-20 money request game: A level-k reasoning study. *The American Economic Review*, 102(7):3561–3573, 2012.

Kaushik Basu. The traveler's dilemma: Paradoxes of rationality in game theory. *The American Economic Review*, 84(2):391–395, 1994.

Colin Camerer, Salvatore Nunnari, and Thomas R. Palfrey. Quantal response and nonequilibrium beliefs explain overbidding in maximum-value auctions. *Games and Economic Behavior*, 98: 243–263, 2016. ISSN 0899-8256.

Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898, 2004.

Andrew S. Caplin and Barry J. Nalebuff. Multi-dimensional product differentiation and price competition. *Oxford Economic Papers*, 38(supp):129–145, 11 1986.

C. Monica Capra, Jacob K. Goeree, Rosario Gomez, and Charles A. Holt. Anomalous behavior in a traveler's dilemma? *The American Economic Review*, 89(3):678–690, 1999.

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

Hans Carlsson and Eric van Damme. Global games and equilibrium selection. *Econometrica*, 61 (5):989–1018, 1993.

Juan D. Carrillo and Thomas R. Palfrey. The compromise game: Two-sided adverse selection in the laboratory. *American Economic Journal: Microeconomics*, 1(1):151–181, 2009.

Olivier Compte. The winner's curse with independent private values, 2001.

Olivier Compte and Andrew Postlewaite. *Ignorance and Uncertainty*. Econometric Society Monographs. Cambridge University Press, 2018. doi: 10.1017/9781108379991.

R. Cressman and K.H. Schlag. The dynamic (in)stability of backwards induction. *Journal of Economic Theory*, 83(2):260–285, 1998.

Tuval Danenberg and Ran Spiegler. Representative sampling equilibrium, 2022.

Ido Erev and Alvin E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4): 848–881, 1998.

Jacob K. Goeree, Charles A. Holt, and Thomas R. Palfrey. Quantal response equilibrium and overbidding in private-value auctions. *Journal of Economic Theory*, 104(1):247–272, 2002. ISSN 0022-0531.

Ed Hopkins. A note on best response dynamics. *Games and Economic Behavior*, 29(1):138–150, 1999.

Ed Hopkins. Two competing models of how people learn in games. *Econometrica*, 70(6):2141–2166, 2002.

David Kreps, Paul Milgrom, John Roberts, and Robert Wilson. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27(2):245–252, 1982.

John Maynard Smith. *On evolution*. Edinburgh University Press, 1972.

Richard McKelvey and Thomas Palfrey. Quantal response equilibria for extensive form games. *Experimental Economics*, 1998.

Richard D. McKelvey and Thomas R. Palfrey. An experimental study of the centipede game. *Econometrica*, 60(4):803–836, 1992.

Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995.

Rosemarie Nagel and Fang Fang Tang. Experimental results on the centipede game in normal form: An investigation on learning. *Journal of Mathematical Psychology*, 42(2):356–384, 1998.

John Nash. Two-person cooperative games. *Econometrica*, 21(1):128–140, 1953.

Martin J. Osborne and Ariel Rubinstein. Games with procedurally rational players. *The American Economic Review*, 88(4):834–847, 1998.

Giovanni Ponti. Cycles of learning in the centipede game. *Games and Economic Behavior*, 30(1): 115–141, 2000.

Robert W Rosenthal. Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, 25(1):92–100, 1981.

Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69 (1):99–118, 1955.

Theodore L. Turocy. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior*, 51(2):243–263, 2005.

Appendix

An alternative to logit with application to the centipede game.

The logit formulation of QR assumes a particular structure on errors. We wish to illustrate the properties of limitQRE with other formulations that implicitly incorporate a notion of satisficing – which may be relevant from a learning perspective. Specifically, we shall fix $c \geq 1$ and examine the family $(h_c^\beta)_\beta$ whereby

$$q_k \text{ is proportional to } \exp -(\lambda(\overline{v} - v_k)^c$$

where $\overline{v}$ is the maximum expected gain. For $c = 1$, this corresponds to the family of logit functions. For c large, the family captures the idea that the agent attempts to discriminate between actions that yield a payoff *close enough* to the max, and actions that yield a payoff *sufficiently away from the maximum*. With $\overline{v}$ normalized to 1, the figure below shows that for $c = 5$ and $\lambda = 5$ (red curves) strategies that yield 90% of the maximum get approximately the same weight (and that weight is thus close to the maximum weight), while this weight quickly drops down for strategies yielding a fraction of the maximum lower than 90%. This is to be contrasted with the blue curve (drawn for $c = 1$ and $\lambda = 5$) which corresponds to the logit response (this response has a much thicker tail).

In summary, the size of c determines the sharpness of the drop for strategies that perform sufficiently below the best one, and β determines the threshold below which the drop occurs.

We illustrate the consequence of a higher c in the centipede. Fixing $a = 0.6$, we compute the limitQRE obtained for different value of c:

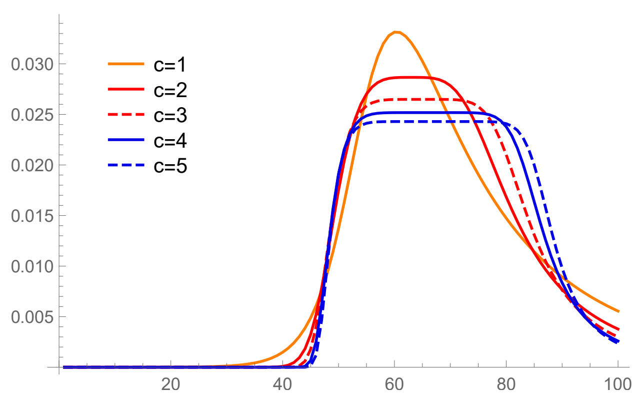

Figure 14: LimitQRE under an alternative to logit

Thus as one raises c, more strategies get played with comparable weight: for large c, the limit QRE comes close to a set-concept that determines a range of *good enough strategies*, which are then all played with comparable weight, as if discriminating between these strategies was too difficult for the agent.

The 11-20 money request game with stubborn types.

In Figure 15 below, we report (orange) the limitQRE distributions obtained when a player has a 5% chance of being a stubborn player claiming $k = 20$ (without trembling). We derive these distributions for the game where players target κ but tremble, with trembling probabilities characterized by $q = 0, 0.1$ etc...0.4. Lighter curves for $q = 0$. The figure reports the unconditional distributions (i.e., including the behavior of the stubborn type). We also recall (in blue) the distributions obtained in the absence of stubborn type. The figure suggests a small (but only tiny) shift towards higher claims.

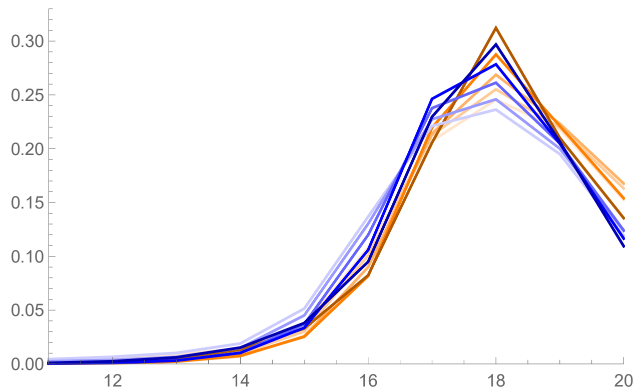

Figure 15: The 11-20 game with stubborn types