THE UNIVERSITY of ADELAIDE

School of Economics

Working Papers

ISSN 2203-6024

Exploration vs Exploitation, Impulse Balance Equilibrium, and a Specification Test for the El Farol Bar Problem

Alan Kirman, École des Hautes Études en Sciences Sociales, CAMS (Paris)

François Laisney, ZEW (Mannheim)

Paul Pezanis-Christou, School of Economics, University of Adelaide

Working Paper No. 2018-11
August 2018

Exploration *vs* Exploitation, Impulse Balance Equilibrium, and a specification test for the *El Farol* bar problem‡

ALAN KIRMAN[#] FRANÇOIS LAISNEY[§] PAUL PEZANIS-CHRISTOU[*]

August, 2018

Abstract: The paper reports on market-entry experiments that manipulate both payoff structures and payoff levels to assess two stationary models of behaviour: Exploration vs Exploitation (EvE, which is equivalent to Quantal Response Equilibrium) and Impulse Balance Equilibrium (IBE). These models explain the data equally well in terms of goodness-of-fit whenever the observed probability of entry is less than the symmetric Nash equilibrium prediction; otherwise IBE marginally outperforms EvE. When assuming agents playing symmetric strategies, and estimating the models with session data, IBE yields more theory-consistent estimates than EvE, no matter the payoff structure or level. However, the opposite occurs when the symmetry assumption is relaxed. The conduct of a specification test rejects the validity of the restrictions on entry probabilities imposed by EvE for agents with symmetric strategies, in 50 to 75% of sessions and it always rejects it in the case of IBE, which indicates that the symmetric variant of these models have little empirical support.

Keywords: congestion games, exploration *vs* exploitation, quantal response equilibrium, impulse balance equilibrium, specification test, experimental economics.

JEL Classification Numbers: C7, C92

‡ We thank Jacob Goeree, Reinhard Selten, Ted Turocy and audiences at Aix-Marseille School of Economics (GREQAM), University of New South Wales, the Economic Science Association International Meeting (Hawaii), the Australia New-Zealand Workshop in Experimental Economics (Melbourne), the Workshop on Behavioral and Experimental Economics (Aix-en Provence), the 11th Workshop on Complexity: Theory and Experimental Analyses (Nice), and the École des Hautes Études en Sciences Sociales (CAMS, Paris) for useful discussions and comments, as well as Nikolaos Georgantzis for access to the Laboratory for Experimental Economics at Universidad Jaume I (Castéllon, Spain), Raquel Barreda and Juan Gómez for assistance with conducting the experiments, and Yohan Mathis and Mathilde Saguin for able research assistance on parts of the statistical analysis at the University of Strasbourg. Support from the European Commission through the *Laboratoires Européens Associés*, the Institute for Economic Analysis (Barcelona) and the University of Côte d'Azur (Nice, Project UCAinACTION of ANR-15-IDEX-01) where parts of this research were conducted, and the Australian Research Council (DP140102949) is gratefully acknowledged. The usual disclaimer applies.

[#]: École des Hautes Études en Sciences Sociales, CAMS (Paris).
email: kirman@ehess.fr
[§]: ZEW (Mannheim).
email: laisney@zew.de
[*]: School of Economics, University of Adelaide.
email: paul.pezanis-christou@adelaide.edu.au [Corresponding author]

Many social and economic activities require some degree of coordination in people's actions to be enjoyed. Congestion games, in which a group of individuals contemplates participating in an event that is enjoyed only if a few participate, characterise such situations well. These games have been widely investigated in the social sciences, and laboratory experiments suggest that despite the existence of multiple equilibria in pure strategies, participants somehow manage to behave almost optimally: overall, the observed participation rates even out the expected profits from entry and from no entry (Ochs, 1990, Arthur, 1994, Sundali, Rapoport and Seale, 1995, and Zwick and Rapoport, 2002). This pattern was first coined as 'magic' (Kahneman, 1988, Meyer, Van Huyck, Battalio and Saving, 1992, and Rapoport, 1995), and has subsequently been explained in terms of reinforcement learning processes (e.g., Erev and Rapoport, 1998, Rapoport, Seale, Erev and Sundali, 1998, Duffy and Hopkins, 2004, and Erev, Ert and Roth, 2010) or probability weighting (Rapoport, Seale and Ordonez, 2002). Yet, some investigations report persistent deviations (see Camerer and Lovallo, 1999, and Fischbacher and Thöni, 2008) for which Goeree and Holt (2005) provide a rationale in terms of Quantal Response Equilibrium (QRE, McKelvey and Palfrey, 1995), a stochastic version of the Nash equilibrium that is remarkably successful in organising the data of numerous experiments (see Goeree, Holt and Palfrey, 2016, and the references therein).

In this paper, we use an alternative rationale for QRE that evokes the 'Exploration *versus* Exploitation' dilemma outlined in Nadal, Chenevez, Weisbuch and Kirman (1996), Kirman (2011) and Bouchaud (2013), and which has been formalised as 'rational inattention' by Matějka and McKay (2015). Unlike QRE which assumes payoff disturbances and noisy best-responding agents, 'Exploration *versus* Exploitation' (henceforth EvE) assumes agents to take account of the information cost (or the entropy) associated with their choice probabilities when maximising their expected payoffs. In a repeated congestion game for example, agents may want to "exploit" one option ('entry' or 'no entry') and sometime "explore" the other to gather information about the payoffs associated to it.[1] Bouchaud (2013) observes that these models are structurally equivalent when agents' payoff disturbances in quantal response models are assumed i.i.d. or equivalently, when agents in EvE models do not interact with each other and take their decisions independently. By discarding the possible effects of time-correlated decisions and agents' heterogeneity, these assumptions permit the determination of a stationary equilibrium in terms of agents' payoff responsiveness (in QRE) or their compromise between 'exploring' and

[1] See Hills, Todd, Lazer, Redish, Couzin, and the Cognitive Search Research Group (2015) for applications of the 'Exploration versus Exploration' dilemma in the social sciences.

'exploiting' (in EvE). The modelling of dynamics in such settings quickly becomes intractable and the determination of equilibria is confined to special cases, see e.g., Bouchaud (2013) and Goeree, Holt and Palfrey (2016).

On the other hand, the modelling of heterogeneity in game-theoretic contexts raises questions about which sort of relaxation of "common knowledge" assumption(s) about what agents believe about others can be used and which still allow one to 'close' the model. Various answers have been proposed, as in e.g., Armantier and Treich (2009), Rogers, Camerer and Palfrey (2009) and Camerer, Nunnari and Palfrey (2016).[2] Although making these modifications shows that assuming heterogeneity in agents' traits considerably improves the model's goodness-of-fit, it does not address the question of *when* one should indeed forego the assumption of homogeneous agents. This approach, is also reminiscent of modifications that have been made to basic modern macroeconomic models to take account of what Angeletos and Lian (2016) describe as **"**the potential fragility of workhorse macroeconomic models to relaxations of common knowledge".[3]

The first goal of this paper is to address the falsifiability of this type of models with regard to homogeneity of agents' behaviour; see Goeree, Holt and Palfrey (2005, 2016) and Haile, Hortaçsu and Kosenock (2008) for discussions on the falsifiability of QRE. We provide a specification test that checks the consistency of individuals' behaviour with the model's prediction for homogenously behaving agents. Our approach is somewhat related to that of Golman (2011) who determines conditions under which the behaviour of the representative agent of a pool of individuals from a heterogeneous population may be rationalised by QRE. These conditions refer to the properties of agents' payoff disturbances that determine whether their aggregation fulfils the i.i.d. assumption of QRE. Unlike Golman's normative approach that is suited to QRE and that generates most useful predictions for the QRE analysis of binary-choice

[2] See McKelvey, Palfrey and Weber (2000) and Weizsäcker (2003) for earlier investigations of QRE models with non-homogenous agents. Rogers et al. (2009) also consider a QRE model with *subjective* beliefs, i.e., where the distributions of others' traits are not common knowledge but each player believes that the others' traits are i.i.d. from the same distribution as her/his (which is private information). Armantier and Treich (2009) study a setting where agents (bidders in first-price auctions) know that their private traits (i.e., agents' risk aversion and probability misperception) are i.i.d. from some commonly known bivariate distribution. In this setting, agents are heterogeneous but remain *ex ante* symmetric, whereas in the non-*subjective* models of Rogers et al. they are *ex ante* asymmetric so the relaxation of the usual common-prior assumption must additionally be dealt with.

[3] We also note a subtle difference between empiricists and experimentalists regarding the definition of *unobserved heterogeneity* in the structural analysis of game-theoretic models. The former define it as information that is available to agents but not to the researcher (Paarsch and Hong, 2006) whereas the latter assume that the researcher is aware of this information and uses it to assess the model's explanatory power (e.g., Armantier and Treich, 2009, and Goeree Holt and Palfrey, 2016). See Kirman (2006) for an overview of the role of heterogeneity in economics and Branch and McGough (2018) for a discussion of the role of heterogeneous expectations in macroeconomics. See also Bookstaber and Kirman (2018) on the difficulties with building both theoretical and computational models of heterogeneous agents even when the nature of the heterogeneity is known and understood.

asymmetric games, ours is based on the means of observable variables and applies more generally to discrete-choice models that assume homogeneous agents.

The paper's second goal is to compare the explanatory power of EvE to that of Impulse Balance Equilibrium (IBE, Selten, Abbink and Cox, 2005, Ockenfels and Selten, 2005) which assumes neither expected profit-maximisation nor best-responding behaviour. IBE basically determines a strategy that equalises the foregone expected payoffs associated to each possible action. It assumes that if at some stage an alternative option would have yielded a higher payoff, then the agent receives an impulse to use this alternative in the next stage. IBE is defined as the long run outcome of such stage-to-stage behaviour which is driven by Learning Direction Theory (Selten and Buchta, 1999) and like EvE, it entails a trade-off and a stationary equilibrium. The relative goodness-of-fit performances of these (and other) models have been investigated in the context of 2×2 constant-sum games by Selten and Chmura (2008), Brunner, Camerer and Goeree (2010), and Selten, Chmura and Georg (2010). Here we set out to compare these models in a similar context, i.e., repeated binary-choice, but with 'many' agents and with treatments that manipulate payoff *levels* (i.e., 'High' or 'Low' payoffs) and *structures* (i.e., where payoffs from entering depend on total attendance). The motivation to analyse a game with 'many' (10) agents is to best allow the structural assessment of agents' homogeneity. Controlling the payoff levels allows to check whether entry is increasing in payoffs, as expected by McKelvey, Palfrey and Weber (2000) and Goeree, Holt and Palfrey (2005), whereas controlling payoff structures aims at providing a broader picture of the problem at hand.

We summarise our findings in the following four points. First, average entry behaviour in the experiments is mostly in line with symmetric Nash equilibrium predictions when payoffs are Low, no matter the payoff structure or whether we use 'sessions' or 'pooled' data (of several sessions). When payoffs are High, behaviour is characterized by a significant under-entry, especially when assuming pooled data or a payoff structure that is non-monotone in entry, i.e., first increasing and then decreasing with entry. Second, when the models are estimated by maximizing (pseudo-) likelihood functions that assume players with homogenous traits who make symmetric choices, IBE yields more consistent estimates than EvE across sessions or when using pooled data. Further, EvE and IBE explain the data *equally* well in terms of goodness-of-fit if the probability of entry is less than or equal to the symmetric Nash mixed-equilibrium prediction. Otherwise, IBE marginally outperforms EvE in terms of goodness-of-fit. The models' pooled estimates also display no particular trends throughout the experiments, which suggests no obvious learning pattern. Third, when the estimations relax symmetry, EvE generates more consistent estimates than IBE, no matter the payoff structure or level and whether the estimates refer to ‘session’ or ‘pooled’ data. Fourth, our specification test rejects the null of consistency with EvE in 50% to 75% of all sessions (mostly with High payoffs), and it rejects the null of consistency with IBE in *all* sessions.

The next section spells out the EvE and IBE models for Arthur’s (1994) *El Farol* bar game. Section 2 lays out the econometric procedures and our specification test for this type of games. The experimental design and procedures are presented in Section 3. Section 4 reports on the estimation outcomes, with and without controlling for the symmetry of players' choices and Section 5 concludes.

1. Two stationary models for congestion games

Assume n agents who independently decide whether to enter a market or not. Agent i's decision is represented by a variable d_i that takes the value 1 if she enters and 0 if not. The payoff from not entering is constant and equal to H, whereas the one from entering is a function $G(\cdot)$ of the number of entrants $A = \sum_i d_i$. A congestion problem arises if for some values c of A, we have $G(c) > H$ and for others $G(c) < H$. With such a reward scheme, any vector of decisions d such that exactly c out of n agents choose to enter constitutes a pure Nash equilibrium. Thus there are exactly $\binom{n}{c}$ such equilibria: each yielding an aggregate payoff equal to $cG(c) + (n-c)H$. There may also exist symmetric mixed-equilibrium strategies which, by definition, equalize an agent's expected payoff from entering, π^E, to that from not entering, $\pi^{NE} = H$. That is, if p stands for the common probability of entry, then p^{Nash} solves:

$$\pi^E(p) \equiv \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} G(k+1) = H \equiv \pi^{NE} \qquad (1)$$

where k is a realization of the random variable K characterizing the number of entrants other than oneself. Note that (1) requires that the n agents behave symmetrically in that they all choose to enter with the same probability p.[4] For reasons that will become clear in Section 2, it is convenient to rewrite this expression as being conditional on vector p_{-i}, the $n-1$ vector of entry probabilities for agents other than agent i:[5]

[4] Clearly, there also are asymmetric mixed-equilibrium strategies which assume that some agents always or never enter.

[5] See Appendix 1 for the expression of the probability $P[k\ go|\ p_{-i}]$ that k agents enter given the vector p_{-i}.

$$\pi_i(d = 1|p_{-i}) = \sum_{k=0}^{n-1} P[k\ go|\ p_{-i}]\, G(k+1).$$

1.1. Exploration *vs* Exploitation: EvE

In this framework, agents aim at finding a compromise between maximizing their current payoff and keeping themselves informed about market conditions to maximize their future payoffs. In our context, we can think of changing market conditions if agents exhibit an irregular or noisy entry behaviour. If that is the case, then they may find it worthwhile to sometimes explore the alternative option, i.e., entering or not entering the market. While the 'exploitation' part of the dilemma, i.e., the maximization of current payoffs, is straightforward, the 'exploration' part hinges upon the *maximum entropy principle* which captures the agent's variety-seeking behaviour (see Anderson, de Palma and Thisse, 1992). In brief, an agent seeking maximal variety in her/his decisions would explore each alternative with equal probabilities so that entropy is maximized whereas an agent who values minimal variety would avoid exploring and would focus on maximizing current payoffs, in which case entropy is minimized.

With p_i standing for an agent i's probability of entry, $\pi^E(p_{-i})$ for the expected payoff in terms of the probabilities of entry of the $n-1$ other agents, and with entropy being defined by $S_i = -p_i \ln p_i - (1-p_i)\ln(1-p_i)$, the agent's objective function to maximise is given by:

$$\begin{aligned} \pi_i &= p_i \pi^E(p_{-i}) + (1-p_i)H + \sigma S_i \\ &= p_i \pi^E(p_{-i}) + (1-p_i)H + \sigma[-p_i \ln p_i - (1-p_i)\ln(1-p_i)]. \end{aligned} \quad (2)$$

where $\sigma \geq 0$ is a parameter capturing the weight that agent i assigns to the preservation of information about market conditions for long term profits. Differentiating this expression with respect to p_i, we obtain the following first-order condition of maximisation:

$$\pi^E(p_{-i}) - H + \sigma[-\ln p_i + \ln(1-p_i)] = 0,$$

or equivalently (if p_i is neither 0 nor 1 and $\sigma = 1/\lambda$)

$$p_i = \frac{1}{1 + \exp\{-\lambda[\pi^E(p_{-i}) - H]\}}. \quad (3)$$

This yields a system of n equations if there are n agents, and given the homogenous weighting parameter λ, this should be solved for the vector $p^* = (p_1, p_2, \dots, p_n)$. Note that this exactly matches McKelvey and Palfrey's definition of the logit QRE, with λ standing for agents' homogenous best-responsiveness. Unlike QRE, EvE yields the logit choice function without assuming noisy payoffs, which is questionable in this setting, since 'not entering' generates a sure payoff H. Under the assumption of symmetry, p_{-i} has all its components equal to p_i, which we simply denote p, and thus p and λ are related by

$$p = \frac{1}{1 + \exp\{-\lambda[\pi^E(p) - H]\}}$$

or equivalently

$$\lambda = \frac{\ln \frac{p}{1-p}}{\pi^E(p) - H}. \tag{4}$$

Thus, if agents who choose symmetrically are rational and do not explore, then p is such that $\pi^E(p) = H$, i.e., $p = p^{Nash}$ and $\lambda \to \infty$. On the other hand, if they maximise exploration, then they choose p such that $p = 1 - p = 0.5$, so that $\lambda \to 0$. If $p > 0.5$, λ is positive if $\pi^E(p) > H$ and it is negative otherwise. The Maximum Likelihood estimate of p, assuming independent observations, is the relative frequency of entry, $\overline{d}_n$, and the MLE of λ follows from (4). Note that $\overline{d}_n$ remains a consistent estimator for $E(d) = p$ for less restrictive covariance structures of the observations, by various weak laws of large numbers.

1.2. Impulse Balance Equilibrium: IBE

According to IBE, agents only take account of foregone payoffs. In the context of a congestion game, an agent receives an impulse for entry if the payoff received from not entering the market is strictly smaller than that from entering. Denoting by I the number of other entrants and by p the common probability of entering the market, the expected magnitude of these impulses is defined as:

$$\begin{aligned} IMP^E(p) &= E[\mathrm{G}(\mathrm{I}+1)\mathbb{I}_{\{G(I+1)>H\}}] \\ &= \sum_{i=0}^{c-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} G(\mathrm{i}+1)\mathbb{I}_{\{G(i+1)>H\}} \end{aligned}$$

or equivalently in terms of p_{-i} rather than p:

$$IMP^E(p_{-i}) = \sum_{k=0}^{c-1} P[k \text{ go } |p_{-i}]\, G(\mathrm{i}+1)\mathbb{I}_{\{G(i+1)>H\}}. \tag{5}$$

Similarly, an agent would receive an impulse for no entry if the payoff received from entering is not larger than that from not entering. The expected magnitude of these impulses is defined as:

$$IMP^{NE}(p) = H.P[G(I+1) < H]$$

$$= H\left[1 - \sum_{i=0}^{c-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} \mathbb{I}_{\{G(i+1)>H\}}\right]$$

or equivalently

$$IMP^{NE}(p_{-i}) = H\left[1 - \sum_{k=0}^{c-1} P[k \text{ go } |p_{-i}]\, \mathbb{I}_{\{G(i+1)>H\}}\right]. \tag{6}$$

Note that these impulses are defined relatively to the game's maximin pure strategy of not entering the market which yields a sure payoff of H. Selten and Chmura (2008) further observe that receiving a payoff lower than this sure payoff should be perceived as a loss. To this extent, and in the light of empirical and experimental evidence of loss aversion in agents' preferences (Bernatzi and Thaler, 1995, and Tversky and Kahneman, 1991), Selten and Chmura assume that the losses relative to the secured payoff count double. Thus, in the context of a congestion game, an IBE could be determined such that agent i is indifferent between entering and receiving $IMP^E(p_{-i})$ and not entering and receiving $2 \times IMP^{NE}(p_{-i})$. Put alternatively, agent i would choose to enter the market with probability p_i that equalises her expected impulses; that is

$$p_i IMP^E(p_{-i}) = (1-p_i) \times 2 \times IMP^{NE}(p_{-i}) \tag{7}$$

Hereafter, we follow Ockenfels and Selten (2005) and consider a parametric version of IBE by estimating the impulse weight κ in the following equation by Maximum Likelihood (as for QRE), i.e., the estimator of p is $\overline{d_n}$ and the MLE of κ (assuming p different from 0) follows from

$$\kappa = \frac{pIMP^E(p)}{(1-p)IMP^{NE}(p)}. \tag{8}$$

2. A specification test

When we assume symmetry, i.e., that agents' entry probabilities $p_1, p_2, \dots, p_n$ are all equal to p, the models we consider only propose a reparametrization $\lambda(p)$ for EvE and $\kappa(p)$ for IBE. Thus, under symmetry, there is no scope for discriminating between these models beyond commenting on implausible values of $\lambda(p)$ and $\kappa(p)$. If we do not impose symmetry, then (3) and (7) can be rewritten as systems of linear restrictions on parameters λ and κ:

$$\ln \frac{p_i}{1-p_i} - \lambda[\pi_i^E(p_{-i}) - H] = 0 \qquad \text{for } i = 1, \dots, n. \tag{9}$$

and

$$p_i IMP^E(p_{-i}) - (1-p_i)\kappa IMP^{NE}(p_{-i}) = 0 \qquad \text{for } i = 1, \dots, n. \tag{10}$$

Both systems can thus be written in the form $y(p) - \theta x(p) = g(p, \theta) = 0$, with $\theta = \lambda$ or κ, and with y, x and g vector functions with values in $\mathbb{R}^n$.

Given the asymptotically normal estimator $\hat{p}_T$ of p, the vector of individual entry frequencies, with asymptotic variance V of which we describe a consistent estimator $\hat{V}_T$ in Appendix 2, an optimal asymptotic least squares estimator of θ is:[6]

$$\begin{aligned} \hat{\theta}_T &= \arg\min_\theta \, g'(\hat{p}_T, \theta)\hat{S}_T^{-1} g(\hat{p}_T, \theta) \\ &= [x'(p)\hat{S}_T^{-1}x(p)]^{-1}x'(p)\hat{S}_T^{-1}y(p) \end{aligned} \tag{11}$$

with $\hat{S}_T$ converging to

$$S = \frac{\partial g(p,\theta)}{\partial p'} V \frac{\partial g'(p,\theta)}{\partial p}.$$

$\hat{\theta}_T$ is thus the GLS estimator in the regression of $y(p)$ on $x(p)$, the variance of the error term being S.

Given a preliminary estimate of θ, say $\tilde{\theta}_T$ obtained by replacing $\hat{S}_T$ in (11) with the identity matrix, i.e., $\tilde{\theta}_T$ is the OLS estimator in the regression of $y(p)$ on $x(p)$, a consistent estimator of S is

$$\hat{S}_T = \frac{\partial g(\hat{p}_T, \tilde{\theta}_T)}{\partial p'} \hat{V}_T \frac{\partial g'(\hat{p}_T, \tilde{\theta}_T)}{\partial p}.$$

[6] See e.g. Gouriéroux and Monfort, 1995, for a justification of all the assertions in this section.

The asymptotic variance of $\hat{\theta}_T$ is given by $V_{asy}(\hat{\theta}_T) = [x'(p)S^{-1}x(p)]^{-1}$ and a consistent estimator is $\widehat{V_{asy}(\hat{\theta}_T)} = [x'(\hat{p}_T)\hat{S}_T^{-1}x(\hat{p}_T)]^{-1}$. Under the null that there exists θ such that $g(p,\theta) = 0$ for the true p, or in other words that the restrictions on entry probabilities embodied by the model are valid,

$$Tg'(\hat{p}_T, \hat{\theta}_T)\hat{S}_T^{-1}g(\hat{p}_T, \hat{\theta}_T) \approx \chi^2(n-1), \tag{12}$$

and this can be used to test the underlying theory. All we need for the implementation of this specification test are thus $\hat{V}_T$ and the derivatives $\partial g_i(p,\theta)/\partial p_i$. The technical details for the determination of these expressions are given in Appendix 3.

As we obtained counter-intuitive results for IBE, with a large number of negative estimates for κ, we explored some variants.[7] One consists in iterating the FGLS procedure, but this produced oscillating results, and thus we turned to optimizing (12) with respect to the argument θ, with θ appearing both in function $g(\hat{p}_T, \theta)$ and in the weighting matrix $\hat{S}_T$, which amounts to obtaining the limit of iterated FGLS.

The above test pertains to the analysis of a session with n agents. When pooling k sessions of n agents, we match agents in different sessions according to their session average entry probabilities. Technically, this is done by renumbering the agents in each session in increasing order of session average entry probabilities. We then treat the observations as if they resulted from the actions of the same n agents over the k sessions. The only change in the estimation and test procedure is that T is replaced with kT.

3. Experimental design and procedures

The experiments involve groups of 10 participants and a 2×3 factorial design which assumes two payoff *levels*, High and Low, and three payoff *structures*, a two-step function (DISC) yielding a positive payoff G from entering if total entry does not exceed the market's capacity c, and 0 otherwise, and two non-monotone ones in which payoffs first increase and then decrease with attendance (NOM1 and NOM2, the latter implying entry if someone already entered). These payoff functions are displayed in Figure 1. There are two things worth mentioning here. First, to keep the payoff structures comparable for a given payoff level, the total payoff from entering the market, measured by $\sum_{i=1}^{9} G(i+1)$, and the market capacity c have been kept constant across

[7] It may appear surprising that regressing the positive variable $p_i IMP^E(p_{-i})$ on the positive variable $(1-p_i)IMP^{NE}(p_{-i})$ may yield a negative coefficient but this can happen with Feasible Generalized Least Squares with a non-diagonal weighting matrix, as is the case here.

payoff structures.[8] Second, DISC and NOM1 have $\binom{10}{6} = 210$ Nash equilibria in pure strategies whereas NOM2 has one more in which all agents choose not to enter.

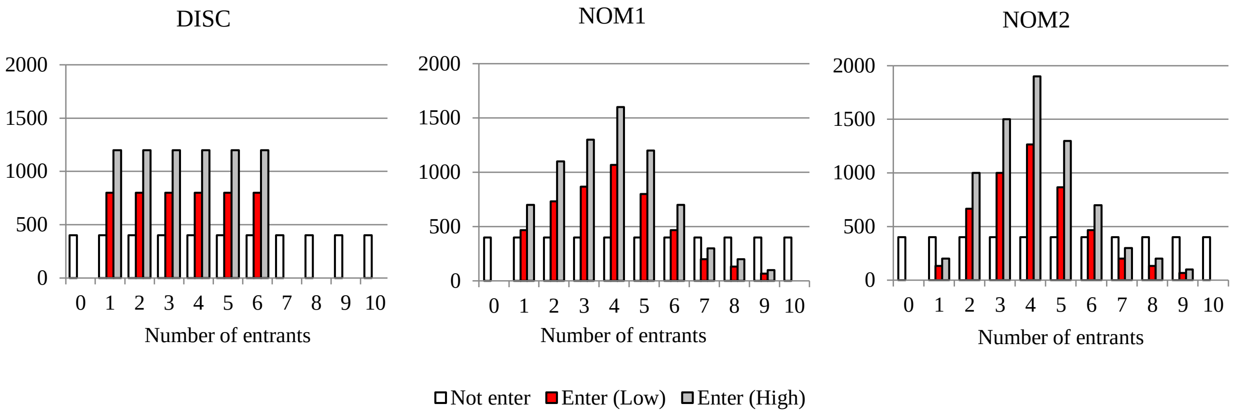
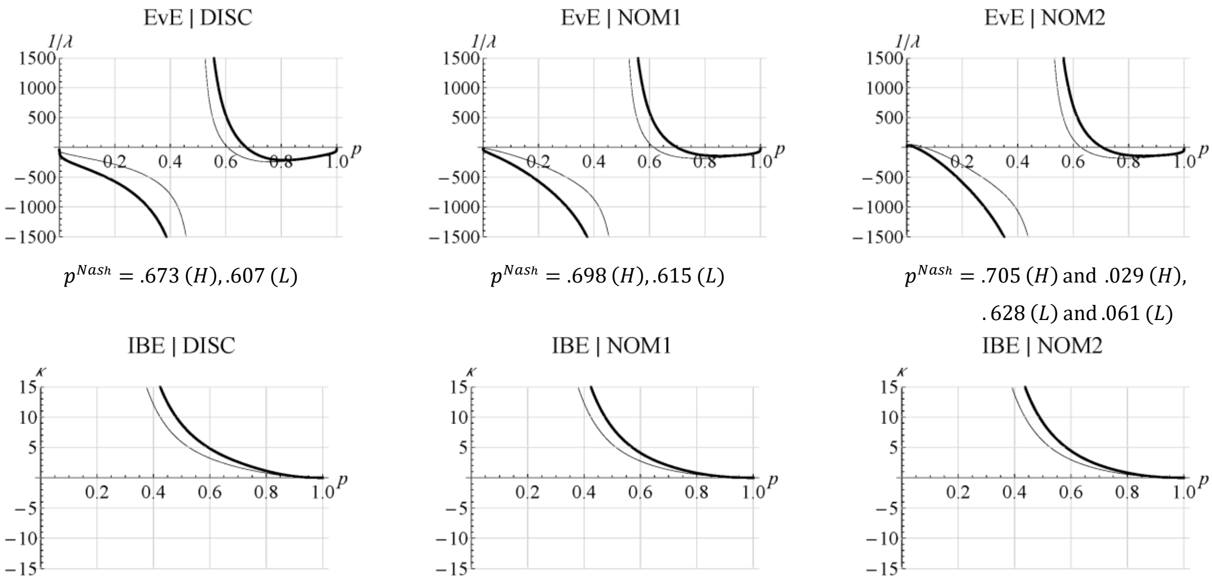


FIGURE 1: PAYOFF LEVELS AND STRUCTURES



Note: Bold (Thin) lines stand for High (Low) payoff levels.

FIGURE 2: RELATIONSHIP BETWEEN p AND $1/\lambda$ (in EvE) OR κ (in IBE).

[8] The alternative to keep the symmetric Nash mixed-equilibrium prediction constant across payoff levels, as in McKelvey, Palfrey and Weber (2001), would require changing either $\sum_{i=1}^{9} G(i+1)$ or c and has therefore not been considered.

Figure 2 displays the relationships between $\sigma = 1/\lambda$ and p in EvE and between κ and p in IBE and reports the mixed-equilibrium strategies for EvE. Recall that for EvE, $\lambda \to \infty$ if $p = p^{Nash}$ and that the observed probability p can be rationalized in terms of EvE only if $\lambda > 0$.[9] Note also that the presence of the additional 'no entry' equilibrium in the NOM2 treatments yields a second symmetric Nash equilibrium in mixed strategies with very low probabilities of entry.

The experiments were conducted at the Laboratory for Experimental Economics of the University of Jaume I (Spain). Participants were undergraduate students in Business Administration, Law or Engineering and were recruited by public advertisement on campus. We conducted eight sessions per payoff structure (DISC, NOM1, NOM2) involving a total of (8×3×10=) 240 individuals. For each payoff structure, we conducted four sessions with Low payoffs and four sessions with High payoffs. At the outset of the experiment, participants were randomly assigned to cubicles equipped with computer terminals and were given instructions that were read aloud.[10] To avoid framing effects, we presented the game in neutral language by asking participants to choose between actions A and B. Each session involved 150 rounds of play, and at the end of each round, participants were informed about the total number of players in their group who chose B ("No entry"), their own choice (A or B), their own payoff in that round and their cumulated payoff. This information was appended to a "History" window that could be seen at any time during the experiment. Each session lasted for a maximum of one hour, including the time needed to read the instructions, and individual average earnings were €12.77 (i.e., €11.94 in the Low payoff sessions and €13.60 in High payoff ones).

4. Results

Figure 3 shows the evolution of average entry in each treatment and reports local linear kernel regression estimates of entry and their 95% confidence interval. Overall, although entry is close to capacity on average (i.e., $c = 6$), it is typically higher in High payoff treatments than in Low payoff ones. The kernel regression estimates also suggest that in most treatments, subjects played cautiously at the outset of a session (by refraining from entering the market) and were more 'outgoing' towards the end of it.

---

[9] For the treatments we experimentally investigate, only values of p greater than 0.5 are relevant. And in the range $[0.5,1]$, the relationship $p \mapsto 1/\lambda$ turns out to be monotonic decreasing over at least $[0.5,0.705]$ for all payoff structures. The relationship is therefore a bijection over the relevant range of p, and p and $1/\lambda$ are equivalent parametrisations of the EvE model.

[10] The experiments were conducted with the *z-Tree* software (Fischbacher, 2007). See Appendix 4 for an English transcript of a set of instructions.

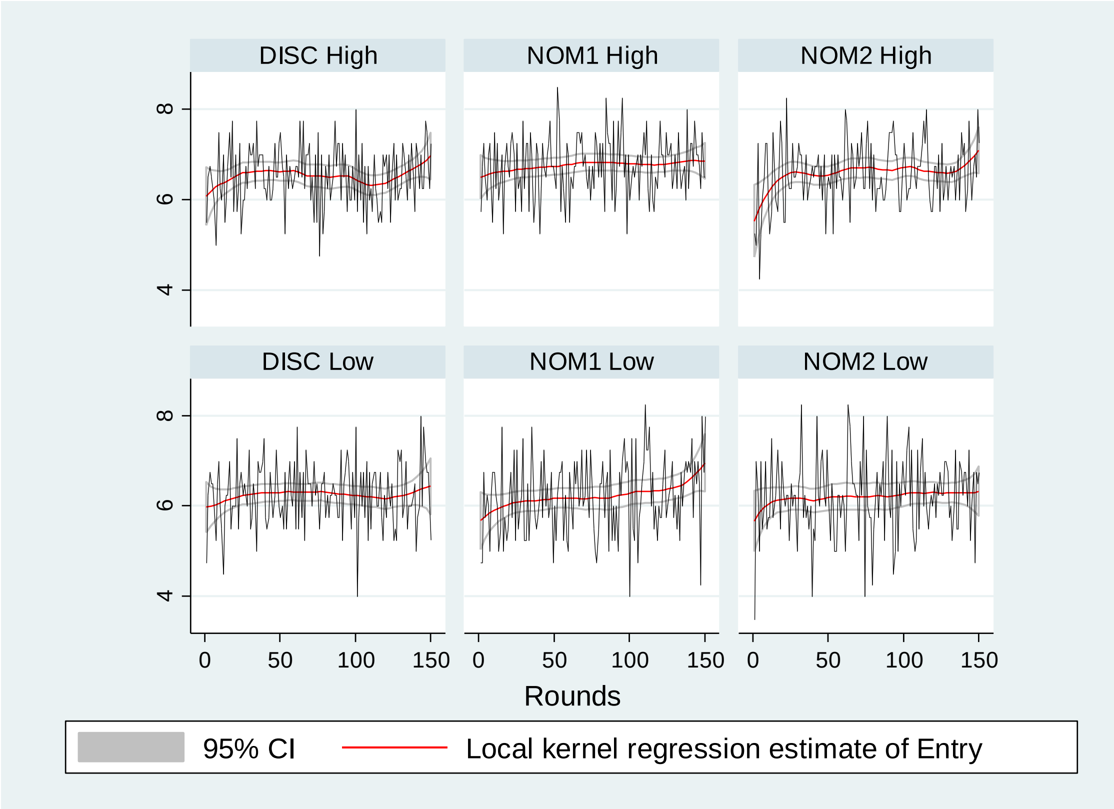


FIGURE 3: AVERAGE ENTRY BEHAVIOUR.

Table 1 reports on average entry probabilities. Looking first at session outcomes, a comparison of the statistics' 95% confidence intervals to equilibrium predictions suggests that the 'magic' of an equilibrium-type of behaviour is verified in 16 out of the 24 sessions conducted. Otherwise we have a significant *under*-entry in six out of the 12 sessions with High payoffs and in one session out of the 12 sessions with Low payoffs, as well as a significant *over*-entry in one Low payoff sessions. Given a payoff structure, the symmetric mixed-equilibrium strategy is typically rejected in only one session out of four; the only exceptions being the NOM2/High treatment where under-entry is observed in all sessions and the NOM1/Low treatment where the null of equilibrium play cannot be rejected at $\alpha = 5\%$ in all sessions. This is partially confirmed by 'pooled' data outcomes, as we find under-entry in all High payoff treatments and in NOM2/Low, and over-entry is only observed in DISC with Low payoffs.

In sum, the data support the prediction of Goeree and Holt (2005) that a significant under-entry occurs when $p^{Nash} > 0.5$, but mostly in High payoff sessions, as in NOM2, or when the data is pooled over sessions, as in DISC, NOM1 and NOM2 (with either High or Low payoffs). In the following section, we check whether EvE and IBE support this finding.

TABLE 1: AVERAGE ENTRY PROBABILITIES.

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
High	DISC *.673*	.656 [.631, .681]	**.647** [.622, .671]	.651 [.617, .685]	.655 [.632, .678]	**.652** [.645, .659]
	NOM1 *.698*	**.669** [.640, .697]	.658 [.619, .669]	.677 [.652, .702]	.677 [.655, .699]	**.676** [.669, .681]
	NOM2 *.705*	**.669** [.644, .695]	**.635** [.609, .660]	**.660** [.639, .681]	**.673** [.653, .693]	**.659** [.653, .665]
Low	DISC *.607*	.630 [.602, .658]	.595 [.577, .614]	**.650** [.622, .678]	.621 [.597, .646]	**.624** [.618, .630]
	NOM1 *.615*	.629 [.605, .654]	.643 [.612, .674]	.620 [.595, .645]	.594 [.566, .622]	.622 [.615, .628]
	NOM2 *.628*	.628 [.604, .652]	**.602** [.581, .623]	.619 [.596, .643]	.634 [.604, .664]	**.621** [.615, .627]

Note: Each 'session' ('pooled') estimate refers to 1500 (6000) observations; Nash mixed-equilibrium prediction in italics; Bold figures indicate a rejection of the null of Nash mixed-equilibrium play at α = *5%*; 95% confidence intervals (based on Newey-West variance estimates) in brackets.

4.1. Structural estimations when imposing symmetry

Table 2 reports the Maximum Likelihood estimation outcomes for EvE (upper panel) and IBE (lower panel). Note that since the models' log-likelihood values are defined as $T[\hat{p}\ln\hat{p} + (1-\hat{p})\ln(1-\hat{p})]$, with T standing for the number of observations, they contain no information about goodness-of-fit beyond the estimated probability of entry $\hat{p}$. In addition, since we correct the estimates' standard deviations for unknown forms of auto-correlation and heteroscedasticity in the observations, these pseudo-likelihood values are of doubtful interest and are therefore not reported.[11] The shaded cells characterize cases that did not reject the mixed-equilibrium predictions on the grounds of entry probabilities, cf. Table 1. Looking at EvE's $\hat{\lambda}$-estimates, they are not significantly different from zero in the High payoff sessions of DISC and NOM1, which suggests maximal exploration and conflicts with the diagnosis of a symmetric mixed-equilibrium behaviour in three out of four sessions of either treatment (cf. Table 1). The estimates for NOM2/High are significantly positive in all sessions, which suggests a contained exploration that is in line with the reported under-entry for this treatment. When payoffs are Low, the $\hat{\lambda}$-

[11] In Appendix 5, we report on the models' estimation outcomes when assuming i.i.d. observations, as is usually done in the literature, so that the models' log-likelihood values become meaningful. As expected from Footnote 9, EvE and IBE generate *exactly* the same goodness-of-fit when the observed probability of entry p lies in $[0.5, p^{Nash}]$, and for the few cases where p lies outside this range, IBE marginally outperforms EvE in terms of goodness-of-fit.

estimates are typically not significantly different from 0 which suggests maximal exploration (in nine out of twelve sessions) or they are significantly negative and thus inconsistent.

TABLE 2: EVE PARAMETER ESTIMATES (λ) WHEN IMPOSING SYMMETRY.

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
High	DISC	**1.258** [-.773, 3.289]	**.762** [-.079, 1.602]	**.924** [-.702, 2.550]	**1.205** [-.538, 2.948]	1.000 [.634, 1.368]
	NOM1	**1.044** [-.712, 2.269]	**2.023** [-1.821, 5.867]	**1.521** [-.513, 3.555]	**1.577** [-.358, 3.512]	1.430 [.981, 1.879]
	NOM2	.812 [.068, 1.557]	.305 [.122, .487]	.595 [.220, .971]	.921 [.194, 1.649]	.580 [.480, .681]
Low	DISC	**-1.139** [-2.278, .0004]	**1.660** [-1.329, 4.649]	-.711 [-1.029, -.392]	**-1.698** [-4.211, .816]	-1.452 [-1.905, -1.001]
	NOM1	**-2.191** [-5.577, 1.195]	-1.236 [-2.288, -.184]	**-6.006** [-36.72, 24.71]	**1.016** [-.668, 2.699]	**-4.504** [-9.050, .0413]
	NOM2	**302.621** [-81383, 81989]	**.806** [-.044, 1.655]	**2.866** [-5.417, 11.150]	**-4.864** [-28.585, 18.86]	3.509 [.281, 6.736]

IBE PARAMETER ESTIMATES (κ) WHEN IMPOSING SYMMETRY

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
High	DISC	3.449 [2.919, 3.978]	3.652 [3.109, 4.196]	3.564 [2.820, 4.308]	3.463 [2.972, 3.954]	3.531 [3.386, 3.677]
	NOM1	2.511 [1.986, 3.036]	1.836 [1.526, 2.146]	2.366 [1.924, 2.808]	2.354 [1.961, 2.747]	2.387 [2.278, 2.496]
	NOM2	2.662 [2.150, 3.173]	3.445 [2.804, 4.087]	2.856 [2.415, 3.298]	2.595 [2.195, 2.995]	2.874 [2.751, 2.998]
Low	DISC	2.693 [2.236, 3.149]	3.306 [2.944, 3.667]	2.386 [1.974, 2.797]	2.836 [2.426, 3.245]	2.790 [2.684, 2.892]
	NOM1	2.224 [1.835, 2.612]	2.013 [1.566, 2.460]	2.374 [1.956, 2.793]	2.846 [2.285, 3.407]	2.347 [2.234, 2.459]
	NOM2	2.412 [1.987, 2.837]	2.908 [2.463, 3.354]	2.568 [2.136, 3.000]	2.309 [1.797, 2.821]	2.541 [2.426, 2.655]

Note: Each 'session' ('pooled') estimate refers to 1500 (6000) observations; shaded cells characterize instances where the symmetric mixed-equilibrium strategy cannot be rejected at the 5% level, cf. Table 1; bold figures indicate instances with maximal exploration in EvE, i.e., λ not significantly different from 0 at α = *5%*; 95% confidence intervals (based on Newey-West variance estimates) in brackets.

Estimating EvE with pooled data supports a contained exploration in all High payoff treatments eventhough this is not supported by the session data of DISC and NOM1. When payoffs are Low, behaviour is inconsistent in DISC and suggests either maximal exploration (in NOM1) or a contained one (in NOM2).

Overall, this first batch of estimations indicates a dissonance between ‘session’ and ‘pooled’ estimates, and between what can be concluded from the observed probabilities of entry and from EvE’s estimated parameters: for the 16 sessions that did not reject the null of a symmetric mixed-equilibrium behaviour, EvE diagnoses a behaviour that is not significantly different from maximal exploration for 15 sessions, and that is inconsistent for one session (#2, NOM1/Low).

Looking at IBE’s $\hat{\kappa}$-estimates and their 95% confidence intervals, we find no significant differences across sessions of a given payoff structure, no matter the payoff level. The estimates are also not significantly different from those pertaining to pooled data, and the latter are significantly different across payoff structures, with $\hat{\kappa}_{DISC} > \hat{\kappa}_{NOM2} > \hat{\kappa}_{NOM1}$, when payoffs are High. Pooled estimates are also significantly different across payoff *levels* in DISC and suggest an increased aversion towards (hypothetical) losses from entering when payoffs are High in this treatment, i.e., $\hat{\kappa}_{High} > \hat{\kappa}_{Low}$. Thus, although the $\hat{\kappa}$-estimates are all significantly larger than Selten and Chmura's benchmark of $\kappa = 2$ in all High payoff treatments (and in six sessions with Low payoffs), they display more consistency across sessions, as well as between ‘session’ and ‘pooled’ data, than EvE’s estimates.

Before investigating how the relaxation of the symmetry assumption affects our conclusions, recall that Figure 3 suggests an increased entry towards the end of an experiment which points to a possible non-stationary behaviour. We therefore check for possible trends in the models’ estimates of successive batches of 15 rounds. A declining trend in EvE’s $\hat{\lambda}$-estimates would suggest that participants explore less as they gain experience of the game they play whereas a decline in IBE’s $\hat{\kappa}$-estimates would suggest a dampening aversion towards the (hypothetical) losses attached to entering. The estimates are reported in Appendix 6 (along with average relative frequencies of entry) and provide no support for such trend patterns in any treatment, which is in line with the findings of McKelvey and Palfrey (1995) for a collection of experiments.

Interestingly, however, the models’ batch-estimates (which assume 600 observations) reproduce the same conflicting patterns as when comparing ‘session’ and ‘pooled’ data estimates in the High payoff treatments: *(i)* EvE’s estimates typically support maximal exploration whereas they indicate a contained one at the pooled level (cf. Table 2), *(ii)* average entry probabilities hardly reject the null of symmetric mixed-equilibrium play in DISC and NOM1, whereas EvE reports a contained exploration (under-entry) for the pooled data of these treatments (cf. Table 2), and *(iii)* IBE's estimates (of all treatments) are very similar to those reported in Table 2 and thus appear more consistent across aggregation levels than EvE's.

4.2. Structural estimations when relaxing symmetry

We now estimate the models without imposing symmetry, i.e., without assuming $p = p_i$ for $i = 1, \ldots, n$, and we run our specification test to assess the consistency of the estimates with the models' predictions. The outcomes reported in the two panels of Table 3 lead to very different conclusions. EvE's $\hat{\lambda}$-estimates are significantly positive in all sessions which suggests a contained exploration. They are also similar across sessions of a given payoff structure as well as across payoff levels, especially in DISC and NOM2, with DISC generating significantly smaller estimates than NOM1 and NOM2. 'Pooled' estimates are also in line with 'session' ones and, overall, those for DISC and NOM2 appear to be invariant to payoff levels whereas those for NOM1 are significantly smaller when payoffs are High, *as if* in this treatment participants explore less when payoffs are scaled up. Yet, the specification test rejects the null of consistency in 17 sessions (i.e., ten with High payoffs and seven with Low payoffs), and in all treatments when pooling the data.

Looking at IBE's $\hat{\kappa}$-estimates, they are less consistent across sessions of a given treatment and they are sometimes significantly negative, mostly in NOM1 and NOM2 with Low payoffs. The estimates pertaining to pooled data are broadly in line with those for session data, and those of DISC and NOM2 are significantly larger in the High payoff treatments, *as if* subjects' impulses for entry in these treatments increase with payoffs. However, running the specification test on pooled data systematically rejects the null of consistency with IBE's prediction. In sum, observed behaviour corresponds better to EvE than to IBE when the symmetry assumption is relaxed.

For each session we could also easily test the assumption that the probability of entry is the same for all session participants, but the rejection of the IBE specification for all cases makes this unnecessary: if the probability of entry is the same for all session participants, the restrictions imposed by IBE vanish, and we should not be able to reject them.

TABLE 3: EVE PARAMETER ESTIMATES (λ) WHEN RELAXING SYMMETRY.

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
	DISC	1.536 [1.452, 1.619] {$<10^{-3}$}	1.379 [1.305, 1.453] {$<10^{-3}$}	1.464 [1.415, 1.513] {.7937}°	1.300 [1.287, 1.313] {$<10^{-3}$}	1.550 [1.533, 1.567] {$<10^{-3}$}
High	NOM1	1.790 [1.760, 1.820] {.0449}	1.935 [1.832, 2.038] {.0454}	1.949 [1.852, 2.046] {.0217}	1.773 [1.728, 1.817] {$<10^{-3}$}	1.699 [1.694, 1.704] {$<10^{-3}$}
	NOM2	1.836 [1.827, 1.845] {.1185}°	1.943 [1.923, 1.963] {$<10^{-3}$}	1.754 [1.736, 1.772] {$<10^{-3}$}	1.610 [1.565, 1.655] {$<10^{-3}$}	1.835 [1.821, 1.849] {$<10^{-3}$}
	DISC	1.738 [1.726, 1.750] {.0034}	1.597 [1.586, 1.609] {$<10^{-3}$}	1.669 [1.621, 1.716] {.0015}	1.572 [1.550, 1.595] {$<10^{-3}$}	1.701 [1.698, 1.705] {$<10^{-3}$}
Low	NOM1	2.081 [2.029, 2.133] {.0860}°	2.050 [1.959, 2.141] {.0386}	2.046 [1.979, 2.112] {.0167}	2.081 [2.030, 2.131] {.1823}°	2.063 [2.056, 2.069] {$<10^{-3}$}
	NOM2	1.879 [1.858, 1.900] {.1739}°	2.115 [2.046, 2.185] {.4252}°	2.046 [1.979, 2.112] {.0167}	1.900 [1.860, 1.940] {.4949}°	1.865 [1.861, 1.869] {$<10^{-3}$}

IBE PARAMETER ESTIMATES (κ) WHEN RELAXING SYMMETRY.

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
	DISC	.857 [.758, .956] {$<10^{-3}$}	-1.231 [-1.265, -1.198] {$<10^{-3}$}	3.042 [2.612, 3.472] {$<10^{-3}$}	2.771 [2.442, 3.100] {$<10^{-3}$}	2.758 [2.677, 2.839] {$<10^{-3}$}
High	NOM1	1.547 [1.287, 1.807] {$<10^{-3}$}	.865 [.745, .984] {$<10^{-3}$}	-.681 [-.705, -.658] {$<10^{-3}$}	-.917 [-.944, -.891] {$<10^{-3}$}	.774 [.740, .809] {$<10^{-3}$}
	NOM2	1.823 [1.582, 2.065] {$<10^{-3}$}	2.373 [2.008, 2.738] {$<10^{-3}$}	-.866 [-.878, -.855] {$<10^{-3}$}	-.970 [-.986, -.954] {$<10^{-3}$}	1.790 [1.731, 1.850] {$<10^{-3}$}
	DISC	2.188 [1.860, 2.516] {$<10^{-3}$}	1.967 [1.832, 2.102] {$<10^{-3}$}	1.862 [1.593, 2.131] {$<10^{-3}$}	1.998 [1.851, 2.145] {$<10^{-3}$}	1.870 [1.810, 1.929] {$<10^{-3}$}
Low	NOM1	2.217 [1.931, 2.504] {$<10^{-3}$}	1.271 [1.028, 1.513] {$<10^{-3}$}	2.095 [1.853, 2.337] {$<10^{-3}$}	-.141 [-.217, -.064] {$<10^{-3}$}	1.668 [1.598, 1.738] {$<10^{-3}$}
	NOM2	1.649 [1.383, 1.914] {$<10^{-3}$}	2.421 [2.161, 2.682] {$<10^{-3}$}	.030 [-.048, .107] {$<10^{-3}$}	1.398 [1.132, 1.664] {$<10^{-3}$}	1.613 [1.546, 1.680] {$<10^{-3}$}

Note: Each 'session' ('pooled') estimate refers to 1500 (6000) observations; 95% confidence intervals (based on Newey-West variance estimates) in brackets; p-value of specification test in curly brackets; °: Non-rejection of the null of correct specification at α = 5%.

Robustness check: We proceed with estimating models with a less rigid specification that allows for an additional constant term in the regression of $y(p)$ on $x(p)$. In the context of EvE, this amounts to replacing (9) with

$$g_i(p,\zeta,\lambda) = \ln\frac{p_i}{1-p_i} - \zeta - \lambda[\pi_i^E(p_{-i}) - H]^{-1},$$

and in the context of IBE to replacing (10) with

$$g_i(p,\zeta,\kappa) = p_i IMP^E(p_{-i}) - \zeta - (1-p_i)\kappa IMP^{NE}(p_{-i}).$$

The estimation outcomes of these augmented versions are reported in Appendix 7. EvE's estimates are similar to those of Table 3 and several occurrences of the estimated constant term (Cste) are not significantly different from zero, this being mostly the case in NOM1 with High payoffs and in NOM2 with Low payoffs. As for IBE's estimates, the constant terms are all significantly positive (and rather large) and the $\hat{\kappa}$-estimates are all significantly negative which renders their interpretation meaningless. In addition, the conduct of our specification test confirms the findings of Table 3: it yields seven more non-rejections (at $\alpha = 5\%$) for EvE, bringing their total to 14 out of 24 sessions, and leaves things unchanged for IBE.[12]

5. Conclusion

We analyse behaviour in the *El Farol* bar game through the lens of stationary models that build on different premises. The main outcome of our study is that the assumption of symmetry of agents' behaviour should be taken seriously when estimating structural equilibrium models, as its non-fulfilment may dramatically affect conclusions. When the estimations impose symmetry, as is typically done in the literature, EvE and IBE 'explain' the data equally well in terms of goodness-of-fit, with IBE yielding stable estimates (that are typically greater than 2) across different levels of data aggregation whereas EvE does not share this property. However, the exact opposite holds when symmetry is relaxed: EvE then yields more stable estimates which suggest a contained exploration or, in QRE terms, a noisy best-response behaviour. The models' falsifiability regarding the symmetry assumption is addressed with a specification test that rejects the null of consistency with EvE (or equivalently QRE) in 50 to 75% of all sessions and rejects the null of consistency with IBE in *all* sessions (and both models overwhelmingly reject the null

[12] We also tried the reverse regressions $x(p)$ on $y(p)$ both with and without constant terms, but the results turned out to be much more unstable. Still, for the case without constant term, we used the reverse regressions as a convergence check, since both the direct and the reverse approaches must yield the same optimum when we optimize the test statistic (12).

when the data is pooled). Thus, our study shows that invoking the oversimplifying (but very convenient) assumption of symmetry to analyse game-like situations may be misleading if it is not properly dealt with and explicitly assessed when estimating the models. To this extent, it confirms that we should be very cautious with results that obtain from gross over-simplifications such as invoking symmetry, as we should when invoking a representative agent in mainstream economic theory.

References

Anderson S., A. de Palma and J.-F. Thisse, 1992, *Discrete choice theory of product differentiation*, MIT Press, New York.

Angeletos G. and C. Lian, 2016, "Incomplete information in macroeconomics: Accommodating frictions in coordination", NBER Working Papers 22297, National Bureau of Economic Research, Inc.

Armantier O. and N. Treich, 2009, "Probability misperception in games: An application to the overbidding puzzle", *International Economic Review*, 50(4), 1079-1102.

Arthur W., 1994, "Inductive reasoning and bounded rationality", *American Economic Review*, 84, 406-411.

Bookstaber R. and A. Kirman 2018, "Modelling a Heterogeneous World" in *Handbook of Computational Economics, vol. 4, Heterogeneous Agent Modelling,* C. Hommes and B. LeBaron, (Eds), North Holland, pp. 769-795.

Bottazzi G. and G. Devetag, 2003, "A laboratory experiment on the minority game", *Physica A*, 124-132.

Bouchaud J.-P., 2013, "Crises and collective socio-economic phenomena: simple models and challenges", *Journal of Statistical Physics*, 151, 3-4, 567-606.

Branch W. and B. McGough, 2018, "Heterogeneous expectations and microfoundations in macroeconomics", in *Handbook of Computational Economics, vol. 4, Heterogeneous Agent Modelling,* C. Hommes and B. LeBaron, (Eds), North Holland, pp. 3-59.

Brunner C., C. Camerer and J. Goeree, 2010, "A correction and re-examination of «Stationary concepts for experimental 2×2 games»", *American Economic Review*, 101, 1029-1040.

Camerer C. and D. Lovallo, 1999, "Overconfidence and excess entry: an experimental approach", *American Economic Review*, 89(1), 306-318.

Camerer C., S. Nunnari and T. Palfrey, 2016, "Quantal response and nonequilibrium beliefs explain overbidding in maximum-value auctions", *Games and Economic Behavior*, 98, 242-263.

Challet D., M. Marsili and G. Ottino, 2003, "Shedding light on El Farol", *Physica A*, 332, 469-482.

Duffy J. and E. Hopkins, 2005, "Learning, information and sorting in market entry games: theory and evidence", *Games and Economic Behavior*, 51, 31-62.

Erev I., E. Ert and A. Roth, 2010, "A choice prediction competition for market entry games: an introduction", *Games*, 1, 117-136.

Erev I. and A. Rapoport, 1998, "Coordination, "Magic," and reinforcement learning in a market entry game", *Games and Economic Behavior*, 23, 146-175.

Fischbacher U., 2007, "z-Tree - Zurich Toolbox for Readymade Economic Experiments", *Experimental Economics*, 10(2), 171-178.

Fischbacher U. and C. Thöni, 2008, "Excess entry in an experimental winner-take-all market", *Journal of Economic Behavior and Organization*, 67, 150-163.

Goeree J. and C. Holt, 2005, "An explanation of anomalous behavior in models of political participation", *American Political Science Review*, 99(2), 201-213.

Goeree J., C. Holt and T. Palfrey, 2005, "Regular quantal response equilibrium", *Experimental Economics*, 8, 347-367.

Goeree J., C. Holt and T. Palfrey, 2016, *Quantal Response Equilibrium: A stochastic theory of games*, Princeton University Press.

Golman R., 2011, "Quantal response equilibria with heterogeneous agents", *Journal of Economic Theory*, 146, 2013-2028.

Gouriéroux C. and A. Monfort, 1995, *Statistics and Econometric Models*, Volumes 1-2, Cambridge University Press.

Haile P., A. Hortaçsu and G. Kosenok, 2008, "On the empirical content of quantal response equilibrium", *American Economic Review*, 98(1), 180-200.

He, X., 2009, "Confidence Intervals for the Binomial Proportion with Zero Frequency", *Paper SP10-2009*, SESUG.

Hills, T., P. Todd, D. Lazer, A. Redish, I. Couzin, and the Cognitive Search Research Group, 2015, "Exploration versus exploitation in space, mind, and society", *Trends in Cognitive Sciences*, 19, 46-54.

Kahneman D., 1988, "Experimental economics: a psychological perspective." In *Bounded Rational Behavior in Experimental Games and Markets*. R. Tietz, W. Albers, and R. Selten (Eds), New-York: Springer-Verlag, pp. 11-18.

Kahneman D. and A. Tversky, 1979, "Prospect theory: an analysis of decision under risk", *Econometrica*, 47, 263-91.

Kirman A., 2006, "Heterogeneity in economics", *Journal of Economic Interactions and Coordination*, 89-117.

Kirman A., 2011, *Complex Economics - Individual and collective rationality*, The Graz Schumpeter Lectures, Routledge.

McKelvey R. and T. Palfrey, 1995, "Quantal response equilibria for normal form games", *Games and Economic Behavior*, 10, 6-38.

McKelvey R., T. Palfrey and R. Weber, 2000, "The effects of payoff magnitudes and heterogeneity on behavior in 2×2 games with unique mixed strategy equilibria", *Journal of Economic Behavior and Organization*, 42, 523-548.

Matějka F. and A. McKay, 2015, "Rational inattention to discrete choices: A new foundation for the multinomial logit model", *American Economic Review*, 105(1), 272-298.

Meyer D., J. Van Huyck, R. Battalio and T. Saving, 1992, "History's role in coordinating decentralized allocation decisions", *Journal of Political Economy*, 100(2), 292-316.

Nadal J.-P., G. Weisbuch, O. Chenevez, and A. Kirman, 1998, "A formal approach to market organization: choice functions, mean field approximation and maximum entropy principle", in *Advances in Self-Organization and Evolutionary Economics*, J. Lesourne and A. Orléan (Eds), Economica, London, pp. 149-159.

Newey W. and K. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55, 703-708.

Newey W. and K. West, 1994, "Automatic Lag Selection in Covariance Matrix Estimation", *Review of Economic Studies*, 61, 631-653.

Ochs J., 1990, "Coordination in market entry games." In *Games and Human Behavior: Essays in Honor of Amnon Rapoport*, D. Budescu, I. Erev and R. Zwick (Eds.), Mahwah, N.J.: Lawrence Erlbaum Associates, pp. 143-172.

Paarsch H. and H. Hong, 2006, *An Introduction to the Structural Econometrics of Auction Data*, MIT Press.

Rapoport A., 1995, "Individual strategies in a market entry game", *Group Decision and Negotiations*, 4, 117-133.

Rapoport A., D. Seale, I. Erev and J. Sundali, 1998, "Equilibrium play in large group market entry games", *Management Science*, 44(1), 119-141.

Rapoport A., D. Seale and L. Ordonez, 2002, "Weighted probabilities in tacit coordination under uncertainty: theory and evidence from market entry games", *Journal of Risk and Uncertainty*, 25, 21-45.

Rogers B., T. Palfrey and C. Camerer, 2009, "Heterogenous quantal response equilibrium and cognitive hierarchies", *Journal of Economic Theory*, 144, 1440-1467.

Schotter A., 2009, "On the relationship between economic theory and experiments." In *Handbook of Experimental Economic Methodology*, G. Fréchette and A. Schotter (Eds), 2015. Oxford University Press, London.

Selten R., K. Abbink and R. Cox, 2005, "Learning direction theory and the winner's curse", *Experimental Economics*, 8, 5-20.

Selten R. and J. Buchta, 1999, "Experimental sealed bid first price auctions with directly observed bid functions", in *Games and Human Behavior: Essays in Honor of Amnon*

Rapoport, D. Budescu, I. Erev and R. Zwick, Mahwah, N.J.: Lawrence Erlbaum Associates, pp.79-104.

Selten R. and T. Chmura, 2008, "Stationary concepts for experimental 2×2 games", *American Economic Review*, 98(3), 938-966.

Selten R., T. Chmura and S. Goerg, 2010, "A correction and re-examination of «Stationary concepts for experimental 2×2 games»: a reply", *American Economic Review*, 101, 1041-1044.

Selten R, M. Schreckenberg, T. Pitz, T. Chmura and S. Kube, 2007, "Commuters' route choice behaviour", *Games and Economic Behavior*, 58, 394-406.

Sundali J., A. Rapoport and D. Seale, 1995, "Coordination in market entry games with symmetric players", *Organizational Behavior and Human Decision Processes*, 64(2), 203-218.

Tversky A. and D. Kahneman, 1991, "Loss aversion and riskless choice: a reference dependent model", *Quarterly Journal of Economics*, 107, 1039-1061.

Weizsäcker G., 2003, "Ignoring the rationality of others: Evidence from experimental normal form games", *Games and Economic Behavior*, 145-171.

Zeileis A., 2004, "Econometric Computing with HC and HAC Covariance Matrix Estimators", *Journal of Statistical Software*, 11(10), 1-17. URL: http://www.jstatsoft.org/v11/i10/.

Zwick R. and A. Rapoport, 2002, "Tacit coordination in a decentralized market entry game with fixed capacity", *Experimental Economics*, 5, 253-272.

Appendix 1: Quantal Response Equilibrium.

In this appendix we provide the usual arguments for the derivation of a Quantal Response Equilibrium for the congestion game we study. According to QRE, agents make mistakes in their best replies and the latter take account of rivals' mistakes. QRE postulates in particular that each agent has a random utility of his decision which is the payoff plus a "standard error" parameter times a random term. The two random terms corresponding to entering and not entering are assumed to take i.i.d. extreme values, and the payoff is computed as expected payoff given the vector p_{-i} of the probabilities of entering of the other agents. Formally, for agent i, we have:

$$U_i(d = 1, p_{-i}) \quad = \pi^E(p_{-i}) + \sigma\varepsilon_1$$

and

$$U_i(d = 0, p_{-i}) \quad = H + \sigma\varepsilon_0$$

Thus, if we denote p_i the probability of entering for agent i, we have:

$$\begin{aligned} p_i &= P[U_i(d = 1, p_{-i}) - U_i(d = 0, p_{-i}) \geq 0] \\ &= P[\sigma(\varepsilon_1 - \varepsilon_0) \geq -\pi_i^E(p_{-i}) + H] \\ &= \frac{1}{1 + \exp\{-\frac{1}{\sigma}[\pi^E(p_{-i}) - H]\}} \end{aligned}$$

since $(\varepsilon_1 - \varepsilon_0)$ is logistic. Note that the last expression is identical to (4) up to the reparametrization $1/\sigma = \lambda$.

Appendix 2: Probability of going conditional on p_{-i}.

To provide an expression for $P[k \text{ go}|p_{-i}]$, we introduce the notation $q = p_{-i}$. Clearly, we then have:

$$P[0 \text{ go}|q] \quad = \prod_{j=1}^{n-1}(1-q_j)$$

and

$$P[n-1 \text{ go}|q] \quad = \prod_{j=1}^{n-1} q_j$$

$$P[n-2 \text{ go}|q] \quad = \prod_{j=1}^{n-1}(1-q_j) \prod_{\substack{l\neq j,\\ l=1}}^{n-1} q_l$$

$$P[n-3 \text{ go}|q] \quad = \prod_{j_1=1}^{n-2}(1-q_{j_1}) \sum_{j_2=j_1+1}^{n-1}\left[(1-q_{j_2}) \prod_{\substack{l\neq j_2,j_1\\ l=1}}^{n-1} q_l\right]$$

and for the intermediate cases $0 < k < n-1$, we have:

$$P[1 \text{ goes}|q] \quad = \sum_{j=1}^{n-1} q_j \prod_{\substack{l\neq j,\\ l=1}}^{n-1}(1-q_l)$$

$$P[2 \text{ go}|q] \quad = \sum_{j_1=1}^{n-2} q_{j_1} \sum_{j_2=j_1+1}^{n-1} q_{j_2} \prod_{\substack{l\neq j_1,j_2,\\ l=1}}^{n-1}(1-q_l)$$

and more generally:

$$P[k \text{ go}|q] \quad = \sum_{j_1=1}^{n-k} q_{j_1} \cdots \sum_{j_k=j_{k-1}+1}^{n-1} q_{j_k} \prod_{\substack{l\neq j_1,\dots,j_k\\ l=1}}^{n-1}(1-q_l).$$

Appendix 3: Technical details.

Here we provide the technical details for the implementation of our specification test; namely the estimation of $\hat{V}_T$ and the derivatives $\partial g_i(p,\theta)/\partial p_i$ and $\partial g_i(p,\theta)/\partial p_j$ with $j \neq i$.

Estimation of $\widehat{V}_T$.

$\hat{p}_T = \sum_{t=1}^{n-1} d_t/T = \hat{d}_T$, with d_t the entry vector at t, with dimension n. We assume independence between agents in each round, thus the variance matrix is diagonal. We estimate each diagonal term using the heteroskedasticity and autocorrelation robust estimator of Newey and West (1987, 1994) with automatic lag selection, using the R program of Zeileis (2004).

A problem arises for participants in a session that choose a constant response (entry or non-entry). For them the frequency estimate of entry probability is 0 or 1 and the corresponding variance estimate is 0. For sessions where this occurred (i.e., NOM1/High/Session 2), we chose to model only the entry probabilities of agents who show some variation in their response, adapting the number of degrees of freedom in (12) accordingly. For the conditional entry probabilities described in Appendix 2, nothing changes. Another possibility would be to depart from the frequency estimator in those cases (see, e.g. He, 2009).

Determination of derivatives.

For EvE, we have:

$$\frac{\partial g_i(p,\lambda)}{\partial p_i} = \frac{1}{p_i(1-p_i)}$$

$$\frac{\partial g_i(p,\lambda)}{\partial p_j} = -\lambda \frac{\partial \pi_i^E(p_{-i})}{\partial p_j}$$

and it is fairly easy to compute $\partial \pi_i^E(p_{-i})/\partial p_j$ analytically: for $i \neq j$,

$$\frac{\partial \pi_i^E(p_{-i})}{\partial p_j} = \sum_{k=0}^{n-1} G(k+1) \frac{\partial P[k \text{ go}|p_{-i})}{\partial p_j}$$

For IBE, we have:

$$\frac{\partial g_i(p,\kappa)}{\partial p_i} = IMP^E(p_{-i}) + \kappa IMP^{NE}(p_{-i})$$

$$\frac{\partial g_i(p,\kappa)}{\partial p_j} = p_i \frac{\partial IMP^E(p_{-i})}{\partial p_j} - \kappa(1-p_i) \frac{\partial IMP^{NE}(p_{-i})}{\partial p_j}$$

where:

$$\frac{\partial IMP^E(p_{-i})}{\partial p_j} = \sum_{k=0}^{n-1} \frac{\partial P[k \text{ go}|p_{-i}]}{\partial p_j} G(k+1) \mathbb{I}_{\{G(k+1)>H\}},$$

and

$$\frac{\partial IMP^{NE}(p_{-i})}{\partial p_j} = -H\sum_{k=0}^{n-1}\frac{\partial P[k\text{ go}|p_{-i}]}{\partial p_j}\mathbb{I}_{\{G(k+1)>H\}}.$$

So the only ingredient still needed is $\partial P[k\text{ go}|p_{-i}]/\partial p_j$. The general expression for this is rather inelegant but it is fairly easily programmed once a program for computing $P[k\text{ go}|p_{-i}]$ is available.

Appendix 4: Instructions for treatment DISC (Low).

You are about to participate in an experiment on decision-making. In this experiment, there are 2 groups of 10 people. You are in one of those 2 groups and you do not know who else is in the same group as you. You will be in the same group of people for the whole experiment.

Please do not communicate in any way with other participants during the experiment.

The experiment is made of 150 rounds of a game that proceeds as follows:

1. In each round, you are asked to choose action A or action B. Once all participants have chosen their actions, the computer will calculate the total number of participants in your group who chose B.

2. The payoff you receive from choosing A or B is determined the following way:

- If you chose A, you will earn 400 points, whatever the other participants in your group have chosen.

- If you chose B and the total number of participants in your group who chose B is less than or equal to 6, then all participants in your group who chose B earn 800 points.

- If you chose B and the total number of participants in your group who chose B is more than 6, then all participants in your group who chose B will earn 0 points.

3. At the end of each round, you will be given the following information/feedback about the round you just played (the round number, the choice you made, how many times action B was chosen in your group, the profit you made in that round and the total profits you made so far),

4. At any time during the experiment, the outcomes of all previous rounds in which you played are displayed on the lower part of your terminal screen.

5. You are allowed to use the calculator we provide you with at the outset of the experiment.

6. The payoffs you earn in each round are quoted in terms of "points". Your reward from participating in this experiment is determined by the sum of your payoffs in points. Your total payoff in points will be exchanged for Euros (€) and paid in cash to you at the end of the experiment at the rate of 0.02 € per 100 points.

7. All other participants in this room received the same instruction sheet.

Please raise your hand if something is unclear or if you have a question to ask.

Appendix 5: Estimation outcomes assuming i.i.d. observations.

EVE PARAMETER ESTIMATES AND GOODNESS-OF-FIT.

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
High	DISC	**1.258** [-.712, 3.227] *965.48*	**.762** [-.072, 1.595] *974.23*	**.924** [-.229, 2.077] *970.55*	**1.205** [-.622, 3.032] *966.12*	1.000 [.340, 1.662] *3876.57*
	NOM1	1.044 [.012, 2.077] *952.68*	**1.770** [-.746, 4.285] *941.06*	**1.521** [-.414, 3.456] *944.03*	**1.577** [-.483, 3.637] *943.29*	1.430 [.559, 2.302] *3781.29*
	NOM2	.812 [.109, 1.516] *951.97*	.305 [.130, .479] *984.64*	.595 [.157, 1.033] *961.55*	.921 [.063, 1.780] *948.41*	.580 [.369, .791] *3849.52*
Low	DISC	∞ [n.a.] *990.12*	**1.660** [-2.326, 5.646] *1012.29*	∞ [n.a.] *977.09*	∞ [n.a.] *995.77*	∞ [n.a.] *3975.69*
	NOM1	∞ [n.a.] *989.59*	∞ [n.a.] *979.73*	∞ [n.a.] *996.17*	**1.016** [-.457, 2.488] *1013.05*	∞ [n.a.] *3979.97*
	NOM2	∞ [n.a.] *990.02*	**.806** [-.177, 1.788] *1008.29*	**2.866** [-5.857, 11.58] *996.58*	∞ [n.a.] *985.30*	**3.509** [-2.839, 9.857] *3981.93*

Note: Each ‘session’ (‘pooled’) estimate refers to 1500 (6000) observations; shaded cells characterize instances where the symmetric mixed-equilibrium strategy cannot be rejected at the 5% level, cf. Table 1; bold figures indicate instances with maximal exploration, i.e., λ not significantly different from 0 at α = *5%*; 95% confidence intervals (assuming i.i.d. observations) in brackets; -Log-Likelihood statistics in italics.

IBE PARAMETER ESTIMATES AND GOODNESS-OF-FIT.

Level	Structure	Session 1	Session 2	Session 3	Session 4	Pooled
High	DISC	3.449 [2.936, 3.962] *965.48*	3.652 [3.113, 4.192] *974.23*	3.564 [3.036, 4.092] *970.55*	3.463 [2.948; 3.978] *966.12*	3.531 [3.269, 3.793] *3876.57*
	NOM1	2.511 [2.069, 2.954] *952.68*	2.319 [1.906, 2.732] *941.06*	2.366 [1.946, 2.786] *944.03*	2.354 [1.936, 2.773] *943.29*	2.387 [2.175, 2.598] *3781.29*
	NOM2	2.662 [2.179, 3.145] *951.97*	3.445 [2.833, 4.058] *984.64*	2.856 [2.342, 3.371] *961.55*	2.595 [2.122, 3.067] *948.41*	2.874 [2.616, 3.133] *3849.52*
Low	DISC	2.693 [2.299, 3.087] *988.43*	3.306 [2.823, 3.788] *1012.29*	2.386 [2.033, 2.739] *971.17*	2.836 [2.422, 3.249] *995.11*	2.790 [2.585, 2.992] *3971.93*
	NOM1	2.224 [1.840, 2.607] *988.96*	2.013 [1.664, 2.362] *977.21*	2.374 [1.966, 2.783] *996.10*	2.846 [2.355, 3.377] *1013.05*	2.347 [2.145, 2.549] *3979.45*
	NOM2	2.412 [1.984, 2.840] *990.02*	2.908 [2.393, 3.423] *1008.29*	2.568 [2.113, 3.023] *996.58*	2.309 [1.899, 2.720] *985.19*	2.541 [2.315, 2.766] *3981.93*

Note: Each ‘session’ (‘pooled’) estimate refers to 1500 (6000) observations; 95% confidence intervals (assuming i.i.d. observations) in brackets; -Log-Likelihood statistics in italics.

Appendix 6: Experience effect and entry probabilities for successive batches of 15 rounds.

EXPERIENCE EFFECT – ENTRY PROBABILITIES FOR SUCCESSIVE BATCHES OF 15 ROUNDS.

	Batch → Treatment ↓	1	2	3	4	5	6	7	8	9	10
High	DISC	**.625** [.586, .664]	.655 [.617, 693]	.663 [.626, .701]	.663 [.626, .701]	.672 [.634, .709]	.643 [.605, .682]	.650 [.612, .688]	**.623** [.585, .662]	.652 [.614, .690]	.675 [.638, .712]
	NOM1	**.658** [.620, .696]	.665 [.627, .703]	.663 [.626, .701]	.677 [.639, .714]	.680 [.643, .717]	.693 [.656, .730]	.678 [.641, .716]	.668 [.631, .706]	.682 [.644, .719]	.690 [.653, .727]
	NOM2	**.608** [.569, .647]	**.667** [.629, .704]	**.660** [.622, .698]	**.635** [.596, .674]	.690 [.653, .727]	**.652** [.614, .690]	.672 [.634, .709]	.677 [.639, .714]	**.647** [.608, .685]	.685 [.648, .722]
Low	DISC	.600 [.561, .639]	.630 [.591, .669]	.627 [.588, .665]	.625 [.586, .664]	.643 [.605, .682]	.622 [.583, .660]	.627 [.588, .665]	.613 [.574, .652]	.622 [.583, .660]	.633 [.595, .672]
	NOM1	.597 [.557, .636]	.603 [.564, .642]	.618 [.579, .657]	.598 [.559, .638]	.645 [.607, .683]	.590 [.551, .629]	.623 [.585, .662]	.647 [.608, .685]	.618 [.579, .657]	.677 [.639, .714]
	NOM2	.598 [.559, .638]	.618 [.579, .657]	.615 [.576, .654]	.608 [.569, .647]	.648 [.610, .687]	.613 [.574, .652]	.633 [.595, .672]	.612 [.573, .651]	.628 [.590, .667]	.633 [.595, .672]

Note: Each estimate refers to 600 observations; bold figures indicate a rejection of the null of Nash mixed-equilibrium play at $\alpha = 5\%$; 95% confidence intervals (assuming i.i.d. observations) in brackets.

EXPERIENCE EFFECT – EVE'S SUCCESSIVE 15-ROUND BATCH ESTIMATES.

	Batch → Treatment ↓	1	2	3	4	5	6	7	8	9	10
High	DISC	**.352** [-.05, .75]	**1.180** [-1.60, 3.97]	**2.300** [-7.02, 11.6]	**2.300** [-7.02, 11.6]	**14.489** [-312, 341]	**.660** [-.38, 1.70]	**.893** [-.83, 2.61]	**.335** [-.04, .71]	**.974** [-1.02, 2.97]	∞ [n.a.]
	NOM1	**.714** [-.18, 1.61]	**.903** [-.39, 2.20]	**.849** [-.32, 2.02]	**1.521** [-1.54, 4.58]	**1.843** [-2.43, 6.11]	**7.470** [-50, 65]	**1.669** [-1.92, 5.26]	**1.030** [-.06, 2.63]	**2.053** [-3.11, 7.22]	**4.370** [-16, 25]
	NOM2	.171 [.03, .31]	**.739** [-.22, 1.70]	**.595** [-.10, 1.29]	.307 [.03, .59]	**2.305** [-4.19, 8.80]	**.468** [-.02, .96]	**.886** [-.39, 2.16]	**1.087** [-.69, 2.86]	0.409 [.00, .82]	**1.653** [-1.95, 5.26]
Low	DISC	**2.914** [-14.7, 20.6]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]
	NOM1	**1.196** [-1.85, 4.24]	**2.002** [-5.38, 9.38]	∞ [n.a.]	**1.338** [-2.33, 5.01]	∞ [n.a.]	**.816** [-.83, 2.47]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]	∞ [n.a.]
	NOM2	**.678** [-.51, 1.87]	**2.548** [-8.6, 13.7]	**1.838** [-4.34, 8.01]	**1.138** [-1.57, 3.85]	∞ [n.a.]	**1.603** [-3.26, 6.46]	∞ [n.a.]	**1.416** [-2.51, 5.34]	∞ [n.a.]	∞ [n.a.]

Note: Each estimate refers to 600 observations; bold figures indicate instances with maximal exploration, i.e., λ not significantly different from 0 at α = *5%*; 95% confidence intervals (assuming i.i.d. observations) in brackets; shaded cells characterize instances where the null of mixed-equilibrium play cannot be rejected at the 5% level (according to the entry probabilities reported in Appendix 5).

EXPERIENCE EFFECT – IBE'S SUCCESSIVE 15-ROUND BATCH ESTIMATES.

	Batch → Treatment ↓	1	2	3	4	5	6	7	8	9	10
High	DISC	4.161 [3.20, 5.12]	3.470 [2.65, 4.27]	3.295 [2.51, 4.08]	3.295 [2.51, 4.08]	3.126 [2.38, 3.87]	3.727 [2.86, 4.60]	3.579 [2.74, 4.42]	4.203 [3.23, 5.17]	3.542 [2.71, 4.37]	3.060 [2.32, 3.80]
	NOM1	2.709 [1.96, 3.46]	2.580 [1.86, 3.30]	2.612 [1.89, 3.34]	2.366 [1.70, 3.03]	2.307 [1.66, 2.96]	2.084 [1.49, 2.68]	2.336 [1.68, 2.99]	2.517 [1.82, 3.22]	2.278 [1.64, 2.92]	2.138 [1.53, 2.75]
	NOM2	4.167 [3.00, 5.33]	2.716 [1.94, 3.49]	2.856 [2.04, 3.67]	3.437 [2.47, 4.40]	2.267 [1.60, 2.93]	3.040 [2.18, 3.90]	2.615 [1.86, 3.37]	2.516 [1.79, 3.24]	3.155 [2.26, 4.05]	2.358 [1.67, 3.05]
Low	DISC	3.216 [2.48, 3.96]	2.693 [2.07, 3.32]	2.747 [2.11, 3.38]	2.774 [2.13, 3.42]	2.485 [1.91, 3.06]	2.830 [2.18, 3.48]	2.747 [2.11, 3.38]	2.973 [2.29, 3.66]	2.830 [2.18, 3.48]	2.640 [2.03, 3.25]
	NOM1	2.794 [2.03, 3.56]	2.667 [1.94, 3.39]	2.402 [1.75, 3.06]	2.762 [2.01, 3.51]	1.989 [1.44, 2.54]	2.926 [2.13, 3.73]	2.320 [1.69, 2.95]	1.966 [1.43, 2.51]	2.402 [1.75, 3.06]	1.578 [1.14, 2.02]
	NOM2	2.985 [2.15, 3.82]	2.587 [1.86, 3.11]	2.650 [1.91, 3.39]	2.779 [2.00, 3.56]	2.078 [1.49, 2.67]	2.681 [1.93, 3.43]	2.320 [1.67, 2.97]	2.714 [1.95, 3.47]	2.406 [1.73, 3.08]	2.320 [1.67, 2.97]

Note: Each estimate refers to 600 observations; 95% confidence intervals (assuming i.i.d. observations) in brackets.

Appendix 7: Estimation outcomes of two-parameter models.

EVE TWO-PARAMETER ESTIMATION RESULTS WHEN RELAXING SYMMETRY

Level	Structure		Session 1	Session 2	Session 3	Session 4	Pooled
	DISC	Cste	**1.218** [-.039, 2.474]	2.381 [1.070, 3.691]	**.225** [-1.173, 1.623]	1.269 [.117, 2.421]	1.266 [.889, 1.644]
		$\hat{\lambda}$	1.630 [1.490, 1.771]	1.715 [1.608, 1.822]	1.479 [1.367, 1.591]	1.526 [1.426, 1.625]	1.645 [1.612, 1.678]
			{$<10^{-3}$}	{$<10^{-3}$}	{.7221}°	{$<10^{-3}$}	{$<10^{-3}$}
High	NOM1	Cste	-1.158 [-1.881, -.435]	**.241** [-.576, 1.058]	**.527** [-.133, 1.187]	**.225** [-.786, 1.236]	.841 [.625, 1.057]
		$\hat{\lambda}$	1.535 [1.404, 1.667]	1.935 [1.760, 2.110]	2.063 [1.856, 2.271]	2.010 [1.765, 2.255]	2.106 [2.055, 2.157]
			{.0563}°	{.0382}	{.0377}	{.0920}°	{$<10^{-3}$}
	NOM2	Cste	**-.159** [-.688, .370]	-2.434 [-3.219, -1.648]	1.241 [.416, 2.067]	2.213 [1.485, 2.940]	-.827 [-.989, -.665]
		$\hat{\lambda}$	1.792 [1.635, 1.948]	1.231 [1.103, 1.359]	2.173 [2.015, 2.331]	2.347 [2.159, 2.535]	1.503 [1.466, 1.540]
			{.0958}°	{.0968}°	{$<10^{-3}$}	{$<10^{-3}$}	{$<10^{-3}$}
	DISC	Cste	2.039 [.811, 3.266]	.758 [.162, 1.354]	3.013 [1.788, 4.237]	3.003 [1.871, 4.136]	1.561 [1.282, 1.840]
		$\hat{\lambda}$	1.882 [1.805, 1.960]	1.587 [1.573, 1.601]	1.977 [1.854, 2.101]	2.076 [1.965, 2.188]	1.791 [1.775, 1.806]
			{.4354}°	{$<10^{-3}$}	{.1583}°	{$<10^{-3}$}	{$<10^{-3}$}
Low	NOM1	Cste	1.985 [.952, 3.017]	2.629 [1.217, 4.041]	1.364 [.189, 2.540]	**.148** [-.921, 1.217]	**.032** [-.228, .291]
		$\hat{\lambda}$	2.581 [2.398, 2.764]	2.527 [2.259, 2.795]	2.348 [2.177, 2.520]	2.080 [1.971, 2.189]	2.066 [2.030, 2.103]
			{.5058}°	{.3663}°	{.3970}°	{.1408}°	{$<10^{-3}$}
	NOM2	Cste	**-.095** [-1.017, .828]	-.117 [-.905, -.671]	**.035** [-.777, .848]	**.138** [-.995, 1.270]	**-.147** [-.368, .074]
		$\hat{\lambda}$	1.868 [1.730, 2.005]	2.115 [1.971, 2.260]	1.849 [1.722, 1.977]	1.916 [1.740, 2.093]	1.861 [1.829, 1.893]
			{.1238}°	{.3623}°	{.0277}	{.4062}°	{$<10^{-3}$}

Note: Each estimate refers to 1500 observations. Bold figures indicate instances with maximal exploration at α = *5%*; 95% confidence intervals in squared brackets; *p*-value of specification test in curly brackets; °: Non-rejection of the specification test at α = *5%*.

IBE Two-Parameter Estimation Results when Relaxing Symmetry

Level	Structure		Session 1	Session 2	Session 3	Session 4	Pooled
	DISC	Cste	6.201	6.266	6.279	5.936	6.398
			[5.784, 6.618]	[5.970, 6.553]	[6.099, 6.459]	[5.722, 6.150]	[6.351, 6.446]
		$\hat{\lambda}$	-2.872	-1.936	-2.841	-2.283	-2.284
			[-3.025, -2.718]	[-2.006, -1.867]	[-2.989, -2.692]	[-2.381, -2.186]	[-2.310, -2.257]
			$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$
High	NOM1	Cste	3.776	4.163	4.408	3.982	4.079
			[3.427, 4.125]	[3.990, 4.336]	[4.225, 4.592]	[3.645, 4.319]	[4.024, 4.135]
		$\hat{\lambda}$	-1.805	-1.903	-1.932	-1.884	-1.899
			[-1.847, -1.763]	[-1.947, -1.860]	[-1.982, -1.881]	[-1.985, -1.782]	[-1.909, -1.888]
			$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$
	NOM2	Cste	5.098	5.669	5.878	2.109	4.708
			[4.675, 5.522]	[5.345, 5.992]	[5.527, 6.230]	[1.993, 2.226]	[4.612, 4.804]
		$\hat{\lambda}$	-2.138	-2.447	-1.976	-1.815	-1.890
			[-2.265, -2.011]	[-2.526, -2.368]	[-2.038, -1.913]	[-1.853, -1.776]	[-1.915, -1.866]
			$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$
	DISC	Cste	4.020	5.680	3.578	5.275	4.498
			[3.914, 4.126]	[5.526, 5.834]	[3.409, 3.748]	[5.167, 5.383]	[4.470, 4.526]
		$\hat{\lambda}$	-1.827	-2.474	-1.710	-1.928	-1.898
			[-1.921, -1.734]	[-2.557, -2.390]	[-1.793, -1.628]	[-2.006, -1.851]	[-1.922, -1.874]
			$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$
Low	NOM1	Cste	3.043	3.475	4.223	4.886	3.839
			[2.843, 3.242]	[3.317, 3.633]	[4.015, 4.431]	[4.665, 5.108]	[3.799, 3.878]
		$\hat{\lambda}$	-1.115	-1.630	-1.759	-3.142	-1.823
			[-1.180, -1.050]	[-1.692, -1.567]	[-1.835, -1.683]	[-3.303, -2.981]	[-1.847, -1.801]
			$\{<10^{-3}\}$	$\{10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$
	NOM2	Cste	3.931	3.936	4.722	3.766	4.121
			[3.780, 4.081]	[3.749, 4.122]	[4.571, 4.872]	[3.609, 3.923]	[4.094, 4.149]
		$\hat{\lambda}$	-1.998	-1.689	-2.835	-1.825	-2.059
			[-2.089, -1.907]	[-1.770, -1.608]	[-2.897, -2.773]	[-1.889, -1.762]	[-2.075, -2.043]
			$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$	$\{<10^{-3}\}$

Note: Each estimate refers to 1500 observations. 95% confidence intervals in squared brackets; *p*-value of specification test in curly brackets.