

Learning in Two-Dimensional Beauty Contest Games: Theory and Experimental Evidence*

Mikhail Anufriev[†] John Duffy[‡] Valentyn Panchenko[§]

January 27, 2022

Abstract

We extend the beauty contest game to two dimensions: each player chooses two numbers to be as close as possible to certain target values, which are linear functions of the averages of the two number choices. One of the targets depends on the averages of both numbers, making the choices interrelated. We report on an experiment where we vary the eigenvalues of the associated two-dimensional linear system and find that subjects can learn the Pareto-optimal Nash Equilibrium of the system if both eigenvalues are stable and cannot learn it if both eigenvalues are unstable. Interestingly, subjects can also learn it if the system has the saddlepath property – with one stable and one unstable eigenvalue - but only if the one unstable eigenvalue is negative. We show theoretically that our results cannot be explained by homogeneous level- k models where all agents apply the same level k depth of reasoning to their choices, including the naïve learning model. However, our results *can* be explained by a mixed cognitive-levels model, including the adaptive learning model. We also run a horserace between many models used in the literature

*An earlier version of this paper circulated under the title “Planar Beauty Contests”. We thank the editor, Pietro Ortoleva, Associate Editor, two anonymous referees, Peter Bossaerts, Marcus Giannatelli, Frank Heinemann, Cars Hommes, Rosemarie Nagel, Jan Tuinstra and audiences at several conferences and workshops for helpful comments and suggestions on earlier drafts. Funding for this project was provided by the UC Irvine School of Social Sciences. Mikhail Anufriev acknowledges financial support from the Australian Research Council through Discovery Project DP140103501.

[†]University of Technology Sydney, Mikhail.Anufriev@uts.edu.au.

[‡]University of California, Irvine, duffy@uci.edu

[§]UNSW Business School, Sydney, v.panchenko@unsw.edu.au

with the winner being a simple mixed model with levels 0, 1, and equilibrium reasoning.

Keywords: Learning; Stability; Simultaneous Equation Systems; Beauty Contest; Complexity; Level- k ; Cognitive Hierarchy Theory.

JEL Classification: C30, C92, D83, D84.

1 Introduction

The process by which individuals learn an equilibrium has been the subject of a large theoretical and experimental literature (see Sargent, 1993; Fudenberg and Levine, 1998; Evans and Honkapohja, 2001; Camerer, 2003; Hommes, 2013). In this paper, we provide new experimental evidence on learning behavior in *coupled, linear multivariate* settings that are of interest to both game theorists and macroeconomists. The equilibrium and the dynamic properties of such settings are often expressed in terms of the eigenvalues of the corresponding system. To fix ideas, consider a 2 variable coupled linear (or linearized) economic model of the form

$$\mathbf{z}_t = \mathbf{M}\bar{\mathbf{z}}_t + \mathbf{d}, \quad \text{with } \mathbf{z}_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix} \quad \text{and} \quad \bar{\mathbf{z}}_t = \begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix}, \quad (1)$$

where \bar{a}_t and \bar{b}_t denote the population average of the expected time t values of the state variables a_t and b_t , and where the 2×2 matrix \mathbf{M} , and the 2×1 vector \mathbf{d} are exogenously given. Let \mathbf{I} denote the identity matrix and assume that $\mathbf{I} - \mathbf{M}$ is invertible. Let $\mathbf{z}^E = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{d}$ and observe that this is a unique point where $\mathbf{z}_t = \bar{\mathbf{z}}_t$. There are a number of economic settings that can be mapped into this basic framework and Appendix A provides two examples one from industrial organization and one from macroeconomics.¹

In the oligopoly market example, firms produce differentiated products for two interrelated markets and compete in prices. Eq. (1) describes, for each firm, the price-setting strategy in the game; its best response depends on the average prices in the two markets, \bar{a}_t and \bar{b}_t . Assuming common knowledge of rationality, \mathbf{z}^E is the symmetric Nash equilibrium profile. The properties of the learning dynamics,

¹Generally, there are many examples, especially in the macroeconomics literature, that use linearized multivariate dynamical systems. The *eigenvalues* of such systems are important for deciding whether the solution is determinate (i.e., locally unique) or not, see Blanchard and Kahn (1980). The examples include optimal growth models (e.g., Alvarez-Cuadrado et al., 2004) and monetary policy models (see Benhabib et al. 2001, where the steady states of the system may be of the *sink*, *saddle* or *source* varieties, the concepts that we use as the basis for the treatments in this paper).

when firms best respond to past average prices, depend on the eigenvalues of \mathbf{M} . The one-dimensional analogue of this model is the so-called “beauty contest” (BC) game, and has been studied extensively. It has been shown that both the *magnitude* and the *sign* of the slope (that is, the eigenvalue) of the *one*-dimensional map matter for whether subjects play the Nash equilibrium from the outset or whether and how fast they converge toward it.²

In the macro example from the monetary policy literature, the state variables are inflation and the output gap and Eq. (1) describes how these variables depend on the expectations formed by a “representative agent” (in the macro literature instead of \bar{z}_t one would typically write z_t^e). Point z^E is then the unique *rational expectations* equilibrium (REE). As coordination on this equilibrium requires strong assumptions, the learning literature in macroeconomics assumes that the representative agent predicts \bar{z}_t using past values of the state variables. Convergence to the REE depends on this learning process. When expectations are naïve, that is, when $z_t^e = z_{t-1}$, convergence depends on the eigenvalues of the matrix \mathbf{M} .

Motivated by these considerations, the main question we ask in this paper is *whether the eigenvalues of the reduced form, two-dimensional multivariate linear system (1) matter for the ability of human subjects to learn the equilibrium*. We address this question by developing an extended, multi-dimensional version of the BC game. Subjects are asked to guess *two* numbers, a_t and b_t , instead of the usual one, in a coupled system (1). We refer to our game as the “Two Dimensional Beauty Contest” or the 2DBC game. By contrast with the univariate version, in our coupled multivariate system, guesses about the realization of one variable can matter for realizations of the other variable, and *vice versa*, which makes the task of coordinating on equilibrium behavior even more complex than in the standard one-dimensional case. It is difficult to know, *a priori*, how subjects will react to the additional complexity of the 2DBC. On the one hand, complexity might slow or prevent learning of the Nash equilibrium,

²The importance of the magnitude of the slope has been shown already in Nagel (1995) and Ho et al. (1998). Sutan and Willinger (2009) and Heemeijer et al. (2009) stress the impact of the sign.

as subjects struggle to understand how their guesses for one variable impact on their guesses for the other. On the other hand, the greater complexity of the 2DBC might promote greater introspection; subjects might respond to the greater complexity by thinking harder about the problem at hand, in the extreme perhaps even applying “fixed–point reasoning” to directly solve for the steady state.³

Our experiment involves four treatments that vary the eigenvalues of the matrix \mathbf{M} . To make the treatments as comparable as possible, we only change one or two elements of this matrix between treatments, and we also adjust the vector \mathbf{d} so that all treatments have the same equilibrium \mathbf{z}^E . To address the question of learning, we ask subjects to play the 2DBC game repeatedly for 15 periods and we examine the dynamics of both variables, a_t and b_t , over time. Consistent with the experimental BC literature, there are no explicit intertemporal dynamical linkages from one period to the next in (1). In such settings, *learning agents* will likely condition their guesses on the *past* history of play. For instance, if agents’ guesses for each variable are equal to the previous period values (as under the best response dynamics or, equivalently, naïve expectations), then model (1) is effectively a first-order dynamical system governed by matrix \mathbf{M} with the steady state \mathbf{z}^E . We label our treatments with reference to such naïve best response dynamics, using the standard terminology of dynamical systems.⁴ In treatment **Sink**, both eigenvalues of \mathbf{M} are stable, so that in theory, all trajectories converge to the steady state. In treatment **Source**, both eigenvalues are unstable and so all trajectories diverge. In treatments **SaddleNeg** and **SaddlePos**,

³The evidence on the relationship between task complexity and task performance is mixed. The majority of studies surveyed by Liu and Li (2011) indicate a negative relationship, but several studies show a positive relationship and there is even some evidence for an inverted U-shaped relationship between task complexity and task performance. Oprea (2020) reports that complexity costs (i.e., differences in willingness to pay to avoid more complex tasks) in individual decision-making experiments are heterogeneous across individuals but do decrease with experience.

⁴See, e.g., Azariadis (1993). Linear systems in discrete time, such as (1), have a unique steady state the stability of which depends on the eigenvalues. The eigenvalue is *stable*, if its absolute value is less than 1, and is *unstable* otherwise. The dynamics are *globally stable*, that is all trajectories converge if and only if all eigenvalues are stable. If all eigenvalues are unstable, then all trajectories diverge (except for the steady state). If there are both stable and unstable eigenvalues, the *saddle-path* is defined as any trajectory starting in the linear space formed by the *eigenvector(s)* corresponding to the stable eigenvalue(s). The saddle-paths converge, all other trajectories diverge.

one eigenvalue is stable and the other is unstable. Theoretically, in both cases, almost all trajectories diverge. The difference between these treatments is in the sign of the unstable eigenvalue, which is negative in **SaddleNeg** and positive in **SaddlePos**.

We find that subjects are able to learn the equilibrium steady state of the system in the **Sink** treatment. Remarkably, they are also able to learn the equilibrium steady state in the **SaddleNeg** treatment. By contrast, in the other two cases, the **SaddlePos** and **Source**, subjects are unable to learn the interior steady state equilibrium. Our findings thus suggest that steady states exhibiting the saddle-path property *can* be learned by agents who do not begin a process of social interaction with rational expectations. This result is important since saddle-path steady states are commonly used in macroeconomic analysis,⁵ but the empirical relevance of such systems can be questioned because they require substantial computational and coordination efforts among participants to be achieved. At the same time, we find that steady state convergence is not a *general* property of saddle-path stable solutions.

We also ask whether a single learning model can explain the different behavior we observe across treatments. We generalize the level- k model of Stahl and Wilson (1994) and the cognitive hierarchy model of Camerer et al. (2004) to our multivariate, dynamic setting. We express the convergent properties of those models in terms of matrix \mathbf{M} and compare the predictions with our experimental data. We also use individual-level data to estimate and compare the fit of these and several other models to our data including the structural level- k model of Gill and Prowse (2016) and the EWA model of Camerer and Ho (1999).

Our paper contributes to a large literature on the Keynesian Beauty Contest game, first explored experimentally by Nagel (1995); see Mauersberger and Nagel (2018) for a recent overview. The classic BC game can be viewed as a univariate version of (1).

⁵The saddle-path environment is particularly attractive since the reaction of the economy to shocks to fundamentals can be uniquely determined only in this environment. In contrast, when the steady state has the sink property, there are infinitely many paths by which the system can adjust following a perturbation.

In that game, each player simultaneously guesses a number in the interval $[0, 100]$, and the average of all guesses, \bar{a}_t , determines the payoff relevant target value, $a_t = p\bar{a}_t + d$. When $p \in (0, 1)$ and $d = 0$, the unique, dominance solvable Nash equilibrium is for all players to guess 0. However, this outcome is rarely observed in the first round of play, $t = 1$. Instead, subjects tend to apply only a limited number of steps of iterated elimination of dominated strategies.⁶ The experimental evidence suggests that there are sizeable fractions of level-0, 1, and 2 types, as well as subjects playing the Nash equilibrium.⁷ In repeated play of the one-dimensional BC game, players use the history of outcomes and continue to apply level- k reasoning to prior period winning numbers. The dynamics converge to the Nash equilibrium for $0 < p < 1$ both in the classic BC game where $d = 0$, and in the game with an “interior” equilibrium, where $d \neq 0$, as in Güth et al. (2002). The dynamics diverge for $p > 1$.

Convergence to REE has also been studied in the related literature on learning-to-forecast (LtF) experiments (Hommes et al., 2005; Hommes, 2011; Anufriev and Hommes, 2012) that seek to understand how human subjects form expectations in self-referential systems. Subjects forecast an endogenous variable that evolves, given the average forecast, \bar{a}_t^e , as $a_t = p\bar{a}_t^e + d$. In contrast to the BC experiments, the LtF literature typically uses a *limited information* setup, where subjects do not know the target equation (Mirdamadi and Petersen, 2018 and Kryvtsov and Petersen, 2021 are exceptions) and are given only qualitative information about whether higher average forecasts results in higher or lower target values. Recent LtF experiments, e.g., Pfajfar and Žakelj, 2016; Assenza et al., 2021; Mauersberger, 2021; Levelt et al., 2021, study expectations in coupled systems arising out of various versions of the New Keynesian model of monetary policy. The aim of these studies is to understand whether certain

⁶“Level-0” players are not strategic and play randomly. Other players begin with an initial reference point \hat{a} as the average choice of “level-0” players. “Level-1” players presume that all other participants are “level-0” and choose the best response, $p \times \hat{a}$, as their guess. “Level-2” types best respond to the “level-1” choice and guess $p^2 \times \hat{a}$. Generalizing, a “level- k ” type guesses $p^k \times \hat{a}$.

⁷Alaoui and Penta (2015, 2021) endogenize the depth of reasoning using a cost-benefit approach. The cost is the cognitive effort needed to employ greater depths of reasoning, while the benefit is the value of additional reasoning as reflected in the payoff incentives of the game.

monetary policy rules or types of central bank communication can move the system to a more desirable equilibrium or reduce the volatility of forecasts. By contrast, in this paper, we are not interested in considering a specific model (e.g., the New Keynesian model) and the role of monetary policy; rather we are interested in the more fundamental question of whether and how subjects can learn the equilibrium of coupled, multivariate systems that are commonplace in economic modeling. Toward that end, we provide our subjects with full information about the data generating process and employ the contemporaneous forecasting setup, as used in the experimental BC literature.

Our paper is also related to theoretical work exploring whether saddle-path stable solutions can be learned under adaptive learning dynamics, e.g., Evans and Honkapohja (2001), Ellison and Pearlman (2011). This work shows that steady state solutions with the saddle-path property can be learned if agents have perceived laws of motion (PLMs) that are correctly specified.⁸ By contrast, in this paper, we do not endow subjects with such specifications, though they do know the data generating equations of the system. We show that simple adaptive learning dynamics, initialized according to an uninformative prior belief that the means of both variables will begin at the midpoint of the guessing interval, closely track the behavior of the subjects.

2 Experimental design

In each experimental session, a fixed group of N individuals participate in T repetitions of the 2DBC game. In each period, each player i submits a pair of numbers, $(a_{i,t}, b_{i,t})$, with each number restricted to lie in the interval $[0, 100]$.⁹ Based on these

⁸For instance, if agents use a minimal state variable representation for the PLM, i.e., that the two target variables are constants so that $E_t(a_t) = a$ and $E_t(b_t) = b$, then, as Evans and Honkapohja (2001) have shown, the system (1) is e-stable if the eigenvalues of matrix $\mathbf{M} - \mathbf{I}$ are negative.

⁹The bounded interval helps to minimize the effects of extreme guesses (mistakes or outliers) and facilitates comparisons with the one-dimensional BC game. From our analysis, we expect that our convergence results would continue to hold without the bounded interval, though learning might take longer and be noisier.

Treatment	Parameters	PONE	Boundary NE	Eigenvalues
Sink	$\mathbf{M} = \begin{pmatrix} 2/3 & 0 \\ -1/2 & -1/2 \end{pmatrix}, \mathbf{d} = \begin{pmatrix} 30 \\ 75 \end{pmatrix}$	$\begin{pmatrix} 90 \\ 20 \end{pmatrix}$	-	$\frac{2}{3}, -\frac{1}{2}$
SaddleNeg	$\mathbf{M} = \begin{pmatrix} 2/3 & 0 \\ -1/2 & -3/2 \end{pmatrix}, \mathbf{d} = \begin{pmatrix} 30 \\ 95 \end{pmatrix}$	$\begin{pmatrix} 90 \\ 20 \end{pmatrix}$	-	$\frac{2}{3}, -\frac{3}{2}$
SaddlePos	$\mathbf{M} = \begin{pmatrix} 2/3 & 0 \\ -1/2 & 3/2 \end{pmatrix}, \mathbf{d} = \begin{pmatrix} 30 \\ 35 \end{pmatrix}$	$\begin{pmatrix} 90 \\ 20 \end{pmatrix}$	$\begin{pmatrix} 90 \\ 0 \end{pmatrix}, \begin{pmatrix} 90 \\ 100 \end{pmatrix}$	$\frac{2}{3}, \frac{3}{2}$
Source	$\mathbf{M} = \begin{pmatrix} 3/2 & 0 \\ -1/2 & -3/2 \end{pmatrix}, \mathbf{d} = \begin{pmatrix} -45 \\ 95 \end{pmatrix}$	$\begin{pmatrix} 90 \\ 20 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 38 \end{pmatrix}, \begin{pmatrix} 100 \\ 18 \end{pmatrix}$	$\frac{3}{2}, -\frac{3}{2}$

Table 1: Characteristics of the four experimental treatments.

“guesses”, the group averages $\bar{a}_t = N^{-1} \sum_i a_{i,t}$ and $\bar{b}_t = N^{-1} \sum_i b_{i,t}$ are computed. The linear, multivariate model (1) determines the “target” values, that is

$$\begin{pmatrix} a_t^* \\ b_t^* \end{pmatrix} = \mathbf{M} \begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} + \mathbf{d} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad (2)$$

where the 2×2 matrix \mathbf{M} and 2×1 vector \mathbf{d} depend on the treatment and are known to all participants. The payoff to participant i in each period is given by

$$\pi_{i,t} = \frac{500}{5 + |a_{i,t} - a_t^*| + |b_{i,t} - b_t^*|} \quad (3)$$

points. Thus, participants are motivated to submit their guesses as close as possible to the two target values. Guessing both targets exactly would bring a maximum reward of 100 points. Deviations from the two target values decrease each participant’s payoffs by an equal amount.¹⁰

We study the behavior of experimental subjects in four treatments that differ in the matrix \mathbf{M} and the vector \mathbf{d} . The treatments and the corresponding parameters

¹⁰We use a distance payoff function as opposed to the tournament structure of the classical BC. The distance payoffs do not change behavior in the BC, see, e.g., Güth et al. (2002). The LtF experiments typically use inverse quadratic distance function truncated at 0 to avoid negative payoffs. Our *hyperbolic* payoff function, involving the inverse of the distance errors, penalizes the smallest deviations from the targets, giving participants robust incentives to guess the targets precisely. Also truncation of payoffs is not necessary, as the function asymptotically approaches 0. Similar functions have been used in Adam (2007) and Assenza et al. (2021).

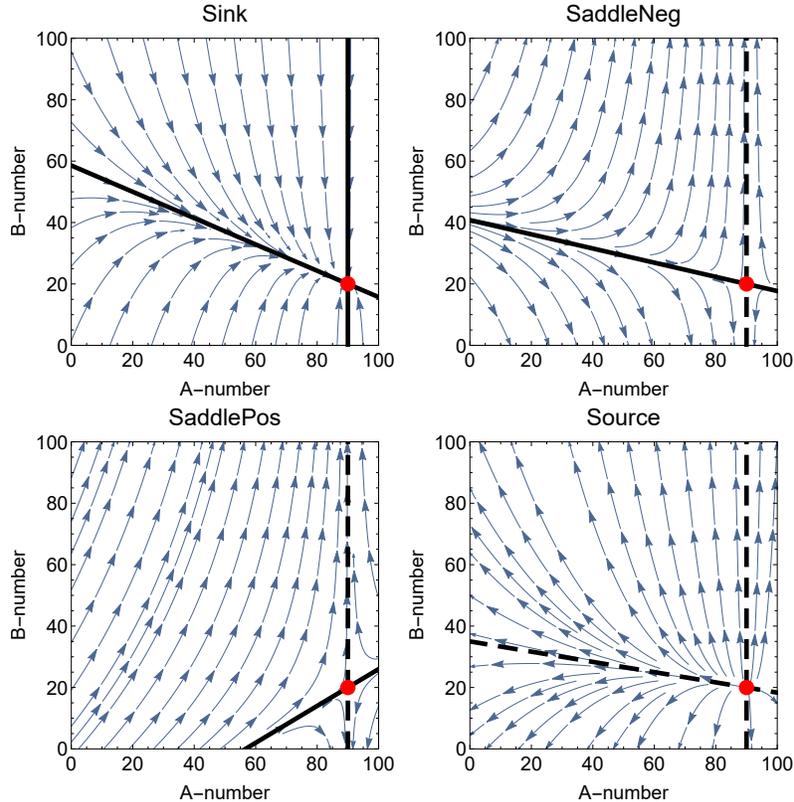


Figure 1: Phase plots in four treatments under best response dynamics. The two straight lines indicate the eigenvectors, stable (solid lines) and unstable (dashed lines); they intersect at the steady state (90, 20). The arrows show the direction of trajectories. As time is discrete in our experiment, the trajectories jump across the eigenvectors with negative eigenvalues.

are presented in the first two columns of Table 1. Fig. 1 illustrates the dynamics in the treatments under the simple baseline assumption that all subjects play a best response to the previous period target. In all treatments, we use a lower triangular matrix by setting $m_{12} = 0$ in (2). This makes our system coupled, but, at the same time, simple enough that subjects could possibly solve for the fixed point. The key difference across our four treatments is in the location of the *eigenvalues* of \mathbf{M} , i.e., its diagonal elements (these are shown in the last column of Table 1). In the first three treatments, we set $m_{11} = 2/3$, so that the first eigenvalue is stable (see footnote 4). With this specific choice for the parameter, the first (uncoupled) equation of our two-dimensional multivariate system is a version of the classic, $2/3$ of the average, BC game (albeit with an interior solution).

Depending on the second eigenvalue m_{22} , we then have three treatments. In the **Sink** treatment, $m_{22} = -1/2$, and so both eigenvalues are stable. In this treatment every trajectory in Fig. 1 converges to the steady state, indicated by the large dot, and so we predict that the steady state would eventually be learned by agents. By contrast, in the two treatments with the “saddle-path property”, one eigenvalue is stable while the other is not. The steady state is then generically unstable; the system converges *only if* it is exactly on the saddle-path, which is the “stable” eigenvector depicted by the solid line in Fig. 1. In the **SaddleNeg** treatment, $m_{22} = -3/2$, while in the **SaddlePos** treatment, $m_{22} = 3/2$. Note that the only difference between these treatments is in the *sign* of the unstable eigenvalue. In particular, the eigenvalues in these two saddle treatments have the same absolute values. The negative eigenvalue introduces “negative feedback” to the b -number strategy in the **SaddleNeg** treatment: that is, higher guesses imply a lower target. Instead, when the eigenvalue is positive in the **SaddlePos** treatment, the b -number guess generates “positive feedback” with higher guesses implying a larger target.¹¹ For the **Source** treatment, we set $m_{11} = 3/2$ and $m_{22} = -3/2$. As both eigenvalues are unstable, we do not predict that agents will converge to the steady state in this case except in the knife-edge case where they all start out there. This treatment can be viewed as the multivariate analogue of the univariate BC game with $p > 1$, see Nagel (1995); Ho et al. (1998).

Given our parametrizations, the matrix $\mathbf{I} - \mathbf{M}$ is invertible in all four treatments. It can be directly checked that if every participant submits the guesses given by

$$\mathbf{z}^E = (\mathbf{I} - \mathbf{M})^{-1} \mathbf{d}, \quad (4)$$

both targets a^* and b^* coincide with the corresponding guesses. Since there are no profitable deviations by any individual subject, (4) defines a Nash equilibrium.¹²

¹¹Positive or negative feedback is typically a property of univariate systems. We use this terminology to characterize the feedback for b -guesses only. While we could characterize the **SaddlePos** system as one of pure positive feedback (as both eigenvalues are positive), the other three systems have mixed feedback properties as one eigenvalue is negative while the other is positive.

¹²The restriction $[0, 100]$ on the guesses motivated us to select the particular Nash equilibrium of

Depending on the treatment, the restriction on the guesses may lead to other, non-interior, or boundary Nash equilibria of the game. These equilibria are shown in the fourth column of Table 1; see Appendix B for formal proofs. In any boundary equilibrium, the target value for at least one variable will be outside of the $[0, 100]$ range. This implies (much) lower payoffs in the boundary equilibria than in the interior equilibrium, and so we refer to \mathbf{z}^E as Pareto-Optimal Nash Equilibrium (PONE).

The experimental sessions were conducted in the Experimental Social Science Laboratory at the University of California, Irvine (UCI) using undergraduate student subjects. We conducted four sessions of each of our four treatments. For each session, a group of 10 subjects played the same 2DBC game for 15 consecutive periods. In each period, participants were incentivized to independently choose their two numbers to be as close as possible to the target values via payoff function (3). As our focus is on the convergence properties of the dynamic environment, we had the same 10 subjects play all 15 periods together.¹³ Every subject participated in one session only and thus we report data from $4 \times 4 \times 10 = 160$ subjects.

At the start of each session, the Instructions (see Appendix C) were read aloud and the procedure by which the target numbers were determined in each period was carefully explained to subjects. We also projected the equations determining the two target values on a screen for all subjects to see. After the instructions were read, subjects had to successfully answer several control questions before they were able to move on to the main experiment. At the end of the experiment, subjects completed a brief survey. The experiment was computerized using the zTree software (Fischbacher, 2007). In every period, the upper part of the main decision screen reminded subjects how the target values a^* and b^* were determined based on their choices (i.e., system (2) was presented, though in a simple, non-matrix way). In the middle part of their

$a^E = 90$ and $b^E = 20$. We wanted to have an interior equilibrium which is also far enough from the focal point of the admissible guessing plane, $(50, 50)$, to observe some learning. At the same time, we wanted to have simple numbers so that the participants could easily solve for the equilibrium.

¹³Rematching subjects each period would make past information less relevant and thus interfere with learning. This approach is standard in this literature, e.g., Nagel (1995).

screen, subjects entered their pair of numbers for the given period, one “ a -number” and one “ b -number”. Subjects could also click on an icon to get access to an online calculator. See the screen shots in Appendix F1.

After all 10 subjects submitted their guesses, the computer program determined the target values for the period and the subjects’ payoffs in points according to the payoff function (3). Period t ended with a second results screen reminding subjects of their submitted pair of numbers and informing them of the group average values for the two numbers, \bar{a}_t and \bar{b}_t , the two target numbers, a_t^* and b_t^* , and the points that the subject had earned for the period. Except for the very first period, subjects could see a history of all previous outcomes (their chosen numbers, the averages for both numbers, the target values, and the points they earned) in the lower part of the main decision screen for each period. Subjects’ total point earnings, from all 15 periods of a session, were converted into dollars at the fixed and known rate of 100 points = \$1. Thus, subjects could earn a maximum of \$15 from their guesses. In addition, subjects were given a \$7 show-up payment for a total maximum of \$22. Actual total earnings (including the show-up payment) varied with the treatment but averaged \$12 across all four treatments for an approximately 75-minute experiment.

3 Results

In this section, we provide an overview of the main results from the experiment, illustrating the dynamics of average guesses and individual choices. Further findings including an analysis of first period choices, can be found in Appendix F2.

Dynamics of Averages. The four panels of Fig. 2 show a representative session (session 1) of the evolution of the averages, \bar{a} and \bar{b} , across all four treatments. The time evolution of \bar{a} (the thick red line) and of \bar{b} (the thin blue line) are graphed in relation to dashed lines indicating the PONE levels, $a^E = 90$ and $b^E = 20$. We observe convergence in both the a and b numbers to those levels in the **Sink** and **SaddleNeg**

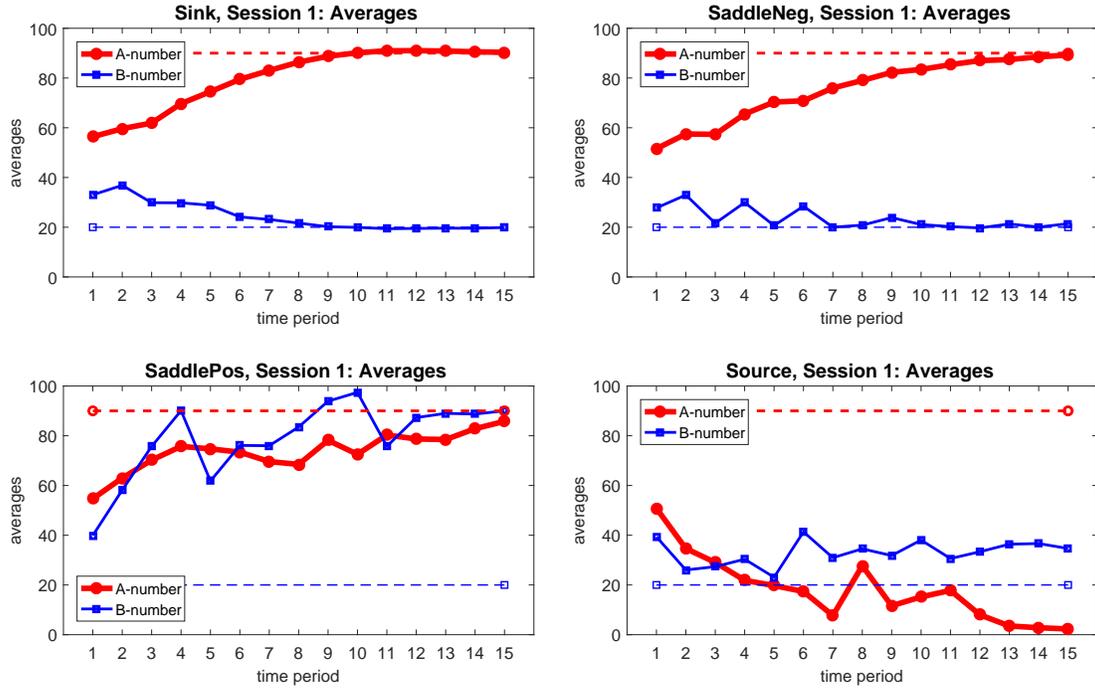


Figure 2: Typical dynamics of the average values, \bar{a} and \bar{b} , observed in four different treatments of the experiment. The dashed lines indicate the PONE.

treatments. By contrast, in the **SaddlePos** and **Source** treatments, the a and b numbers approach payoff-dominated Nash equilibria instead – see Table 2. Overall, the dynamics in the four sessions of each of the four treatments are very similar to those of session 1; see Figs. F.6 to F.9 in Appendix F2.

The most remarkable finding is the contrast in dynamics between two saddle treatments. Within the time frame of our experiment, participants are able to learn a steady state with the saddle-path property under negative feedback quickly. By contrast, participants *are not* able to learn a steady state with the saddlepath property under positive feedback.

Table 2 shows the average guesses for the a - and b -numbers in the first and the last periods of each session of the experiment (as well as the treatment mean, standard deviation, and median). In parentheses, we also report the deviations of the average guesses from the PONE. From this table we can make the following observations.

Treatment	Session	a -number ($a - a^E$)		b -number ($b - b^E$)		Payoff		
		period 1	period 15	period 1	period 15	period 1	period 15	periods 1–15
Sink	Sess. 1	56.4 (-33.6)	90.3 (0.3)	33.0 (13.0)	19.9 (-0.1)	23.38	92.68	923.43
	Sess. 2	65.3 (-24.7)	84.0 (-6.0)	37.6 (17.6)	30.1 (10.1)	16.97	35.70	549.32
	Sess. 3	49.8 (-40.2)	80.4 (-9.6)	32.6 (12.6)	28.5 (8.5)	23.33	41.55	734.63
	Sess. 4	58.0 (-32.0)	82.2 (-7.8)	36.7 (16.7)	23.2 (3.2)	24.38	62.05	758.11
	Mean	57.4 (-32.6)	84.2 (-5.8)	35.0 (15.0)	25.4 (5.4)	22.02	58.00	741.37
	Std Dev	17.7	14.4	20.0	17.4	14.15	25.77	191.90
	Median	60.0 (-30.0)	86.9 (-3.1)	30.0 (10.0)	21.5 (1.5)	17.84	48.80	727.24
SaddleNeg	Sess. 1	51.6 (-38.4)	89.2 (-0.8)	27.8 (7.8)	21.4 (1.4)	14.18	61.39	581.20
	Sess. 2	60.6 (-29.4)	89.7 (-0.3)	39.1 (19.1)	20.0 (-0.0)	11.69	92.59	888.13
	Sess. 3	51.0 (-39.0)	89.9 (-0.1)	37.5 (17.5)	20.0 (-0.0)	17.02	95.06	791.62
	Sess. 4	56.9 (-33.1)	90.4 (0.4)	19.9 (-0.1)	19.9 (-0.1)	13.40	93.32	892.81
	Mean	55.0 (-35.0)	89.8 (-0.2)	31.1 (11.1)	20.3 (0.3)	14.07	85.59	788.44
	Std Dev	22.3	0.5	21.5	1.5	11.81	15.79	151.13
	Median	56.0 (-34.0)	89.9 (-0.1)	29.0 (9.0)	20.0 (0.0)	12.47	92.12	833.75
SaddlePos	Sess. 1	54.7 (-35.3)	85.8 (-4.2)	39.9 (19.9)	90.0 (70.0)	10.85	13.19	189.63
	Sess. 2	61.5 (-28.5)	86.4 (-3.6)	40.9 (20.9)	79.5 (59.5)	14.40	21.16	210.00
	Sess. 3	50.6 (-39.4)	88.3 (-1.7)	49.7 (29.7)	99.8 (79.8)	15.98	10.70	165.70
	Sess. 4	58.2 (-31.8)	88.8 (-1.2)	41.4 (21.4)	99.9 (79.9)	10.86	10.80	168.94
	Mean	56.3 (-33.7)	87.3 (-2.7)	43.0 (23.0)	92.3 (72.3)	13.02	13.96	183.57
	Std Dev	20.5	3.1	23.2	24.9	10.88	6.64	44.81
	Median	58.8 (-31.2)	88.0 (-2.0)	46.7 (26.7)	100.0 (80.0)	9.65	10.95	175.30
Source	Sess. 1	50.8 (-39.2)	2.3 (-87.7)	39.4 (19.4)	34.7 (14.7)	15.65	9.04	160.09
	Sess. 2	47.7 (-42.3)	12.2 (-77.8)	52.8 (32.8)	35.6 (15.6)	6.11	11.88	162.27
	Sess. 3	38.2 (-51.8)	10.7 (-79.3)	34.6 (14.6)	37.3 (17.3)	9.56	11.21	150.24
	Sess. 4	42.4 (-47.6)	4.6 (-85.4)	36.9 (16.9)	34.7 (14.7)	11.06	9.39	162.51
	Mean	44.8 (-45.2)	7.5 (-82.5)	40.9 (20.9)	35.6 (15.6)	10.59	10.38	158.78
	Std Dev	21.0	13.5	23.2	6.0	11.47	2.75	34.92
	Median	45.8 (-44.3)	0.0 (-90.0)	45.0 (25.0)	35.0 (15.0)	7.93	10.04	153.41
PONE		90 (0)	90 (0)	20 (0)	20 (0)	100.00	100.00	1500.00

Table 2: Statistics on average guesses and deviations (in parentheses) from the PONE (90, 20) for both the a and b numbers and subjects' payoffs for each session of the experiment.

First and most importantly, there is a consistent pattern in the dynamics of all four sessions of each treatment, comparable with the representative session 1 shown in Fig. 2. Specifically, in the first period the average a - and b -guesses in all treatments are significantly different from the PONE.¹⁴ We observe convergence to the PONE (90, 20) in the **Sink** and **SaddleNeg** treatments and we do not observe convergence to this equilibrium in the **SaddlePos** and **Source** treatments.¹⁵ In the **SaddlePos** treatment, the b -guesses end up even further away from the equilibrium than they

¹⁴For each treatment, using the $4 \times 10 = 40$ individual a -guesses and a two-sided t-test, we reject the null hypothesis that the mean guess is initially 90. We also reject the null hypothesis that the mean initial b -guess is 20 for each treatment. All p-values are less than 0.0025.

¹⁵Using the four session-level observations (as these are independent) for the **Sink** and **SaddleNeg** treatments, we do not reject the null hypothesis that the average a - and b -guesses in period 15 are different from 90 and 20, respectively. However, we reject this same hypothesis for a - or b -guesses of the **SaddlePos** and **Source** treatments. We use two-sided t-tests. The p-values for the tests for the **Sink**, **SaddleNeg**, **SaddlePos** and **Source** treatments for a -numbers are 0.076, 0.488, 0.034, 0.0001, respectively, and for the b -numbers are 0.106, 0.457, 0.001, 0.0001, respectively.

were in the initial period. Similarly, the a -guesses in the **Source** treatment move further away from the equilibrium. Finally, the b -guesses in the **Source** treatment move only very slightly towards the PONE. Second, the heterogeneity in guesses (as measured by the standard deviations) is similar across all treatments for first period choices and typically decreases to a very small number in the two converging cases.

Third, the differences in behavior between treatments translate into differences in subjects' payoffs. The last three columns of Table 2 show the average payoffs per experimental session (in points) for periods 1, 15, and over all 15 periods.¹⁶ The largest total payoffs are consistently achieved in **Sink** and **SaddleNeg** treatments where the dynamics converged. The payoffs in **SaddlePos** and **Source** treatments are much lower. When we look at how payoffs changed during the experiment, we observe that the initial payoffs did not improve in the **SaddlePos** and **Source** treatments. In the two other convergent treatments, the payoffs improved. Interestingly, the initial payoffs in the **Sink** treatment were almost twice as large as in the **SaddleNeg** treatment, indicating that the one shot game was much easier in the **Sink** treatment than in the **SaddleNeg** treatment (as well as in all other treatments). However, the convergence observed in the **SaddleNeg** treatment was apparently even quicker which allowed subjects to earn the largest total payoff across all four treatments.

Individual Choices and Levels. We illustrate the dynamics of individual guesses in two ways. First, the eight panels of Fig. 3 show, for all four treatments, the cumulative frequencies of individual choices in periods $t = 1$ (magenta thick line), $t = 5$ (blue dotted line), $t = 10$ (green thin line), and $t = 15$ (black dashed line) for both guesses: the a -number (top panels) and the b -number (bottom panels). The PONE values are indicated by vertical (red) lines. Notice that the variance of individual choices decreases over time, with the greatest reduction occurring during the first 5 periods.

¹⁶One point corresponded to 1 US cent in the experiment. Recall from Eq. (3) that the maximum payoff per period was equal to 100 points, and that even small deviations in predictions from target values were costly. For instance, a prediction error of 1 for only one of the two targets reduces the maximum possible payoff for the round to 83.3 cents, see the payoff Table presented to subjects in the instructions, Appendix C.

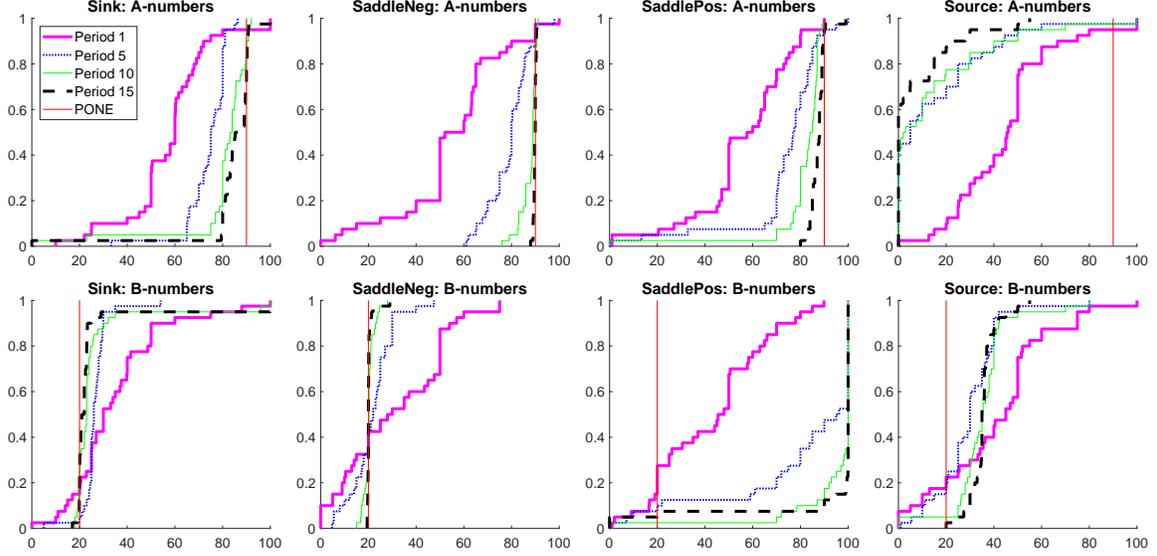


Figure 3: Cumulative frequencies of individual choices in periods 1, 5, 10, and 15 for four treatments (from left to right) for a -numbers (*top*) and b -numbers (*bottom*).

In the **Sink** and **SaddleNeg** treatments, individual choices converge to the steady state and stay relatively close to each other. In **SaddlePos** and **Source** treatments subjects' choices are more dispersed even in the last period of the experiment.

Second, we *classify* subjects' play using level- k reasoning extended to the dynamic environment. For each session s , we define the level-0 choices for round $t > 1$ as the averages of the previous round, $\mathbf{z}_t^s(0) := \bar{\mathbf{z}}_{t-1}^s$. The higher level- k choices are best responses to level $k - 1$ choices, where we assume that subjects use the same level in their best response for both the a and b numbers. It means that recursively the level k choices are defined as $\mathbf{z}_t^s(k) = \mathbf{M}\mathbf{z}_t^s(k-1) + \mathbf{d}$. We compare now the participant's choice in round t of the experiment, $(a_{i,t}, b_{i,t})$, with such a benchmark choice for his/her session, $\mathbf{z}_t^s(k) = (a_t^s(k), b_t^s(k))$, using the distance function

$$d_t(k) := |a_{i,t} - a_t^s(k)| + |b_{i,t} - b_t^s(k)|, \quad (5)$$

and identify the k at which this function is minimized among all levels $0 \leq k \leq 4$ and $k = E$ which denotes the equilibrium play $\mathbf{z}^E = (90, 20)$. We consider seven periods,

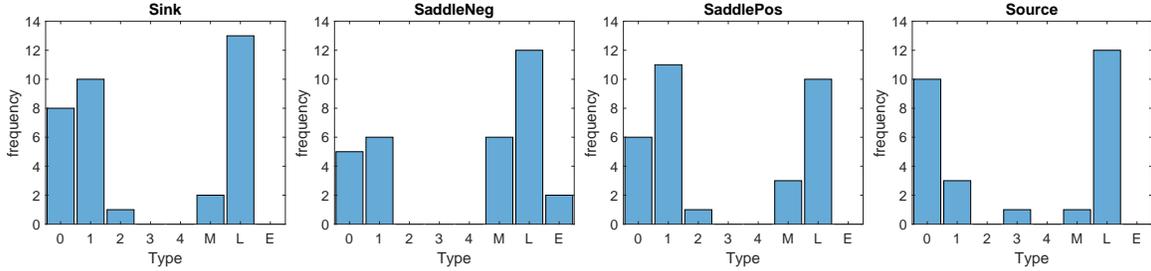


Figure 4: Frequencies of different types of subjects in the four treatments.

from 2 to 8, and classify every subject as follows:¹⁷

- “Level- k ” type ($k = 0, 1, 2, 3, 4$): if at least in 5 periods, level- k was identified;
- “Equilibrium” type: if at least in 5 periods, $k = E$ was identified;
- “Learning” type: none of the above, and the median k in periods 2–5 is smaller than the median k in periods 5–8 (E is excluded from the median calculations);
- “Mixing” type: none of the above, and at least in 6 periods levels $k = 0$ or $k = 1$ were identified.

The distributions of types for each of the treatments are shown in Fig. 4. Remarkably, these distributions are similar across treatments. Almost half of the participants (20 out of 40 in the **Sink** and **SaddlePos** treatments, 17 in the **SaddleNeg** and 14 in the **Source**) are classified as either consistent level 0 or 1 types, or “mixing” types, who go back and forth. About one quarter of all participants, (between 10 in the **SaddlePos** and 13 in the **Sink**) are classified as “learning” types. With the exception of the **SaddleNeg** treatment, there are no “equilibrium” types.

We summarize our findings as follows:

1. First-period choices are heterogeneous with concentrations at the mid-point of the guessing interval, 50. In subsequent periods, heterogeneity is reduced, with the majority of subjects choosing numbers close to level-0 (the average choice of the previous period) or level-1. Some subjects increase their level over time.

¹⁷See Table 7 in Appendix D for individual levels during first eight periods and resulting classification. We look only at periods 2-8 because after convergence (or divergence) all the levels are close to each other, and we do not use period 1 in the classification, as behavior then is not dynamic.

2. The dynamics converge to the PONE only in the **Sink** and **SaddleNeg** treatments. Convergence is fastest in the **SaddleNeg** treatment.
3. The a -number converges to its PONE level of 90 almost monotonically in the **Sink**, **SaddleNeg** and **SaddlePos** treatments.
4. The b -number converges to its PONE level of 20 in the **Sink** and **SaddleNeg** treatments with a pronounced oscillatory path in the latter treatment.
5. The b -number does not converge to its PONE level in the **SaddlePos** and **Source** treatments.

4 Behavioral Models

In this section, we analyze the dynamic properties of various learning models.¹⁸ Our focus is on whether such models are consistent with the features of data as found in Section 3. Understanding the properties of the models is important as it allows for generalizing beyond the specific parametrizations used in the experiment.

4.1 Homogeneous level- k learning model

In Section 3, we used a simple classification of the participants based on the idea of level- k thinking. Applying that same mechanism, we now define a homogeneous level- k learning model. Following Nagel (1995), we define level-0 players in period $t > 1$ as guessing the average of the a and b -numbers from the previous period.¹⁹

¹⁸Thus we study *theoretical trajectories* of the models. These trajectories depend *only* on the model's parameters and initial conditions. In every subsequent period, the trajectories are conditioned on the model's realization. In contrast, in econometric analysis of Section 5, the models are conditioned, at each time step, on actual experimental data.

¹⁹One may argue that our level-0 is too simplistic. We have chosen this definition over possible alternatives based on our experimental data as well as for the sake of exposition. Recall that the participants had access to the past averages of the a and b numbers, as well as to the past target numbers. Classification in Section 3 suggests that both pieces of information were used. The definition of level-0 as guessing the past averages, allows us to define level-1 players as those whose guesses equal the past target values, following the idea that higher-level agents best respond to lower-level agents.

Using vector notation and denoting the level-0 players' choices as $\mathbf{z}_t(0)$, we have

$$\mathbf{z}_t(0) = \bar{\mathbf{z}}_{t-1}. \quad (6)$$

Agents who follow the level-0 choice in each time period can be labeled as “stubborn”, since their guesses do not change throughout the experiment. In a *homogeneous* level-0 model, *all* agents are level-0 and so $\bar{\mathbf{z}}_t = \bar{\mathbf{z}}_{t-1} = \dots = \bar{\mathbf{z}}_1$. Thus, given the first period average guesses for both numbers, under the homogeneous level-0 model all subsequent guesses would remain the same in all remaining periods of the experiment.²⁰ Obviously, such a homogeneous level-0 model generates trajectories that are very different from those observed in our experiment.

Next we define the level-1 choice at time t as a best response to the level-0 choice at time t . It follows that level-1 players submit guesses that are equal to the previous period's targets, that is

$$\mathbf{z}_t(1) = \mathbf{z}_{t-1}^*. \quad (7)$$

The homogeneous level-1 model assumes that all agents make level-1 choices in each period. Using (2), we write this model as a two-dimensional dynamical system

$$\bar{\mathbf{z}}_t = \mathbf{M}\bar{\mathbf{z}}_{t-1} + \mathbf{d}. \quad (8)$$

In the interpretation of the experiment using expectations, this means that agents, in each period, expect the realized targets from the previous period. This homogeneous level-1 model is equivalent to the well-known *naïve* expectations model.

Proceeding further, we define for $k > 1$ the level- k choices for time t as guesses

²⁰The average of the first period numbers plays a role in the initial conditions of the learning models. Motivated by our analysis of the first period choices (see Appendix F2.2), we can further take (50, 50) as the first round level-0 choice. In the standard BC game experiment, Burchardi and Penczynski (2014) elicit level-0 beliefs of participants and find that 50 (the mid-point of the admissible range) is the modal level-0 belief.

Treatment	Eigensystem				Attractor and dynamics type		Data	
	μ_1	\mathbf{v}_1	μ_2	\mathbf{v}_2	for a	for b	for a	for b
Sink	$\frac{2}{3}$	$\begin{pmatrix} -7 \\ 3 \end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	a^E monotone	b^E oscillatory	✓	✓
SaddleNeg	$\frac{2}{3}$	$\begin{pmatrix} -13 \\ 3 \end{pmatrix}$	$-\frac{3}{2}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	a^E monotone	2-cycle $\{0, 50\}$ oscillatory	✓	✗
SaddlePos	$\frac{2}{3}$	$\begin{pmatrix} 5 \\ 3 \end{pmatrix}$	$\frac{3}{2}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	a^E monotone	0 or 100^\dagger monotone	✓	✓
Source	$\frac{3}{2}$	$\begin{pmatrix} -6 \\ 1 \end{pmatrix}$	$-\frac{3}{2}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	0 or 100^\dagger monotone	2-cycle $\{45, 95\}^\ddagger$ oscillatory	✓	✗

[†] depending on the initial conditions

[‡] assuming that the a -number dynamics converges to 0

Table 3: Properties of the homogeneous level-1 model for four experimental treatments. We report the eigensystem (two eigenvalues and the corresponding eigenvectors), the attractor of the model when guesses are truncated as in (10), and the type of dynamics (monotone or oscillatory). The last two columns verify whether the dynamics match the experimental data (✓) or not (✗).

that best respond to the choices of the previous level, $k - 1$, that is,

$$\mathbf{z}_t(k) = \mathbf{M}\mathbf{z}_t(k - 1) + \mathbf{d}. \quad (9)$$

In the homogeneous level- k model, all agents make level- k choices at each period.

What do the dynamics of these homogeneous learning models look like? We start with the case of $k = 1$, as described by system (8). This system has a unique steady state corresponding to the PONE, \mathbf{z}^E . The system is linear and its dynamic properties can be understood from the eigensystem of matrix \mathbf{M} , see Table 3. Since for all four treatments matrix \mathbf{M} is triangular, its eigenvalues, μ_1 and μ_2 , coincide with the diagonal elements. In fact, as explained in Section 2, we designed and named our treatments referring to the best response dynamics, which is exactly the homogeneous level-1 model (7). In particular, the phase diagrams of Fig. 1 correspond to the dynamics (8) in our four treatments.²¹ We observe that in the **SaddleNeg**

²¹The black lines are the eigenvectors, \mathbf{v}_1 and \mathbf{v}_2 , reported in Table 3. Since the matrix \mathbf{M} is lower triangular, the eigenvector \mathbf{v}_2 is a vertical line in all treatments.

treatment, almost all trajectories diverge in contrast to the experimental results.

The level- k guesses in (9), can be written as $\mathbf{z}_t(k) - \mathbf{z}^E = \mathbf{M}(\mathbf{z}_t(k-1) - \mathbf{z}^E)$. Applying this iteratively, we find that the dynamics of the homogeneous level- k model is given by $\bar{\mathbf{z}}_t - \mathbf{z}^E = \mathbf{M}^k(\bar{\mathbf{z}}_{t-1} - \mathbf{z}^E)$. Obviously, the dynamics have a unique steady state corresponding to the PONE and are governed by the linear system with matrix \mathbf{M}^k . It follows that the homogeneous level- k model has the *same* conditions for convergence, in terms of the eigenvalues, as the homogeneous level-1 model.²² Therefore, the discrepancy between the homogeneous level- k model and the experimental data in the **SaddleNeg** treatment will persist for any value of k .

Fig. 5 column 1 illustrates the dynamics of the homogeneous level-1 (naïve expectations) model and compares these with the experimental data, column 2. The panels in column 2 show the trajectories for average guesses in a representative experimental session (session 1) of each of the four treatments (from the top to the bottom row). The panels in column 1 show the simulated trajectories of the level-1 model in each treatment. The simulated trajectories start at the point of the first period averages of a and b -numbers from the experimental session, but otherwise they are not informed by the data. To take into account the restriction of our experiment that guesses must be in the interval $[0, 100]$, we truncate the choices in (7) as

$$\bar{a}_t = \max \{0, \min \{100, a_{t-1}^*\}\} \quad \text{and} \quad \bar{b}_t = \max \{0, \min \{100, b_{t-1}^*\}\}, \quad (10)$$

and simulate model (8) with these averages. Several observations can be made from a comparison of the simulations of the level-1 model and experimental data.

First, the dynamic paths of the a - and b -numbers in the simulations and in the data look very similar to one another in the **Sink** treatment (top row of Fig. 5). They both converge, and, remarkably, along the same direction, given by the eigenvector \mathbf{v}_1 of the model. Second, the a -number converges both in the model and in the

²²It is well-known that if μ is an eigenvalue of \mathbf{M} with eigenvector \mathbf{v} , then μ^k is an eigenvalue of \mathbf{M}^k with the same eigenvector. Moreover $|\mu| < 1$ if and only if $|\mu^k| < 1$.

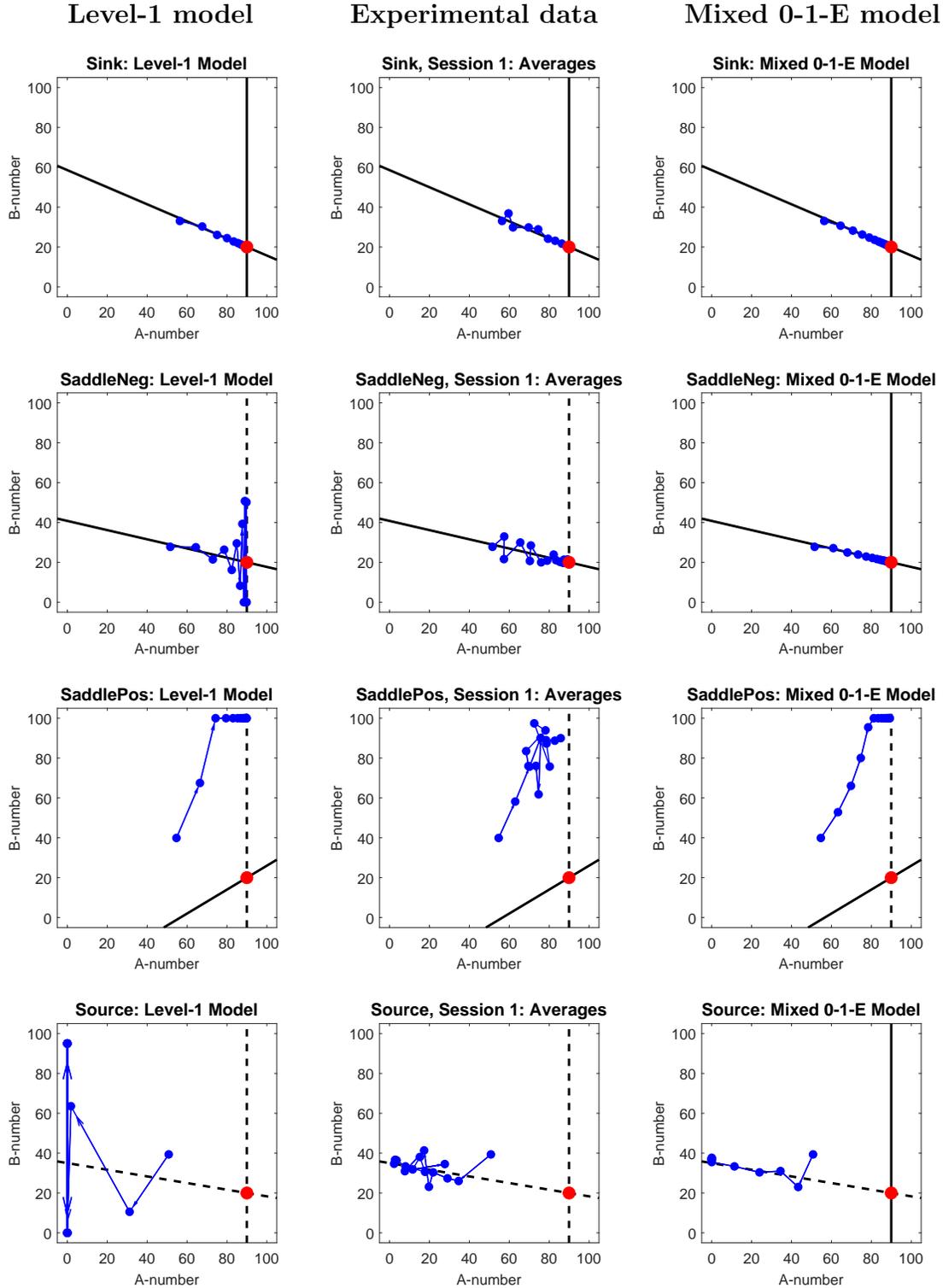


Figure 5: Dynamics of averages in an experimental group (*middle panels*) as compared with the dynamics of the homogeneous level-1 model (*left panels*) and mixed levels 0-1-E model with $f_0 = 0.41$, $f_1 = 0.52$ and $f_E = 0.07$ (*right panels*). For both learning models, 14 periods are simulated with the initial point given by the a and b -averages in the experimental group. The black lines are the eigenvectors, stable (solid) and unstable (dashed).

experiment to the PONE level of 90 in both saddle treatments (the two middle rows) and to 0 in the **Source** treatment (bottom row). But the model does *not* match the data in the dynamics of the b -number in two treatments. In the **SaddleNeg** treatment, the second eigenvalue of the model is negative, causing the simulated path for the b -number to jump back and force through \mathbf{v}_1 . The dynamics in the experiment exhibits similar oscillations. However, in the experiment these oscillations converge to the PONE, while in the model they converge to a limiting 2-cycle between 0 and 50. In the **Source** treatment, the limiting b -number dynamics of the model oscillate between 0 and 95, whereas in the experiment the dynamics stay close to a boundary Nash equilibrium value of 38. Third, the model accurately characterizes the dynamics in the **SaddlePos** treatment, apart from small details such as the speed of divergence and dynamics at the boundary. Table 3 summarizes this discussion. The *attractors* of the simulated dynamics are the sets to which the simulated trajectories converge. We distinguish between two types of dynamics, monotone or oscillatory, depending on whether the eigenvalue is positive or negative. The last two columns indicate whether the dynamics of the simulated model agrees with the experimental data.

As mentioned before, the dynamics of the homogeneous level- k model for larger values of k exhibit similar properties. Further, if k is equal to 2 or any other even number, all eigenvalues of the model are positive, which rules out the oscillatory behavior observed in several experimental treatments. We formalize this analysis as

Result 1. *The homogeneous level- k learning model for any finite k cannot reproduce all the features of our experimental data. In particular, this model does not predict convergence to the PONE in the **SaddleNeg** treatment, and there are discrepancies in other treatments as well, including very different dynamics in the **Source** treatment.*

Note that this result holds only if we restrict ourselves to the model with a *homogeneous* level of iterative thinking. Given the evidence for different levels of thinking in our experimental data, presented in Fig. 4, it is natural to relax the homogeneity assumption, which we do below.

4.2 Mixed levels 0-1 learning model

Let us assume that in each period t , a fraction $1 - \lambda$ of agents make the level-0 choice of the previous period average as in (6) and the remaining fraction λ of agents play the best response and guess the previous period targets as in (7). We call this model the *mixed levels 0-1* learning model. The averages in this model evolve as

$$\bar{\mathbf{z}}_t = \lambda \mathbf{z}_{t-1}^* + (1 - \lambda) \bar{\mathbf{z}}_{t-1} = (\lambda \mathbf{M} + (1 - \lambda) \mathbf{I}) \bar{\mathbf{z}}_{t-1} + \lambda \mathbf{d}, \quad (11)$$

where we substituted the target values from (2) to get the last expression.

Note that the dynamics (11) can be written as $\bar{\mathbf{z}}_t = \bar{\mathbf{z}}_{t-1} + \lambda(\mathbf{z}_{t-1}^* - \bar{\mathbf{z}}_{t-1})$. When guesses are interpreted as expectations, this relation is *identical to the model of adaptive expectations*.²³ According to this model, a representative agent adapts their expectations in a constant proportion, λ , to the forecast error. This observation provides a novel and useful bridge between the level- k models of strategic thinking and the model of adaptive expectations.

The dynamics of the mixed levels 0-1 model (11) have a unique steady state, the PONE \mathbf{z}^E , as in the homogeneous level- k models. The linear dynamics are now governed by the matrix $\lambda \mathbf{M} + (1 - \lambda) \mathbf{I}$. The eigenvalues of this matrix are easy to determine, as in all our treatments the matrix \mathbf{M} is lower triangular. They are the weighted averages of 1 (with weight $1 - \lambda$) and the eigenvalues of \mathbf{M} (with weight λ). Furthermore, matrix $\lambda \mathbf{M} + (1 - \lambda) \mathbf{I}$ has the same system of eigenvectors as the matrix \mathbf{M} , see Appendix F3.1.

Table 4 summarizes important convergence properties with the last two columns comparing the model with the data. We find that to obtain convergence to the PONE

²³The adaptive expectations model is quite popular in macroeconomics; see Nerlove (1958) and Hommes (1994) for theoretical treatment and Pfajfar and Žakelj (2016) and Bao and Duffy (2016) for recent experimental evidence. Adaptive expectations are closely related to constant gain learning models that are increasingly used in contemporary macroeconomic modeling, see Section 3.3 in Evans and Honkapohja (2001). The naïve expectations model is a special case of adaptive expectations where $\lambda = 1$.

Treatment	Eigenvalues		Attractor and dynamics type		Data	
	μ_1	μ_2	for a	for b	for a	for b
Sink	$1 - \frac{\lambda}{3}$	$1 - \frac{3\lambda}{2}$	a^E monotone	b^E oscillates for $\lambda > 2/3$	✓	✓
SaddleNeg	$1 - \frac{\lambda}{3}$	$1 - \frac{5\lambda}{2}$	a^E monotone	b^E for $\lambda < 4/5$ oscillates for $\lambda > 2/5$	✓	✓ $\lambda \in (\frac{2}{5}, \frac{4}{5})$ ✗ otherwise
SaddlePos	$1 - \frac{\lambda}{3}$	$1 + \frac{\lambda}{2}$	a^E monotone	0 or 100^\dagger monotone	✓	✓
Source	$1 + \frac{\lambda}{2}$	$1 - \frac{5\lambda}{2}$	0 or 100^\dagger monotone	38 or 18^\ddagger for $\lambda < 4/5$, oscillates for $\lambda > 2/5$	✓	✓ $\lambda < 4/5$ ✗ otherwise

[†] depending on the initial conditions

[‡] depending on whether the a -number dynamics converges to 0 or 100, respectively

Table 4: Properties of the mixed levels 0-1 model. See the caption of Table 3 for details.

in the **SaddleNeg** treatment, the fraction λ of level-1 agents in the learning model cannot be too large. Specifically, when $\lambda < 4/5$, the dynamics for the b -number will converge. The ‘jumping’ behavior along the b -dimension, visible in the phase diagrams, and translating into the oscillatory convergence observed in the **SaddleNeg** treatment is consistent with the model only when λ is not too small, i.e., for $\lambda > 2/5$. Also, the median numbers in the final experimental period (see Table 2) are the same as those predicted by the mixed levels 0-1 model for the diverging b -number in the **SaddlePos** treatment (100) and the diverging a -number in the **Source** treatment (0). The diverging b -number in the **Source** treatment is predicted to be 38 for this model, which is very close to the average value of 35 observed in the experiment. We summarize these findings as follows:

Result 2. *The mixed levels 0-1 learning model with proportion of the level-1 agents $\lambda \in (\frac{2}{5}, \frac{4}{5})$ can reproduce all features of the experimental data.*

The mixed levels 0-1 model effectively averages the last period information that was available to the participants in the experiment. In Appendix F3.3 we discuss the models that average available information from *all past periods*.

4.3 Higher order mixed level- k models

Generalizing the mixed levels 0-1 model, we assume that there is an exogenous distribution, $\mathcal{F} = \{f_k\}_{k \geq 0} \cup f_E$, of agents with different levels of reasoning in the population. Level- k agents are present in proportion $f_k \in [0, 1)$. We also include proportion $f_E \geq 0$ of an ‘equilibrium type’ who always submits the fixed point, \mathbf{z}^E as their guess.²⁴ Level-0 agents use the previous period averages as in (6). Depending on assumptions made about higher level agents, we consider the two families of models, mixed level- k and cognitive hierarchy models.

Mixed level- k model. In this model, level- k agents always assume that all of the other agents are of the next lower level $k - 1$, and thus they best respond to level $k - 1$ agents, as in (9). The dynamics of the model is given by:

$$\bar{\mathbf{z}}_t - \mathbf{z}^E = \left(f_0 \mathbf{I} + f_1 \mathbf{M} + f_2 \mathbf{M}^2 + \dots + f_k \mathbf{M}^k + \dots \right) (\bar{\mathbf{z}}_{t-1} - \mathbf{z}^E). \quad (12)$$

The dynamics of the general, mixed level- k model have a unique steady state (the PONE) and are linear. The eigenvalues, μ_1 and μ_2 , are the convex combinations of the k -th powers of the eigenvalues of matrix \mathbf{M} , weighted by the corresponding fractions, f_k . This property enables us to reproduce the features of the experimental data in all four treatments, similarly to the mixed levels 0-1 model of Section 4.2, which is a special case with $f_0 = 1 - \lambda$ and $f_1 = \lambda$.

Whereas the mixed level- k model provides more flexibility in general, there are some restrictions on the distribution of levels needed in or to match our experimental data. Consider, for example, the oscillating convergence in the **SaddleNeg** experimental treatment. To reproduce it, we should have both $\mu_2 > -1$ and $\mu_2 < 0$ in the model. The former is ruled out when *all* levels k and hence powers of matrix \mathbf{M}

²⁴We observed the presence of this type in small measure in the experiment in the classification in Section 3. The equilibrium guesses, \mathbf{z}^E , were also occasionally submitted by participants who were not classified as being an equilibrium type, see Table 7 in Appendix D.

are odd; the latter is ruled out when all levels k are even.²⁵ Therefore, as a minimal requirement, *both even and odd levels need to be present in the mixed model*. Also note that the presence of the equilibrium type helps to stabilize the model dynamics in all treatments. Tables F.3 and F.4 of Appendix F3.2 present the properties of the mixed levels 1-2 and 0-1-E models, respectively.

As we will see in Section 5, the 0-1-E model with a small fraction of equilibrium types is the best in describing our data quantitatively. Fig. 5 column 3 illustrates the trajectories of the mixed levels 0-1-E model with the parameters estimated from the data (see the caption). The trajectories start at the point of the first period averages of the experimental session 1, so that the dynamics can be compared with those from the homogeneous level-1 model (column 1) and the experiment (column 2). The 0-1-E model reproduces the main features of the experimental dynamics much better than the homogeneous level-1 model. Most importantly, the dynamics of the mixed levels model are converging to the steady state in the **SaddleNeg** treatment as in the experimental data and exhibit divergent dynamics in the **Source** treatment that are similar to those observed in the experimental data (not just for session 1, but for all four sessions of these treatments as well). Truncation according to (10) implies that for the diverging treatments, simulations result in more regular dynamics than in the experiment. Similarly, in the converging treatments, simulations are smoother and closer to the eigenvector than in the experiment. This is because, aside from the initial condition, the trajectories of the learning model do not use the experimental data, and thus cannot account for discrepancies between the model’s predictions and the noisier experimental dynamics at every time step.

We summarize the discussion of the mixed level- k models as the following result.

Result 3. *The mixed level- k model reproduces all features of our experimental data for many, but not all, distributions of levels. Oscillating convergence in the **SaddleNeg***

²⁵When all levels are odd, $\mu_2 = f_1 m_{22} + f_3 m_{22}^3 + \dots$. Since $m_{22} = -3/2$ in the **SaddleNeg** treatment, $\mu_2 \leq -3/2$. When all levels are even, $\mu_2 = f_0 + f_2 m_{22}^2 + f_4 m_{22}^4 + \dots$ and so $\mu_2 > 0$. In this discussion, we define the equilibrium type as even.

treatment is reproduced if and only if both even and odd levels of thinking have non-zero proportions. The equilibrium type stabilizes the dynamics; its fraction should be small enough to be consistent with divergence in the **SaddlePos** and **Source** treatments.

Cognitive Hierarchy Model. The second family of higher order mixed-level k models introduced in Camerer et al. (2004), builds a “cognitive hierarchy” (CH), where agents of level k best respond not simply to agents of level $k - 1$, but rather to a *distribution over all lower levels*. The equilibrium type continues to submit \mathbf{z}^E , but all other types assume that *all* other agents’ levels are lower than theirs, with the different level k types present in the proportions given by the suitably adjusted distribution \mathcal{F} . Thus, every level- k agent first builds a perceived distribution of the levels of the other agents in the population, $\{g_\ell^{(k)}\}_{\ell=0}^{k-1}$, with the fraction of level- ℓ agents (with $\ell = 0, 1, \dots, k - 1$) given by

$$g_\ell^{(k)} = \frac{f_\ell}{f_0 + f_1 + \dots + f_{k-1}}.$$

Then, the agent best responds to the perceived behavior of the others. Therefore, the guess of the level- k type, in deviations from the equilibrium, is given by

$$\mathbf{M}(g_0^{(k)}\mathbf{I} + g_1^{(k)}\mathbf{M} + \dots + g_{k-1}^{(k)}\mathbf{M}^{k-1})(\bar{\mathbf{z}}_{t-1} - \mathbf{z}^E).$$

This guess is weighted with the actual fraction, f_k , to represent an effect of all level- k agents on the total dynamics. When these effects are summed up over all levels of thinking in the population, we obtain the dynamics of the CH model.

With only two levels, 0 and 1, the CH model is identical to the mixed levels 0-1 model.²⁶ The next result shows that, in general, the CH model is a mixture of several mixed levels 0-1 models with appropriate coefficients. (The proof is in Appendix E.)

²⁶Camerer et al. (2004) focus on the special case where \mathcal{F} is a Poisson distribution, truncated at the maximum level of thinking in the population. The only parameter of this distribution, τ , characterizes the average cognitive level of the population. The Poisson CH model truncated to the levels 0 and 1 is equivalent to the mixed levels 0-1 model with the fraction of level-1 type given by $\lambda = \tau/(1 + \tau)$, establishing a one-to-one correspondence between the parameters of the two models.

Proposition 4.1. *The dynamics of the CH model with distribution $\mathcal{F} = \{f_k\}_{k=0}^K \cup f_E$ are given by*

$$\bar{\mathbf{z}}_t - \mathbf{z}^E = (1 - f_E) \prod_{i=1}^K (\lambda_i \mathbf{M} + (1 - \lambda_i) \mathbf{I}) (\bar{\mathbf{z}}_{t-1} - \mathbf{z}^E), \quad (13)$$

with $\lambda_K = f_K$, and other weights defined as follows: for any $\ell < K$, $\lambda_\ell = f_\ell / \sum_{j=0}^\ell f_j$.

Proposition 4.1 shows that the CH model generates linear dynamics with the steady state \mathbf{z}^E . Apart from the stabilizing effect of having some equilibrium types, its convergence properties are determined from interaction of the underlying levels 0-1 models. The coefficients (λ 's) of these models are $f_K, g_{K-1}^{(K-1)}, g_{K-2}^{(K-2)}$, etc., i.e., the fraction of the highest level of thinking in the population, and then, for each lower level, the relative fraction of this level in the subset of the population consisting of this level and the levels below. Since matrix \mathbf{M} is always triangular, all factors on the right side of (13) are triangular matrices. Hence, the eigenvalues of the dynamics of (13) are the products of the eigenvalues of those matrices and $1 - f_E$. Assume that $f_E = 0$. It follows from Proposition 4.1 that the dynamics of the CH model will converge if all levels 0-1 models on the right side of (13) converge, and they will diverge if all of these models diverge. Combining this with Result 2, we obtain

Result 4. *Consider the CH model with $\mathcal{F} = \{f_k\}_{k=0}^K$ and λ_i defined as in Proposition 4.1. If $\lambda_i \in (\frac{2}{5}, \frac{4}{5})$ for all i , the CH model is consistent with all of our experimental findings (except for oscillations in the **SaddleNeg** treatment if K is even). If $\lambda_i \notin (\frac{2}{5}, \frac{4}{5})$ for all i , the CH model is not consistent with the experimental data.*

If some λ_i 's are in the set $(\frac{2}{5}, \frac{4}{5})$ for which the mixed levels 0-1 model is consistent with the data, and some λ_i 's are outside of this set, the properties of the model in relation to the experimental data depend on the exact distribution for \mathcal{F} .

5 Estimation

In this section, we compare models in terms of their *quantitative fit* to the experimental data at the individual subject level. For a given session and period, each model \mathcal{M} generates an individual predicted guess by conditioning on some, if any, of the experimental data available to the participant in that session and period, i.e.,

$$\mathbf{z}_{i,t}^m = \mathcal{M} [\mathbf{z}_{t-1}^*, \bar{\mathbf{z}}_{t-1}, \mathbf{z}_{i,t-1}, \dots; \boldsymbol{\theta}] , \quad (14)$$

where we omit the session-specific index, and $\boldsymbol{\theta}$ denotes any model parameters. Each model generates a one-period ahead prediction, $\mathbf{z}_{i,t}^m = (a_{i,t}^m, b_{i,t}^m)$, for each individual i at time t in a given treatment and session. We compare the following models:

- the Pareto-Optimal Nash Equilibrium (PONE) prediction;
- the homogeneous level- k models (9) with $0 \leq k \leq 4$;
- the mixed level models (12) with various combinations of levels;
- the cognitive hierarchy (CH) model (13) with a Poisson distribution over levels;
- the average models in Appendix F3.3: the moving average (MAve) model (F.5); and the exponential weighted moving average (EWMA) model (F.4).

For completeness, we also include the following popular models from the literature, whose detailed specifications are reported in Appendix F4:

- the quantal response equilibrium (QRE) model of McKelvey and Palfrey (1995);
- the noisy introspection (NI) model of Goeree and Holt (2004);
- the experience-weighted attraction (EWA) model of Camerer and Ho (1999);
- and the structural mixed level- k learning model of Gill and Prowse (2016).

Similarly to the PONE, the QRE and NI are not learning models, as they do not condition on data from past rounds. We include them as common benchmarks to see whether learning can improve the fit. The EWA model is a popular learning model, which explains behavior in many games well and was often estimated for

one-dimensional BC games. Finally, the recent Gill and Prowse (2016) model is particularly interesting for us, because it structurally specifies individual behavior and explicitly allows for “learning types” who increase their levels over time.

For the models that contain parameters, we estimate them in two ways. For the mixed level- k , EWMA and the CH models, we do not specify the error distribution explicitly and estimate the parameters by minimizing the sum of squared errors (SSE) between the model predictions and the actual data from the experiment.²⁷ For the QRE, NI, Gill-Prowse, and EWA models, where the distributional assumptions are explicit, we use maximum likelihood estimation. Since estimating most models requires conditioning on the lagged data, we estimate all models with the *dependent* variable starting from period $t = 2$. We assume that choices are independent between individuals, conditional on the common past information. The SSE is thus defined as

$$\sum_i \sum_{t=2}^{15} \left((a_{i,t}^m(\boldsymbol{\theta}) - a_{i,t})^2 + (b_{i,t}^m(\boldsymbol{\theta}) - b_{i,t})^2 \right),$$

where the first sum is over all individuals in the experiment. The log-likelihoods for the ML estimation depend on the specific models and are given in Appendix F4.

We seek universal parameters for all treatments and sessions to ensure the models are parsimonious, yet sufficiently general. For this reason, in estimating the model’s parameters $\boldsymbol{\theta}$, we minimize the SSEs or maximize the likelihoods over all treatments and sessions. Parameter estimates resulting from this estimation for different models are shown in Table 6 with associated standard errors reported in parentheses.²⁸

²⁷This procedure is equivalent to the maximum likelihood (ML) estimation under the assumption of normal or Student’s t errors (Zellner, 1976). This procedure, however, does not allow for a dynamic specification for the variance in the errors (cf. Ho et al., 1998; Gill and Prowse, 2016).

²⁸We compute standard errors using a nonparametric bootstrap method (Hansen, 2019, Chapter 10.9). To account for time and session dependence, we perform bootstrap random sampling with replacement on the level of one session in each treatment. In particular, each new bootstrap sample is generated by drawing randomly with replacement four sessions for each treatment from the four actual sessions corresponding to a particular treatment. We compute parameter estimates for each of the bootstrap samples and then take their standard deviation as standard errors. We use 1,000 bootstrap samples. This is sufficient for computing standard errors up to the two decimal places.

After estimating each model’s parameters, we compare the performance of the models with the data. To facilitate the comparison between the models that make no explicit distributional assumptions and the more detailed models that assume specific distributions and account for over-fitting, we use the out-of-sample one-step-ahead prediction root mean squared error (RMSE) criterion. To compute this RMSE, we use the so-called *leave-one-out* cross-validation procedure of Stone (1974).²⁹ The RMSE is defined as

$$\text{RMSE}_T = \sqrt{\frac{1}{4} \sum_{s \in S_T} \frac{1}{10} \sum_{i \in I_s} \frac{1}{14} \sum_{t=2}^{15} \frac{1}{2} \left((a_{i,t}^m(\boldsymbol{\theta} | \Omega_{-s}) - a_{i,t})^2 + (b_{i,t}^m(\boldsymbol{\theta} | \Omega_{-s}) - b_{i,t})^2 \right)},$$

where S_T is the set of all sessions in treatment T , I_s is the set of all individuals in session s , and $\boldsymbol{\theta} | \Omega_{-s}$ denotes the parameters, used to generate $a_{i,t}^m$, $b_{i,t}^m$, that are estimated from Ω_{-s} , all available experimental data except for data from session s .³⁰ In particular, we estimate the parameters with one of the sessions *left out*, and then compute MSE of the estimated model on the data of the left-out session. After this procedure is performed four times, leaving out each session of the treatment, we average the resulting MSEs and compute the RMSE for the treatment. Consistent with the estimation, we evaluate RMSEs starting from period 2 for all models.

The out-of-sample RMSEs for all models are reported in Table 5. In the last column of the table, we report the RMSEs over all treatments, that is, the square root of the average MSE over all 16 sessions.³¹ The RMSE of the best model (i.e., the smallest RMSE) for each treatment and overall is shown in boldface.

²⁹This procedure is “out-of-sample” because the data of the session for which the MSE is computed are not used to estimate the model parameter(s). This procedure has an implicit penalty for the number of parameters, i.e., over-parametrized models perform better in-sample, but worse out-of-sample because of over-fitting. AIC and BIC are popular model selection criteria that have explicit penalties for the number of parameters. In contrast to these criteria, leave-one-out cross-validation does not require any distributional assumptions. Under certain conditions, it is asymptotically equivalent to AIC (Stone, 1977).

³⁰For the models that make explicit distributional assumptions and generate the whole predictive distribution rather than a point forecast, that is, QRE, NI, Gill-Prowse, and EWA, we use the conditional means of these predictive distributions as point predictions $a_{i,t}^m$ and $b_{i,t}^m$.

³¹We also computed the RMSEs of the applicable models for period 1 choices, see Appendix F4.5.

Models		Out-of-sample RMSEs				
		Sink	SadNeg	SadPos	Source	Overall
PONE		16.93	11.61	51.03	57.20	39.68
QRE		15.74	11.50	28.33	21.04	20.27
NI		16.21	16.87	34.01	28.82	25.17
Average	MAve(2)	12.59	8.22	18.39	20.75	15.77
	MAve(3)	12.61	8.27	18.24	20.67	15.71
	MAve(4)	12.60	8.40	18.18	20.55	15.67
	MAve(5)	12.66	8.54	18.16	20.58	15.70
	MAve(14)	12.98	9.40	18.37	20.58	15.95
	EWMA	12.62	8.42	18.20	20.52	15.67
Level- k	0	12.85	9.25	17.54	20.02	15.49
	1	12.75	8.27	18.57	21.32	16.05
	2	13.61	11.52	20.92	25.24	18.66
	3	14.53	11.18	21.43	24.21	18.58
	4	15.25	15.26	21.80	31.15	21.86
Mixed	0 1	12.52	7.92	16.76	18.79	14.61
	0 1 2	12.52	7.92	16.76	18.79	14.61
	0 1 2 3	12.57	7.82	16.69	18.71	14.56
	0 1 2 3 4	12.57	7.82	16.69	18.71	14.56
	1 2	12.85	7.99	18.84	20.87	15.96
	0 E	12.85	9.24	17.55	20.04	15.49
	1 E	13.04	8.34	17.20	18.98	14.97
	0 1 E	12.56	7.78	16.61	18.42	14.43
	0 1 2 E	12.58	7.80	16.61	18.41	14.44
	0 1 2 3 E	12.58	7.80	16.61	18.42	14.44
	0 1 2 3 4 E	12.58	7.80	16.61	18.41	14.44
	1 2 E	13.18	8.21	17.39	18.50	14.88
Gill-Prowse	0-1 Learning	12.67	7.89	17.21	21.67	15.72
	0-1 E Learning	12.68	7.88	17.03	21.24	15.53
	0-2 Learning	12.76	7.93	17.39	20.75	15.48
	0-2 E Learning	12.77	7.94	17.13	20.25	15.24
	0-3 Learning	12.77	7.93	17.40	20.79	15.50
	0-3 E Learning	12.79	7.94	17.14	20.28	15.26
	0-4 Learning	12.77	7.93	17.40	20.79	15.50
	0-4 E Learning	12.79	7.94	17.14	20.28	15.26
CH-Poisson	0 1	12.52	7.92	16.76	18.79	14.61
	0 1 2	12.53	7.99	16.75	18.82	14.62
	0 1 2 3	12.53	8.00	16.74	18.82	14.62
	0 1 2 3 4	12.53	8.00	16.74	18.82	14.62
EWA		12.19	8.32	24.28	23.06	18.66
EWA, reinforcement learning		27.17	28.49	37.41	32.66	31.69
EWA, belief-based learning		12.22	8.69	26.80	17.57	17.69

Table 5: Performance of different models using individual-level data for periods 2-15. The models are compared in terms of their out-of-sample, one-step-ahead RMSE using the leave-one-out procedure. The smallest RMSE for each treatment and overall is shown in boldface. Parameter estimates of the models with parameters are reported in Table 6. The mixed levels 0-1 and CH-Poisson 0-1 models have the same RMSEs as they are equivalent, see Section 4.3.

Models		Parameter Estimates						
QRE		λ						
		0.14 (0.01)						
NI		λ		μ				
		0.13 (0.01)		0.95 (0.00)				
Average EWMA		λ						
		0.42 (0.06)						
Mixed		f_0	f_1	f_2	f_3	f_4	f_E	
		0 1	0.56 (0.03)	0.44 (0.03)				
		0 1 2	0.56 (0.03)	0.44 (0.03)	0.00 (0.00)			
		0 1 2 3	0.64 (0.03)	0.12 (0.06)	0.00 (0.00)	0.23 (0.03)		
		0 1 2 3 4	0.65 (0.03)	0.12 (0.06)	0.00 (0.00)	0.23 (0.03)	0.00 (0.00)	
		1 2		0.84 (0.03)	0.16 (0.03)			
		0 E	1.00 (0.00)					0.00 (0.00)
		1 E		0.86 (0.01)				0.14 (0.01)
		0 1 E	0.41 (0.02)	0.52 (0.02)				0.07 (0.01)
		0 1 2 E	0.40 (0.03)	0.51 (0.02)	0.03 (0.02)			0.07 (0.01)
		0 1 2 3 E	0.41 (0.04)	0.48 (0.05)	0.02 (0.02)	0.02 (0.03)		0.07 (0.01)
		0 1 2 3 4 E	0.41 (0.04)	0.48 (0.04)	0.02 (0.02)	0.02 (0.03)	0.00 (0.00)	0.07 (0.01)
		1 2 E		0.71 (0.02)	0.16 (0.03)			0.14 (0.01)
		Gill-Prowse		f_0	f_1	f_2	f_3	f_4
0-4 Learning	0.11 (0.02)			0.23 (0.06)	0.06 (0.05)	0.00 (0.00)	0.00 (0.00)	
0-4 E Learning	0.11 (0.02)			0.21 (0.05)	0.06 (0.04)	0.00 (0.00)	0.00 (0.00)	0.02 (0.01)
	$f_{0 \rightarrow 1}$			$f_{1 \rightarrow 2}$	$f_{2 \rightarrow 3}$	$f_{3 \rightarrow 4}$		
0-4 Learning	0.22 (0.07)			0.36 (0.05)	0.02 (0.02)	0.00 (0.00)		
0-4 E Learning	0.23 (0.06)			0.35 (0.04)	0.02 (0.02)	0.00 (0.00)		
Gill-Prowse		α	β	δ	ν			
		0-4 Learning	1.84 (0.19)	-0.59 (0.34)	3.39 (0.37)	0.50 (0.00)		
		0-4 E Learning	1.75 (0.17)	-0.91 (0.81)	3.20 (0.36)	0.50 (0.00)		
CH-Poisson		τ						
		0 1	0.78 (0.08)					
		0 1 2	0.58 (0.05)					
		0 1 2 3	0.57 (0.05)					
		0 1 2 3 4	0.56 (0.05)					
EWA		δ	ϕ	κ	λ			
		0.65 (0.04)	0.33 (0.11)	0.00 (0.00)	0.22 (0.03)			
		EWA, reinforcement learning	set to 0	0.29 (0.15)	0.00 (0.00)	0.14 (0.06)		
EWA, belief-based learning		set to 1	0.31 (0.04)	0.00 (0.00)	0.19 (0.03)			

Table 6: Parameter estimates of the models for period 2-15 guesses conditional on all observations including period 1 observations in the learning models. Standard errors are reported in parentheses. Appendix F4 specifies the parametrizations of the QRE, NI, EWA and Gill-Prowse models. For the Gill-Prowse model, we report the parameter estimates of the most comprehensive 0-4 and 0-4 E specifications. The parameters of other specifications are similar and are reported in Table F.5 in Appendix F4.4.

Comparing the RMSEs across all considered models, we find, unsurprisingly that the PONE prediction is the least accurate. The QRE and NI models also do not perform very well, This suggests that *learning* is important in our experimental setting. Moving to the learning models, we observe that the homogeneous level- k models perform rather poorly, which is consistent with Result 1 in Section 4. Among these models, the homogeneous level-0 model performs better overall due to better performance in the non-convergent **SaddlePos** and **Source** treatments, whereas the level-1 model performs better in the convergent **Sink** and **SaddleNeg** treatments. The average models generally perform better than the homogeneous level- k models. In particular, the model with a 4-lag moving window and the EWMA model are the best among all average models.

All of the remaining models have an even better overall performance except for the EWA model, which performs poorly in the two non-convergent treatments.³² The EWA model does outperform all models for the **Sink** treatment. The reinforcement learning model (EWA with $\delta = 0$) performs worse than the unrestricted EWA model which is to be expected in this dynamic learning environment. By contrast, the belief-based learning model (EWA with $\delta = 1$) performs well, and attains the smallest RMSE for the **Source** treatment.³³ We observe that, in agreement with Result 3 in Section 4, the class of the mixed level models generally perform better than the other models considered in Table 5. The Poisson CH model class, which is a single-parameter disciplined version of the mixed models, performs slightly worse than the

³²Wilcox (2006) shows how the parameters of the EWA model may be biased due to individual heterogeneity. In models that pool across all individuals, any idiosyncratic errors in individual choices will make them correlated over time. Further, if such models use individual past choices to predict individual current choices, there may be an upward bias in the parameter estimates on the individual's past choices. Following Wilcox (2006), we checked an alternative specification of the EWA model by allowing individual heterogeneity in parameter λ of the model. The other parameters remained within their standard error bounds indicating no substantial bias in our case. Importantly, none of the other models that we estimate are subject to this critique because the individual choices in these models do not condition on the individual's *own* past choices but rather condition on the past *average* choices (or the functions of those averages).

³³Since δ is set to 1 and the κ estimate is 0, this belief-based model corresponds to weighted fictitious play, which is related to the adaptive expectations model, see footnote 9 in Camerer et al. (2002). The mixed levels 0-1 model, which also yields a good performance, is closely related to the adaptive expectations model as well, see Section 4.2.

mixed models. Note that the mixed 0-1 model and CH-Poisson 0-1 model are equivalent. The mixed 0-1-E model has the lowest RMSE in all treatments except for **Sink** and has the lowest RMSE overall.³⁴ From Table 6, this model is a mixture of 41% level-0s, 52% level-1s and 7% equilibrium types. In terms of the parametrization in Table F.4 in Appendix F3.2, for this model $\lambda = 0.56$. This model satisfies the necessary and sufficient condition (F.3) to reproduce the data features. We illustrate the simulated path of this model (not correcting for any discrepancies with experimental data during the simulations) in the right, column 3 panels of Fig. 5. Interestingly, the same, mixed 0-1-E model shows the lowest RMSE for period 1 only (Table F.7 in Appendix F4.5), but in that case the mixture of levels 0, 1 and the equilibrium type are 68%, 18% and 14%, respectively. This indicates that there is some individual learning over the levels, as there is a decrease in level-0s and an increase level-1s moving from period 1 to the remaining periods. The fraction of the equilibrium type also reduces, which may indicate that, although some participants can compute the PONE, they soon realize that this is not a best response to the behavior of others.

By contrast with the mixed levels models, the Gill-Prowse model is a structural model that aims to model individual behavior consistently and specifies both fixed level types as well as learning types who increase their levels over time. Since the Gill-Prowse model imposes structural restrictions, its fit is not as good as the mixed 0-1-E model, but it is very close. The best performing version of the Gill-Prowse model includes levels 0, 1 and 2 and the learning types between them. Adding the equilibrium type to the Gill-Prowse model provides an even better fit to the data. The model estimates that 40% of subjects are fixed level types up to level-2 with the highest fraction being level-1 types, a small fraction of equilibrium types and 60% who are learning types. These results are generally consistent with the individual data classification, see Fig. 4.

³⁴Note that many models with more levels than in the mixed 0-1-E model, e.g., mixed 0-1-2-3-4-E model, have estimated proportions of these additional levels that are close to 0, indicating that these levels are “redundant”. That is why the fit of these models is close to the fit of the mixed 0-1-E model, but not as good because of the implicit penalty for the over-parametrization.

6 Conclusion

We have designed and reported on an experiment examining learning behavior in linear, coupled, multivariate systems, which are widely used in game theory, microeconomics, macroeconomics and other fields. Our experiment makes use of and generalizes the well-known Beauty Contest game to *two* dimensions, resulting in a richer and more complex, coupled multivariate beauty contest game.

We conducted four experimental treatments with different eigenvalues for this coupled multivariate system. We found convergence (divergence) to the unique PONE in the **Sink (Source)** treatments. More surprising, we found that subjects are able to learn a steady state with the saddlepath property when the sign of the unstable eigenvalue is negative, as in our **SaddleNeg** treatment, but not when it is positive as in our **SaddlePos** treatment. This is an important finding since steady states with the saddlepath property are common in theoretical models used across a variety of different fields in economics. We find that, as in uni-variate models such as the original beauty contest game, negative feedback seems to play a role in disciplining convergence to equilibrium in coupled multivariate linear systems.

Finally, we have compared the performance of a variety of different learning models found in the literatures on learning in games and macroeconomics in terms of their convergence dynamics as well as the fit of these estimated models to individual level data. Indeed, we provide a kind of “Rosetta stone” linking, e.g., level- k models from behavioral game theory with the adaptive learning approach used in the econometric learning literature in macroeconomics. Among the many learning models we consider, we find that a heterogeneous-agent, mixed level- k learning model (with some measure of levels 0, 1 and the equilibrium type), outperforms all other considered learning models in terms of the out-of-sample root mean squared error of its predictions relative to individual level data and we provide some intuition for this result.

References

- Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *The Economic Journal* 117(520), 603–636.
- Alaoui, L. and A. Penta (2015). Endogenous Depth of Reasoning. *The Review of Economic Studies* 83(4), 1297–1333.
- Alaoui, L. and A. Penta (2021). Cost-benefit analysis in reasoning. Barcelona GSE Working paper.
- Alvarez-Cuadrado, F., G. Monteiro, and S. J. Turnovsky (2004). Habit formation, catching up with the joneses, and economic growth. *Journal of Economic Growth* 9(1), 47–80.
- Anufriev, M. and C. Hommes (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics* 4(4), 35 – 64.
- Assenza, T., P. Heemeijer, C. H. Hommes, and D. Massaro (2021). Managing self-organization of expectations through monetary policy: a macro experiment. *Journal of Monetary Economics* 117, 170–186.
- Azariadis, C. (1993). *Intertemporal Macroeconomics*. Blackwell Publishing Company.
- Bao, T. and J. Duffy (2016). Adaptive versus eductive learning: Theory and evidence. *European Economic Review* 83, 64–89.
- Benhabib, J., S. Schmitt-Grohé, and M. Uribe (2001). The perils of taylor rules. *Journal of Economic Theory* 96(1-2), 40–69.
- Blanchard, O. J. and C. M. Kahn (1980). The solution of linear difference models under rational expectations. *Econometrica* 48(5), 1305–1311.
- Branch, W. A. and B. McGough (2009). A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control* 33(5), 1036–1051.

- Bulow, J. I., J. D. Geanakoplos, and P. D. Klemperer (1985). Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political economy* 93(3), 488–511.
- Burchardi, K. B. and S. P. Penczynski (2014). Out of your mind: Eliciting individual reasoning in one shot games. *Games and Economic Behavior* 84, 39–57.
- Camerer, C. and H. T. Ho (1999). Experience-weighted attraction learning in normal form games. *Econometrica* 67(4), 827–874.
- Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Russell Sage Foundation.
- Camerer, C. F., T.-H. Ho, and J.-K. Chong (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory* 104(1), 137–188.
- Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics* 119(3), 861–898.
- Ellison, M. and J. Pearlman (2011). Saddlepath learning. *Journal of Economic Theory* 146(4), 1500–1519.
- Eusepi, S. and B. Preston (2018). The science of monetary policy: An imperfect knowledge perspective. *Journal of Economic Literature* 56(1), 3–59.
- Evans, G. W. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics* 10(2), 171–178.
- Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games*. MIT press.

- Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- Gill, D. and V. Prowse (2016). Cognitive ability, character skills, and learning to play equilibrium: A level- k analysis. *Journal of Political Economy* 124(6), 1619–1676.
- Goeree, J. K. and C. A. Holt (2004). A model of noisy introspection. *Games and Economic Behavior* 46(2), 365–382.
- Güth, W., M. Kocher, and M. Sutter (2002). Experimental ‘beauty contests’ with homogeneous and heterogeneous players and with interior and boundary equilibria. *Economics Letters* 74(2), 219–228.
- Hansen, B. (2019). *Econometrics*. University of Wisconsin.
- Heemeijer, P., C. Hommes, J. Sonnemans, and J. Tuinstra (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic Dynamics and Control* 33(5), 1052–1072.
- Ho, T.-H., C. Camerer, and K. Weigelt (1998). Iterated dominance and iterated best response in experimental ‘p-beauty contests’. *The American Economic Review* 88(4), 947–969.
- Hommes, C. (1994). Dynamics of the cobweb model with adaptive expectations and nonlinear supply and demand. *Journal of Economic Behavior & Organization* 24(3), 315–335.
- Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control* 35(1), 1–24.
- Hommes, C. (2013). *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press.

- Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005). Coordination of expectations in asset pricing experiments. *The Review of Financial Studies* 18(3), 955–980.
- Kryvtsov, O. and L. Petersen (2021). Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics* 117, 760–780.
- Kurz, M., G. Piccillo, and H. Wu (2013). Modeling diverse expectations in an aggregated new keynesian model. *Journal of Economic Dynamics and Control* 37(8), 1403–1433.
- Levelt, E., C. H. Hommes, and M. Hennequin (2021). Planar learning to forecast market games. Technical report. <https://dx.doi.org/10.2139/ssrn.3882461>.
- Liu, P. and Z. Li (2011). Toward understanding the relationship between task complexity and task performance. In *International Conference on Internationalization, Design and Global Development*, pp. 192–200. Springer.
- Mauersberger, F. (2021). Monetary policy rules in a non-rational world: A macroeconomic experiment. *Journal of Economic Theory*, 105203.
- Mauersberger, F. and R. Nagel (2018). Levels of reasoning in keynesian beauty contests: A generative framework. In C. Hommes and B. LeBaron (Eds.), *Handbook of Computational Economics*, Volume 4, pp. 541–634. North-Holland.
- McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior* 10(1), 6–38.
- Mirdamadi, M. and L. Petersen (2018). Macroeconomic literacy and expectations. Available at SSRN 3187131. <http://dx.doi.org/10.2139/ssrn.3187131>.
- Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review* 85(5), 1313–1326.

- Nerlove, M. (1958). Adaptive expectations and cobweb phenomena. *The Quarterly Journal of Economics* 72(2), 227–240.
- Oprea, R. (2020). What makes a rule complex? *American Economic Review* 110(12), 3913–51.
- Pfajfar, D. and B. Žakelj (2016). Inflation expectations and monetary policy design: Evidence from the laboratory. *Macroeconomic Dynamics* 22, 1–41.
- Sargent, T. J. (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press.
- Sonnemans, J. and J. Tuinstra (2010). Positive expectations feedback experiments and number guessing games as models of financial markets. *Journal of Economic Psychology* 31(6), 964–984.
- Stahl, D. and P. Wilson (1994). Experimental evidence on players’ models of other players. *Journal of Economic Behavior & Organization* 25(3), 309–327.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)* 36(2), 111–133.
- Stone, M. (1977). An asymptotic equivalence of choice of model by cross-validation and akaike’s criterion. *Journal of the Royal Statistical Society: Series B (Methodological)* 39(1), 44–47.
- Sutan, A. and M. Willinger (2009). Guessing with negative feedback: An experiment. *Journal of Economic Dynamics and Control* 33(5), 1123–1133.
- Wilcox, N. T. (2006). Theories of learning in games and heterogeneity bias. *Econometrica* 74(5), 1271–1292.
- Zellner, A. (1976). Bayesian and non-bayesian analysis of the regression model with multivariate student-t error terms. *Journal of the American Statistical Association* 71(354), 400–405.

APPENDIX

A Motivating Examples

In this appendix, we provide two motivating examples of economic models that map into the basic framework described in the Introduction.

A.1 Oligopoly market example

The first example concerns price expectations in two interrelated oligopoly markets, one for product group A and the other for product group B . Assume that N firms compete in each of these two markets on price, but produce differentiated products. For simplicity, we assume that the cost of producing products is the same across all firms, $C_i(q_i^A, q_i^B) = F + c(q_i^A + q_i^B)$, where the parameter c denotes the marginal cost of production, and q_i^A and q_i^B are the quantities produced by firm i for markets A and B , respectively.

The demand function for the product of firm i in market A is given by

$$D_i^A(p_i^A, \bar{p}^A) = \alpha^A - \beta^A p_i^A + \gamma^{AA} \bar{p}^A, \quad \beta^A > 0, \quad (\text{A.1})$$

where p_i^A is the price of the product and $\bar{p}^A = \frac{1}{N} \sum_j p_j^A$ is the average price charged by all firms in market A . Parameter γ^{AA} specifies the relationship between the products of the various firms in this market: if $\gamma^{AA} > 0$ the goods are substitutes, whereas if $\gamma^{AA} < 0$ they are complements. The linear demand as in (A.1) can be derived assuming that consumers have a linear-quadratic utility function.

The demand function for the product of firm i in market B is given by

$$D_i^B(p_i^B, \bar{p}^A, \bar{p}^B) = \alpha^B - \beta^B p_i^B + \gamma^{AB} \bar{p}^A + \gamma^{BB} \bar{p}^B, \quad \beta^B > 0, \quad (\text{A.2})$$

where p_i^B is the price of the product in market B and \bar{p}^B is the average price charged by all firms in this market. Parameter γ^{BB} specifies the relationship between the products of various firms in market B in the same way that γ^{AA} does for market A : if $\gamma^{BB} > 0$ the goods are substitutes, whereas if $\gamma^{BB} < 0$ they are complements.

When $\gamma^{AB} \neq 0$ in (A.2) markets A and B become interconnected in the sense that the demand in market B depends on the average price in market A .³⁵ When $\gamma^{AB} > 0$ the higher average price of products sold in market A leads to a higher demand for products in market B , and so the markets become substitutes, whereas for $\gamma^{AB} < 0$ the markets are complements. Such between-markets connections can occur when products traded in market A are the intermediate goods used in the production of goods traded in market B . Examples include the markets for oil and gasoline, or the markets for milk and cheese.

Firm i maximizes its total profit

$$\pi_i(p_i^A, p_i^B; \bar{p}^A, \bar{p}^B) = (p_i^A - c) D_i^A(p_i^A, \bar{p}^A) + (p_i^B - c) D_i^B(p_i^B, \bar{p}^A, \bar{p}^B). \quad (\text{A.3})$$

Assuming that the number of firms is sufficiently large so that firms can ignore their own impact on average prices, the first-order conditions are given by

$$\begin{aligned} \frac{\partial \pi_i}{\partial p_i^A} &= \alpha^A - \beta^A p_i^A + \gamma^{AA} \bar{p}^A - (p_i^A - c) \beta^A = 0 \\ \frac{\partial \pi_i}{\partial p_i^B} &= \alpha^B - \beta^B p_i^B + \gamma^{AB} \bar{p}^A + \gamma^{BB} \bar{p}^B - (p_i^B - c) \beta^B = 0. \end{aligned}$$

Solving this system of equations, we find that the optimal prices of firm i for markets A and B , denoted as p_i^{A*} and p_i^{B*} , can be computed as

$$\begin{pmatrix} p_i^{A*} \\ p_i^{B*} \end{pmatrix} = \mathbf{M} \begin{pmatrix} \bar{p}^A \\ \bar{p}^B \end{pmatrix} + \mathbf{d}, \quad (\text{A.4})$$

³⁵Another way to connect the markets would be by modeling economies/dis-economies of scope in the cost functions, say, as in Bulow et al. (1985).

where

$$\mathbf{M} = \begin{pmatrix} \frac{\gamma^{AA}}{2\beta^A} & 0 \\ \frac{\gamma^{AB}}{2\beta^B} & \frac{\gamma^{BB}}{2\beta^B} \end{pmatrix} \quad \text{and} \quad \mathbf{d} = \begin{pmatrix} \frac{\alpha^A + c\beta^A}{2\beta^A} \\ \frac{\alpha^B + c\beta^B}{2\beta^B} \end{pmatrix}$$

There exist a Nash Equilibrium, where all firms set the same prices in each of the markets, which are given by the elements of $(\mathbf{I} - \mathbf{M})^{-1}\mathbf{d}$. Regardless of whether firms are in this equilibrium, given the average prices in both markets, each firm maximizes its own profit by setting prices as close as possible to the target prices, p_i^{A*} and p_i^{B*} , as defined in (A.4). Indeed, from (A.3), the firm's profit can be written as

$$\pi_i(p_i^A, p_i^B; \bar{p}^A, \bar{p}^B) = -\beta^A (p_i^A - p_i^{A*})^2 - \beta^B (p_i^B - p_i^{B*})^2 + \text{Const},$$

where ‘Const’ denotes the terms that do not depend on p_i^A and p_i^B . Thus, our 2DBC game provides subjects with the same incentives as they have in the price-setting game, to be as close to each target as possible. Our payoff function (3) penalize them for deviations from the targets on both markets in the same way, which corresponds to $\beta^A = \beta^B$.

A.2 New Keynesian model example

For another motivating example of a 2D linearized system studied by macroeconomists, consider a New Keynesian (NK) model. The model consists of equations for inflation, π_t , the output gap, y_t , and a policy rule for the nominal interest rate, i_t :

$$\begin{aligned} \pi_t &= \pi_{t+1}^e + \kappa y_t, \\ y_t &= y_{t+1}^e - \varphi(i_t - \pi_{t+1}^e - r), \\ i_t &= \lambda(\pi_t - \pi). \end{aligned}$$

In these equations, π_{t+1}^e denotes expected inflation, y_{t+1}^e is the expected output gap, r is the exogenously fixed real rate of interest, π is the central bank's inflation target

and $\kappa > 0$, $\varphi > 0$ and $\lambda > 0$ are parameters. The adaptive learning literature (see Eusepi and Preston, 2018 for a recent review) treats the expectations as subjective, and not necessarily rational. Branch and McGough (2009) and Kurz et al. (2013) show that when expectations are heterogeneous, the dynamics can be described by the above relations with average expectations, $\bar{\pi}_{t+1}^e$ and \bar{y}_{t+1}^e , replacing π_{t+1}^e and y_{t+1}^e on the right side. Such a system can be rewritten in the matrix form as

$$\begin{pmatrix} \pi_t \\ y_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} \bar{\pi}_{t+1}^e \\ \bar{y}_{t+1}^e \end{pmatrix} + \mathbf{d}, \quad (\text{A.5})$$

where

$$\mathbf{M} = \Omega \begin{bmatrix} 1 + \kappa\varphi & \kappa \\ 1 & \varphi(1 - \lambda) \end{bmatrix}, \quad \mathbf{d} = \Omega \begin{pmatrix} \kappa\varphi \\ \varphi \end{pmatrix} (\lambda\pi + r), \quad \Omega \equiv \frac{1}{1 + \kappa\varphi\lambda}.$$

Individuals in the NK model with heterogeneous expectations aim to minimize errors in their expectations about inflation and output gap, π_t , and y_t . These variables correspond to the ‘target’ numbers in our 2DBC, and they indeed depend in (A.2) on the average guesses. Our 2DBC model then maps to the NK model with heterogeneous expectations, albeit with different timing. Indeed, whereas the NK model has forward looking expectations, expectations in our version of the 2DBC game are contemporaneous.³⁶

Suppose we consider standard calibrations of the New Keynesian model where $0 < \kappa \leq 1$, $0 < \varphi \leq 1$ (see, e.g., Galí, 2015). Then, one can show that for values of the policy parameter $0 < \lambda < \hat{\lambda}$, the system is a saddle with a positive unstable root. For values of $\lambda > \hat{\lambda}$, the system has two stable eigenvalues $-1 < \mu_1 < 0 < \mu_2 < 1$, so that the system becomes a sink and the steady state is globally stable. More generally,

³⁶Some recent “learning to forecast” experiments use systems like (A.5) and ask participants to predict the next period values of both variables. We do not pursue this approach here, as we want to stay close to the standard BC experiments. The simpler timing of our experiment leads to a cleaner theoretical analysis. Sonnemans and Tuinstra (2010) note, that for one-dimensional forecasting game experiments, the difference in the timing of forecasts (contemporaneous or one-step ahead) is not crucial for convergence properties.

we wish to abstract from a particular macroeconomic application and to consider a variety of different 2D linear systems as characterized by whether the steady state is globally stable (a sink), saddle-path stable, or unstable, two explosive roots (a source).

B Equilibria of the 2DBC Game

Consider the 2DBC game with N players, individual strategies in $[0, 100] \times [0, 100]$, the payoff function (3), and having matrix \mathbf{M} from one of our four treatments. The strategy profile where each player submits $\mathbf{z}_i = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{d} = (90, 20)$ is a Nash equilibrium. Indeed, as the averages for this profile coincide with the guesses, the targets become

$$\mathbf{z}^* = \mathbf{M}(\mathbf{I} - \mathbf{M})^{-1}\mathbf{d} + \mathbf{d} = \mathbf{M}(\mathbf{I} - \mathbf{M})^{-1}\mathbf{d} + (\mathbf{I} - \mathbf{M})(\mathbf{I} - \mathbf{M})^{-1}\mathbf{d} = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{d}.$$

Thus, the target values coincide with the individual guesses for both numbers, ensuring the maximum possible payoff of 100 points for each participant, according to (3). Any deviation from this profile will lead to a lower payoff of the deviating participant.

We will now show that asymmetric equilibria are impossible in this game. Assume that there is a pair of different strategies in an equilibrium profile of a participant. Without loss of generality, we assume that the a -numbers differ. Denote the left-most and right-most among all a -numbers in the profile as a_L and a_R , respectively. Then for the average of all a -numbers, \bar{a} , we must have $a_L < \bar{a} < a_R$. Let us show that if the target $a^* \leq \bar{a}$, then the participant increases his own payoff, by replacing a_R with $a_R - \varepsilon$ with a small enough $\varepsilon > 0$ and keeping the same b -number.³⁷ Indeed, when (a_R, b) changes to $(a_R - \varepsilon, b)$, the average of the a -numbers decreases by ε/N , leading to a new target $a^* - \varepsilon m_{11}/N$. The distance between the new target and

³⁷If $a^* \geq \bar{a}$, the payoff is increased replacing (a_L, b) with $(a_L + \varepsilon, b)$.

the new strategy is $a_R - a^* - \varepsilon(N - m_{11})/N$. The target for the b -number changes to $b^* - \varepsilon m_{21}/N$. Overall, there is a gain of $\varepsilon(N - m_{11})/N$ for the a -target and there might be a maximal loss of $\varepsilon|m_{21}|/N$ for the b -target. In total, there is an increase in performance as measured by the sum of absolute deviations of the a - and b -numbers by at least $\varepsilon(N - m_{11} - |m_{21}|)/N$ and this quantity is strictly positive in all our treatments, when $N = 10$.

Since only symmetric equilibria are possible, an argument similar to the one above shows that all participants submit the same a and the same b -numbers in any equilibrium. Thus, \mathbf{z}^E is the unique NE in the interior of the strategy space. Boundary-value NE are possible only when participants, submitting boundary guesses, receive targets outside the admissible strategy space. It is then straightforward to check the following:

- There are no boundary equilibria for the a -number in our treatments, except for the **Source** treatment where 0 and 100 are two such equilibria. In this treatment, when $a = 0$, strategy $b = 38$ completes the equilibrium profile, and when $a = 100$, strategy $b = 18$ completes the equilibrium profile.
- There are no boundary equilibria for the b -number in our treatments, except for the **SaddlePos** treatment where 0 and 100 are two such equilibria. In this treatment, when $b = 0$ or $b = 100$, strategy $a = 90$ completes the equilibrium profile.

This completes the proofs of the statements of Section 2 as summarized in Table 1.

C Experimental Instructions

The following instructions (for treatment **Sink**) were distributed to every participant and read loudly. The equations were explained and displayed on the screen during the whole experiment.

Welcome to this experiment in economic decision-making. Please read these instructions carefully as they explain how you earn money from the decisions you make in today's experiment. There is no talking for the duration of this session. If you have a question at any time during the experiment, please raise your hand and your question will be answered in private. Kindly silence and put away all mobile devices.

General information: You are in a group of 10 participants including you. There are 15 successive time periods $1, 2, \dots, 15$ in this experiment. The same participants will be in your group during this experiment in all 15 periods. In each period you have to choose two numbers, an "A-number" and a "B-number". Your choice for each number must be between 0 and 100 inclusive which means that 0 or 100 are also allowed. Every participant in your group also chooses a pair of numbers, an "A-number" and a "B-number", between 0 and 100 inclusive. After you and every other participant in your group have chosen a pair of numbers, two target values will be determined as explained below, one for the "A-numbers" and another for the "B-numbers". Your earnings from this experiment will depend on how close your chosen numbers are to the corresponding target values. The closer your numbers are to the target values, the greater will be your earnings.

Determination of the target values: In each time period, you and all other participants choose one "A-number" and one "B-number". After all participants have chosen their numbers, the target values for the "A-numbers" and "B-numbers" are determined on the basis of the average values of all 10 "A-numbers" and all 10 "B-numbers" (including your own choices). The average of all "A-numbers" is computed as the sum of all ten "A-numbers" chosen by the participants in your group in this period and divided by 10. The average of all "B-numbers" is determined as the sum of all ten "B-numbers" chosen by participants in your group in this period and divided

by 10. The corresponding target values, A^* and B^* , are computed as follows:

$$A^* = 30 + (2/3) \times \text{Average of all "A-numbers"}$$

$$B^* = 75 - (1/2) \times \text{Average of all "A-numbers"} - (1/2) \times \text{Average of all "B-numbers"}.$$

Note that the target value, A^* , depends on the choices made by all participants in your group (including yourself) of “A-numbers”, whereas the target value B^* depends on the choices of all participants in your group (including yourself) of both “A-numbers” and “B-numbers”.

Here is an example: For simplicity suppose that there are only 3 participants in a group and in some period they submit the following “A-numbers”:

50.50 100 20.25

and “B-numbers”

10.79 30 0.

Then, the average of the “A-numbers” is $(50.50 + 100 + 20.25)/3 = 170.75/3 = 56.92$ and the average of the “B-numbers” is $(10.79 + 30 + 0)/3 = 40.79/3 = 13.6$. Given these averages, the target values are:

$$A^* = 30 + (2/3) \times 56.92 = 67.95; \text{ and}$$

$$B^* = 75 - (1/2) \times 56.92 - (1/2) \times 13.6 = 39.74.$$

All results are rounded to 2 decimals. Please, note that this example is for illustration purposes only. The actual group size in the experiment is 10 participants.

About your task: The experiment lasts for 15 periods. Each period your only task is to choose two numbers, an “A-number” and a “B-number”. Your goal is to choose these numbers to be as close as possible (in absolute value) to the target values A^* and B^* which will be determined in each period, using the procedure explained above

after all participants in your group have made their choices. Both numbers you choose should be between 0 and 100, inclusive. You may enter a real number with up to 2 decimals.

About your earnings: Your decision will determine how many points you receive each period. Your earnings will be based on the sum of all your points over the 15 periods, with 1 point = 1 US cent. In addition, you are guaranteed to receive \$7 as a show-up payment. Your points in each of the 15 period are based on how close your “A-number” and “B-number” are to the target values A^* and B^* and are calculated (by the computer program) as follows:

$$\text{Your payoff in points each period} = \frac{500}{5 + |\text{your “A-number”} - A^*| + |\text{your “B-number”} - B^*|}$$

where $|\cdot|$ is an absolute value (deviation), e.g., $|3 - 5| = 2$, $|5 - 1| = 4$. Notice several things. First, if you submit the exact target values for A^* and B^* , you receive the maximum payoff of 100 points. Second, deviations from the target values, A^* or B^* , have an equal effect on your payoffs; the further you are away from either target value, the lower is your payoff in points. Third, the payoff of all 10 participants in your group is determined in a similar way. Finally, all 10 participants (including you) can earn the maximum of 100 points if all choose the exact target values for A^* and B^* . For your convenience we provide a table on page 4 showing how your payoff changes depending on the deviations of your A and B choices from the target values, A^* and B^* .

Example continued: In the example above, a participant who submitted the “A-number” 50.50 and the “B-number” 10.79, misses the target value $A^*= 67.95$ by 17.45 and the target value $B^*= 39.74$ by 28.95, and, so his payoff in points is given by:

$$\frac{500}{5 + |50.50 - 67.95| + |10.79 - 39.74|} = \frac{500}{5 + 17.45 + 28.95} = 9.73 \text{ “points”}$$

(rounded to 2 decimals).

Information and Record Keeping: At the end of each period, you will see a screen that reports the results of the just completed period. Specifically, you will be informed of:

- The “A-number” and “B-number” that you submitted for the period
- The average of all “A-numbers” and the average of all “B-numbers” submitted by group members for the period
- The computed target values A^* and B^* for the period
- Your points earned for the period

Please record this information on your record sheet for each period under the appropriate headings. When you are done recording this information, click on the OK button in the bottom right corner of your screen.

So long as the 15th period has not yet been played, we will move to the next period decision screen. On that screen you will have to type the “A-number” and “B-number” for the current period. Additionally, you will see a history table displaying for each prior period:

- your chosen “A-number” and your chosen “B-number”
- averages of all “A-numbers” and all “B-number”
- computed target values A^* and B^*
- your points earned

Points Table The table gives the number of points for a given discrepancy of “A-number” from the target value A^* (the first column) and a given discrepancy of “B-number” from the target value B^* (the first row).

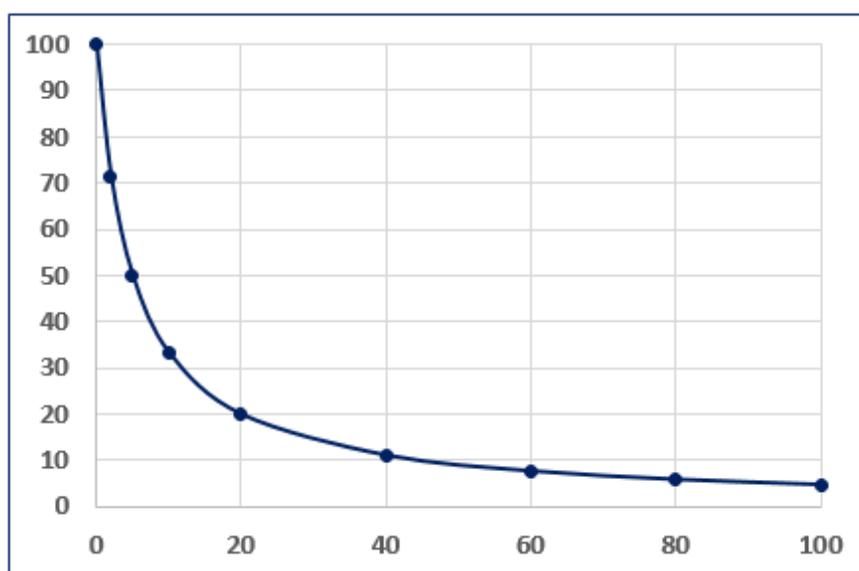
The figure below shows the relation between the number of points you score (vertical

 A-A* \ B-B* 	0	1	2	3	4	5	10	20	30	40	50	60	80	100
0	100.00	83.33	71.43	62.50	55.56	50.00	33.33	20.00	14.29	11.11	9.09	7.69	5.88	4.76
1	83.33	71.43	62.50	55.56	50.00	45.45	31.25	19.23	13.89	10.87	8.93	7.58	5.81	4.72
2	71.43	62.50	55.56	50.00	45.45	41.67	29.41	18.52	13.51	10.64	8.77	7.46	5.75	4.67
3	62.50	55.56	50.00	45.45	41.67	38.46	27.78	17.86	13.16	10.42	8.62	7.35	5.68	4.63
4	55.56	50.00	45.45	41.67	38.46	35.71	26.32	17.24	12.82	10.20	8.47	7.25	5.62	4.59
5	50.00	45.45	41.67	38.46	35.71	33.33	25.00	16.67	12.50	10.00	8.33	7.14	5.56	4.55
10	33.33	31.25	29.41	27.78	26.32	25.00	20.00	14.29	11.11	9.09	7.69	6.67	5.26	4.35
20	20.00	19.23	18.52	17.86	17.24	16.67	14.29	11.11	9.09	7.69	6.67	5.88	4.76	4.00
30	14.29	13.89	13.51	13.16	12.82	12.50	11.11	9.09	7.69	6.67	5.88	5.26	4.35	3.70
40	11.11	10.87	10.64	10.42	10.20	10.00	9.09	7.69	6.67	5.88	5.26	4.76	4.00	3.45
50	9.09	8.93	8.77	8.62	8.47	8.33	7.69	6.67	5.88	5.26	4.76	4.35	3.70	3.23
60	7.69	7.58	7.46	7.35	7.25	7.14	6.67	5.88	5.26	4.76	4.35	4.00	3.45	3.03
80	5.88	5.81	5.75	5.68	5.62	5.56	5.26	4.76	4.35	4.00	3.70	3.45	3.03	2.70
100	4.76	4.72	4.67	4.63	4.59	4.55	4.35	4.00	3.70	3.45	3.23	3.03	2.70	2.44

axis) and the combined discrepancy

$$|\text{your "A-number"} - A^*| + |\text{your "B-number"} - B^*|$$

of your chosen numbers and the targeted values (horizontal axis). Notice that the table presents only some possibilities for your point earnings (the table is not exhaustive) and that the number of points you earn decreases more slowly as your discrepancies from the two target values increase.



Additional Information

- Before the experiment starts you will have to take a short quiz which is designed to check your understanding of the instructions.
- At the end of the experiment you will be asked to answer a questionnaire before you are paid. Your answers will be processed in nameless form only. Please fill in the correct information.
- During the experiment any communication with other participants, whether verbal or written, is forbidden. The use of phones, tablets or any other gadgets is not allowed. Violation of the rules can result in removal from the experiment.
- You may use the back side of your record sheet as scratch paper if you wish. Do not write your name on this, only write your ID number on the front side.
- Please follow the instructions carefully at all the stages of the experiment. If you have any questions or encounter any problems during the experiment, please raise your hand and the experimenter will come to help you.

Please ask any question you have now!

D Classification of participant choices

Table 7 shows the results of the classification of participants' choices during periods 2 to 8, as defined in Section 3. The first two columns report treatment, session and participants ID in the session. Next, for each period t from 1 to 8, we report the k that minimizes the distance $d_t(k)$ in Eq. (5). The row **Average** reports an average level k for the corresponding period for each treatment excluding 'E' from the computation. The column **Type** shows our type classification based on the data for periods 2–8. This classification is an integer reflecting the level k or the type is labeled 'E', 'L', 'M' for **E**quilibrium, **L**earning and **M**ixing types, respectively). If no such type could be identified, then 'n/c' is reported. The frequencies found in this classification are shown, for each treatment, in Fig. 4.

In addition to the classification based on distance, we analyze the choices of every participant using a regression analysis. For each participant i , we estimated the following regression assuming the same coefficients for a and b -choices

$$\begin{aligned} a_{i,t+1} &= \beta_{i,0} a_t^s(0) + \beta_{i,1} a_t^s(1) + (1 - \beta_{i,0} - \beta_{i,1}) a_t^s(2) + \varepsilon_{i,t,a}, \\ b_{i,t+1} &= \beta_{i,0} b_t^s(0) + \beta_{i,1} b_t^s(1) + (1 - \beta_{i,0} - \beta_{i,1}) b_t^s(2) + \varepsilon_{i,t,b}. \end{aligned} \tag{D.1}$$

Recall that $a_t^s(k)$ and $b_t^s(k)$ denote the level- k choices in period $t \geq 1$ in session s , where individual i was present. The regression is estimated on the same periods 2-8 that we used to classify subjects, that is in (D.1) we use $t = 1, \dots, 7$, leading to 14 data points for each individual. Regression (D.1) describes the individual guesses as a combination of level-0, level-1 and level-2 choices. The regression coefficients, $\beta_{i,0}$ and $\beta_{i,1}$, are shown in the corresponding columns of Table 7. Coefficients that are significantly different from zero at a 5% significance level are reported in boldface.

We then test, for each participant, whether their choices can be assigned to one of three levels, 0, 1 or 2, and we report the results in the last column of Table 7, **Supported models**. We list there the restricted models, out of the general model (D.1), that cannot be rejected at a 5% significance level with the F-test. Integers ‘0’, ‘1’ and ‘2’ denote level-0, level-1 and level-2 models, that is, the model (D.1) with restrictions for ‘0’: $\beta_{i,0} = 1$ and $\beta_{i,1} = 0$, for ‘1’: $\beta_{i,1} = 1$ and $\beta_{i,0} = 0$, and for ‘2’: $\beta_{i,0} = \beta_{i,1} = 0$. In addition, ‘M’ denotes the mixed level 0-1 model, that is, the model where (D.1) is restricted with $\beta_{i,0} + \beta_{i,1} = 1$. This model includes both the level-0 and level-1 models. Of course, participant’s behavior may be consistent with several of these models, which can also be seen from the last column of Table 7.

Fig. 6 visualizes the results of this individual level regression analysis in the simplex for our four treatments. This scatter plot shows the estimated coefficients $(\beta_{i,0}, \beta_{i,1})$ from regression (D.1), with one point corresponding to one participant. The shape of the marker reflects the treatment that each participant belongs to (see the legend for details). We observe that for all treatments, there is a substantial heterogeneity

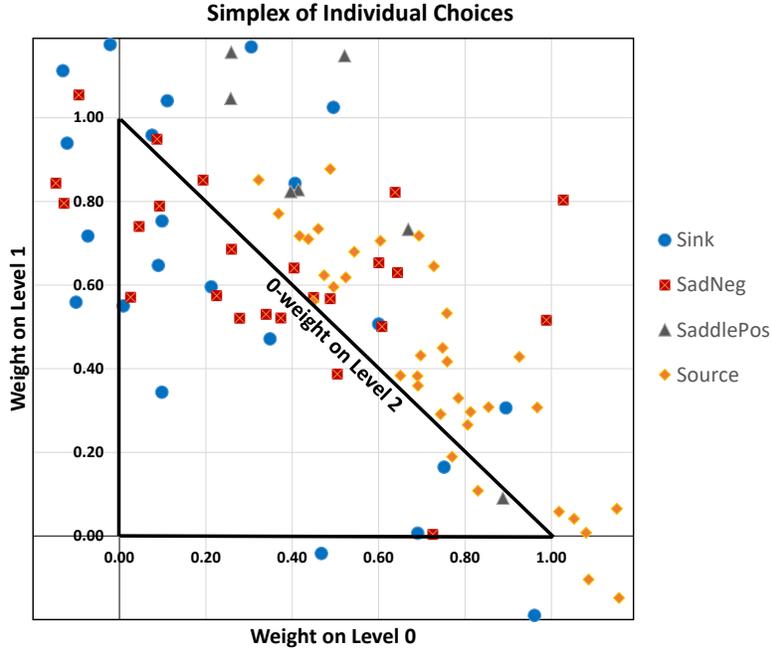


Figure 6: Simplex of individual choices illustrating the results of regression (D.1).

of participants in the estimated behavior. Another tendency that is common across treatments, is that many choices are located along the edge of the simplex with 0 weight on level 2 (i.e., far from the origin), showing that very few subjects have behavior consistent with level-2. In the converging, **Sink** and **SaddleNeg** treatments, more points are observed in the top-left corner of the simplex, those that correspond either to level 1 or to the mixed level 0-1 model. By contrast, in the **Source** treatment, many choices are close to the bottom-right corner with many level-0 choices. Overall, it can be seen that the regression-based analysis is consistent with the classification based on the distances to the level- k choices illustrated in Fig. 3. Many participants are consistently type 0 or 1, but some participants are also mixing between these types.

Session	ID	Period								Type	Regr. coeffs		Supported models
		1	2	3	4	5	6	7	8		$\beta_{i,0}$	$\beta_{i,1}$	
Sink 1	1	1	0	0	0	0	0	0	0	0	0.28	3.56	
	2	0	1	0	2	1	2	1	2	L	0.35	0.47	M, 1
	3	1	0	1	1	0	0	0	2	M	0.41	0.84	M
	4	1	3	0	0	0	0	1	2	L	0.31	1.17	0, M, 1
	5	2	2	0	0	1	3	3	4	L	-0.02	1.17	M, 1
	6	1	1	1	1	1	1	1	2	1	-0.09	1.3	M, 1
	7	0	0	2	2	2	3	1	1	n/c	-0.12	1.26	M, 1
	8	1	0	0	1	1	2	2	3	L	0.96	-0.19	0, M
	9	1	1	0	1	2	1	1	2	L	0.69	0.01	0, M, 1, 2
	10	0	0	1	1	1	2	1	2	L	-0.15	1.8	M, 1
Sink 2	1	0	0	0	0	0	0	1	1	0	1.01	-0.22	0, M
	2	1	1	1	0	2	1	3	1	L	0.47	-0.04	M, 1, 2
	3	0	1	1	1	1	1	1	1	1	-0.13	1.11	M, 1
	4	1	0	0	0	1	1	1	1	L	0.6	0.51	0, M
	5	3	1	1	1	1	1	1	1	1	0.09	0.65	M
	6	E	0	4	0	0	0	0	E	0	3.34	-4.16	0, M, 1, 2
	7	2	1	4	0	0	0	0	0	0	-0.57	5.41	0, 1
	8	2	1	1	0	1	1	1	1	1	0.21	0.6	M
	9	E	1	1	1	0	1	1	0	1	0.05	2.56	0, M, 1
	10	0	1	0	0	0	0	1	1	L	0.5	1.02	0
Sink 3	1	3	0	0	0	0	0	0	0	0	0.89	0.31	0
	2	1	1	3	0	1	1	0	0	M	-0.38	1.59	M
	3	0	1	2	2	1	1	0	3	n/c	-0.12	0.94	M, 1
	4	1	2	2	2	1	1	1	0	n/c	0.01	0.55	1
	5	1	2	2	2	1	1	1	2	n/c	-0.1	0.56	
	6	1	1	0	2	4	4	E	E	L	-0.34	1.51	M, 1, 2
	7	0	0	0	0	0	0	0	0	0	1.36	1.27	0, M
	8	1	1	0	1	1	2	1	1	1	0.08	0.96	M, 1
	9	1	1	1	2	3	3	1	1	L	0.1	0.34	
	10	1	1	1	1	1	1	1	3	1	0.1	0.75	M, 1
Sink 4	1	1	1	1	0	0	1	1	1	1	0.11	1.04	M, 1
	2	0	0	0	0	0	1	1	1	L	0.75	0.16	0, M
	3	1	1	1	1	1	1	1	1	1	0	1.22	M, 1
	4	0	0	0	0	0	0	0	0	0	1.55	-0.02	0, M
	5	1	2	2	1	1	1	2	1	n/c	-0.32	1.13	
	6	1	3	1	1	1	0	1	1	1	-0.33	1.29	M, 1, 2
	7	0	0	0	2	1	1	2	1	L	-0.16	4.14	0, M, 1
	8	0	E	0	0	0	0	0	0	0	1.73	2.05	0, M, 1, 2
	9	2	2	2	2	1	1	1	1	n/c	-0.07	0.72	
	10	2	2	2	2	2	2	1	2	2	-0.23	0.64	2
Average		0.9	0.9	1.0	0.8	0.9	1.1	0.9	1.2				

Session	ID	Period								Type	Regr. coeffs		Supported models
		1	2	3	4	5	6	7	8		$\beta_{i,0}$	$\beta_{i,1}$	
SadNeg 1	1	0	1	1	1	1	1	2	2	1	0.09	0.79	M, 1
	2	0	0	1	1	0	0	1	0	M	0.99	0.52	0, M
	3	4	1	1	1	2	0	0	0	M	0.28	0.52	
	4	0	0	1	0	0	0	0	0	0	0.61	0.5	M
	5	0	3	0	0	1	1	1	2	L	0.34	0.53	M, 1
	6	1	0	1	1	1	1	1	1	1	0.63	1.3	0, M, 1
	7	3	3	1	1	1	2	1	0	n/c	-0.23	0.91	M
	8	E	E	E	E	E	E	E	E	E	-1.2	1	M
	9	1	0	0	1	1	2	1	2	L	2.06	0.3	0, M, 1
	10	1	0	0	0	0	0	0	0	0	1.25	0.09	
SadNeg 2	1	0	E	1	1	1	0	0	0	M	0.05	0.74	M, 1
	2	0	0	0	1	E	1	0	1	L	0.5	0.39	0, M
	3	0	2	3	1	1	2	1	1	n/c	0.73	0	0, M
	4	0	3	2	1	1	2	2	1	n/c	-0.59	1.17	M, 1
	5	0	1	1	0	0	1	0	0	M	0.64	0.63	M
	6	E	E	E	E	E	1	E	2	E	-1.17	1.02	M
	7	1	1	0	1	E	1	4	E	L	0.45	0.57	M, 1
	8	E	1	1	1	E	2	2	2	L	-0.15	0.84	M
	9	1	E	1	1	E	1	2	1	n/c	-0.13	0.8	
	10	0	0	0	0	0	0	1	1	0	1.88	0.01	M
SadNeg 3	1	0	1	3	3	1	0	1	E	n/c	-0.29	1.37	M, 1
	2	1	1	2	3	2	1	2	E	n/c	0.03	0.57	
	3	0	1	3	2	1	2	3	0	n/c	-0.32	0.8	M
	4	0	1	1	1	1	1	3	1	1	0.09	0.95	M, 1
	5	4	1	0	3	0	1	0	1	M	0.4	0.64	M, 1
	6	0	1	1	2	1	3	3	3	L	0.19	0.85	M, 1
	7	1	1	3	1	1	2	3	2	L	-0.26	0.88	M
	8	1	4	0	0	0	0	0	0	0	0.86	-0.34	0, 2
	9	0	0	2	1	E	3	2	1	L	0.49	0.57	M
	10	0	0	0	0	1	0	0	1	0	1.76	0.09	0, M, 1
SadNeg 4	1	0	1	3	1	4	1	3	1	n/c	-0.09	1.05	M, 1
	2	1	3	2	0	0	3	E	3	L	0.23	0.57	0, M, 1
	3	1	0	1	0	1	1	0	E	L	0.37	0.52	
	4	1	1	1	E	2	E	E	1	L	-0.22	0.8	M
	5	0	1	E	0	3	3	2	3	L	-0.24	0.74	
	6	3	1	1	0	1	1	1	2	1	0.26	0.69	M
	7	1	2	1	2	2	1	2	4	n/c	-0.37	0.7	M
	8	0	1	1	1	2	1	0	1	1	1.03	0.8	
	9	0	1	1	0	1	0	1	0	M	0.6	0.65	M
	10	1	1	1	1	1	1	1	2	1	0.64	0.82	
Average		0.7	1.1	1.1	0.9	1.1	1.1	1.3	1.2				

Session	ID	Period								Type	Regr. coeffs		Supported models
		1	2	3	4	5	6	7	8		$\beta_{i,0}$	$\beta_{i,1}$	
SadPos 1	1	0	0	0	4	E	0	E	E	n/c	4.73	-5.63	0, M
	2	0	1	1	2	0	1	1	1	1	-0.91	2.91	M
	3	0	0	0	2	3	1	1	1	n/c	0.06	1.3	
	4	1	1	0	0	0	0	2	3	L	-0.13	2.01	0
	5	E	1	2	2	2	1	2	1	n/c	-1.56	3.94	
	6	0	1	1	1	E	1	0	1	1	-1.95	5.04	0
	7	0	1	2	2	2	2	1	1	n/c	-1.57	3.89	M
	8	E	1	0	0	0	0	0	0	0	-1.14	3.77	0
	9	0	1	0	0	E	0	0	1	M	0.67	0.73	0
	10	0	1	2	3	0	1	0	1	n/c	-2	4.87	M
SadPos 2	1	E	0	1	0	1	1	0	0	M	0.26	1.05	
	2	1	0	1	0	1	1	1	1	1	-0.05	1.46	
	3	0	0	1	0	0	1	1	1	L	2.28	-2.43	0
	4	0	0	0	0	0	0	0	0	0	1.12	-0.23	0
	5	E	1	1	0	2	1	1	E	n/c	-3	6.92	0
	6	0	0	0	0	1	1	0	0	0	0.41	0.83	0
	7	0	1	1	0	0	1	0	0	M	-1.34	4.03	0, M
	8	E	E	1	0	E	0	4	1	L	4.42	-5.63	0, M
	9	E	E	1	1	E	1	E	E	n/c	-2.34	6.17	0
	10	1	0	1	0	1	1	1	0	L	-0.62	2.51	
SadPos 3	1	0	2	1	1	1	1	1	2	1	-0.86	2.86	M
	2	0	1	1	1	2	1	1	0	1	-1.36	3.76	M
	3	0	0	0	0	1	1	1	1	L	0.89	0.09	0
	4	1	1	1	1	0	3	1	1	1	-1.5	4.09	M
	5	1	2	2	3	2	2	2	1	2	-2.77	6.18	M
	6	1	0	4	2	2	2	0	2	n/c	-1.28	3.7	
	7	0	0	0	0	1	E	1	0	L	-0.28	2.51	0
	8	0	1	0	1	1	1	2	1	1	0.4	0.82	
	9	0	0	1	0	0	0	1	0	0	0.26	1.16	0
	10	0	0	0	0	0	E	4	4	L	0.52	1.15	0
SadPos 4	1	1	1	1	2	3	2	4	1	L	-1.3	3.66	
	2	E	1	2	1	0	2	1	1	n/c	-1.14	3.37	
	3	E	1	0	0	1	1	1	1	1	-0.57	2.63	0
	4	0	0	1	0	0	E	0	0	0	1.73	-0.94	0
	5	E	1	1	0	0	0	0	0	0	-1.34	3.88	
	6	0	1	0	4	4	4	0	4	L	-1.74	4.49	M
	7	0	1	1	1	2	1	1	1	1	-1.6	4.12	M
	8	0	1	2	1	3	0	E	2	L	-3.69	7.9	0, M
	9	E	1	0	1	1	1	1	1	1	-0.44	2.26	
	10	E	1	1	1	4	1	1	1	1	-0.71	2.68	
Average		0.2	0.7	0.9	0.9	1.2	1.0	1.0	1.0				

Session	ID	Period								Type	Regr. coeffs		Supported models		
		1	2	3	4	5	6	7	8		$\beta_{i,0}$	$\beta_{i,1}$			
Source 1	1	0	1	0	0	0	0	0	0	E	0	0.93	0.43		
	2	1	1	0	0	0	0	0	3	1	L	0.6	0.71	0, M	
	3	0	0	0	0	0	0	0	0	0	0	1.02	0.06	0, M	
	4	0	0	0	4	4	1	1	0	0	n/c	0.77	0.19	0, M	
	5	1	2	1	3	3	3	0	0	0	E	0.81	0.27	0, M	
	6	0	1	0	0	0	1	0	0	3	L	0.75	0.45	0	
	7	0	1	1	4	1	1	1	0	0	1	0.47	0.62	M	
	8	E	1	E	0	0	1	2	1	1	L	0.76	0.42	0, M	
	9	0	1	4	4	4	1	1	0	0	n/c	0.69	0.38	0, M	
	10	0	1	4	4	4	E	1	4	4	n/c	0.46	0.73	0, M	
Source 2	1	0	1	0	0	1	3	1	3	3	L	0.52	0.62		
	2	2	1	3	0	1	0	1	0	0	M	0.78	0.33	0	
	3	0	0	0	1	1	4	0	3	3	L	0.83	0.11	0, M	
	4	0	1	3	1	1	1	1	3	3	1	0.37	0.77		
	5	0	1	3	3	2	4	1	1	1	n/c	0.42	0.72		
	6	0	2	0	2	2	E	1	E	1	n/c	1.2	-0.13	0, M	
	7	1	3	0	1	2	0	0	0	3	n/c	0.69	0.36	0, M	
	8	E	0	0	0	0	0	0	0	0	0	1.15	0.06	M	
	9	2	0	0	0	0	4	0	0	0	0	1.09	-0.1	0, M	
	10	0	1	E	4	2	3	4	0	0	L	0.85	0.31	0, M, 1	
Source 3	1	0	0	0	3	0	0	0	0	E	0	0.97	0.31	0, M	
	2	1	2	1	1	1	4	1	1	1	1	0.45	0.56	M, 1	
	3	1	0	1	1	0	1	4	E	1	L	0.81	0.3	0, M, 1	
	4	0	0	3	1	1	4	3	1	1	L	0.5	0.6	M	
	5	0	1	1	0	1	4	E	1	1	n/c	0.7	0.43	0, M	
	6	0	0	0	1	1	0	0	0	0	0	0.76	0.53		
	7	3	3	3	E	1	0	0	0	0	n/c	0.69	0.72	0	
	8	2	3	1	1	E	3	1	1	1	n/c	0.54	0.68	0, M, 1	
	9	0	0	0	1	2	3	0	E	1	L	1.08	0.01	0, M	
	10	0	1	0	0	1	2	0	2	2	L	0.65	0.38	0, M	
Source 4	1	0	1	0	2	3	1	1	1	1	n/c	0.32	0.85		
	2	0	1	4	3	1	1	0	E	1	n/c	0.49	0.88	0, M	
	3	2	2	0	0	0	0	0	0	0	0	1.31	-0.5	0, M	
	4	E	0	0	1	E	0	1	1	1	L	0.73	0.64	0, M	
	5	3	2	0	0	0	0	0	0	0	0	1.32	-0.49	0, M	
	6	0	E	0	2	3	1	0	2	2	n/c	0.74	0.29	0, M	
	7	0	3	3	3	3	2	3	1	1	3	0	1.31		
	8	1	3	1	3	3	2	0	1	1	n/c	0.44	0.71	0, M, 1	
	9	0	0	0	0	3	0	0	0	0	0	1.16	-0.15	0, M	
	10	0	1	3	0	0	0	0	0	0	0	1.05	0.04	0, M	
Average		0.5	1.1	1.1	1.4	1.4	1.5	0.8	1.0						

Table 7: Classification of individual types and individual regressions.

E Proof of Proposition 4.1

Let K be the largest level of thinking in population and let $\mathbf{H}(K)$ denote the operator that describes the dynamics of the CH model (the map from period $t - 1$ to period t variables) in deviations from the PONE, in the case where $f_E = 0$. We need to prove that $\mathbf{H}(K) = \prod_{i=1}^K (\lambda_i \mathbf{M} + (1 - \lambda_i) \mathbf{I})$ with λ 's as defined in the Proposition.

The level-0 agents play the previous averages and hence contribute with operator \mathbf{I} . The level-1 agents play the best response and contribute with operator \mathbf{M} . If level 1 is the highest level, then we conclude that $\mathbf{H}(1) = f_1 \mathbf{M} + (1 - f_1) \mathbf{I}$, as required.

Assume now that level 1 is not the highest level. Note that from the perspective of agents of level 2, the average play of 0 and 1 level agents is $\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}$ with $\lambda_1 = f_1 / (f_1 + f_0)$ and $1 - \lambda_1 = f_0 / (f_1 + f_0)$. The level-2 agents' best response is then $\mathbf{M}(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I})$. If level 2 is the highest level in population, then we can complete the model, by weighting this best response with f_2 and assigning the remaining weight $1 - f_2$ to the average play of level-0 and 1 agents as defined above. This leads to the operator

$$\begin{aligned} \mathbf{H}(2) &= f_2 \mathbf{M}(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}) + (1 - f_2)(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}) = \\ &= (f_2 \mathbf{M} + (1 - f_2) \mathbf{I})(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}). \end{aligned}$$

This is exactly what the proposition claims for this case, where level 2 is the highest level in population.

Assume now that level 2 is not the highest level. From the perspective of agents of level-3, the average play of levels 0, 1 and 2 agents is given by

$$\begin{aligned} \lambda_2 \mathbf{M}(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}) + (1 - \lambda_2)(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}) = \\ = (\lambda_2 \mathbf{M} + (1 - \lambda_2) \mathbf{I})(\lambda_1 \mathbf{M} + (1 - \lambda_1) \mathbf{I}). \end{aligned}$$

The first term in the first line multiplies the normalized fraction $\lambda_2 = f_2 / (f_0 + f_1 + f_2)$

to the best reply that agents of level-2 play (as we determined above), whereas the second term in the first line assigns the remaining weight to the average play of the level 1 and level 0 agents.³⁸ The level 3 agents' best response is then $\mathbf{M}(\lambda_2\mathbf{M} + (1 - \lambda_2)\mathbf{I})(\lambda_1\mathbf{M} + (1 - \lambda_1)\mathbf{I})$. If level 3 is the highest level in the population, then we can complete the model, by weighting this best response with f_3 and assigning the remaining weight $1 - f_3$ to the average play of levels 0, 1 and 2 agents. This leads to the operator

$$\begin{aligned} \mathbf{H}(3) &= f_3\mathbf{M}(\lambda_2\mathbf{M} + (1 - \lambda_2)\mathbf{I})(\lambda_1\mathbf{M} + (1 - \lambda_1)\mathbf{I}) + \\ &\quad + (1 - f_3)(\lambda_2\mathbf{M} + (1 - \lambda_2)\mathbf{I})(\lambda_1\mathbf{M} + (1 - \lambda_1)\mathbf{I}) = \\ &= (f_3\mathbf{M} + (1 - f_3)\mathbf{I})(\lambda_2\mathbf{M} + (1 - \lambda_2)\mathbf{I})(\lambda_1\mathbf{M} + (1 - \lambda_1)\mathbf{I}). \end{aligned}$$

This is exactly what the statement claims in this case, when level 3 is the highest level in population.

Continuing in the same manner, the statement is verified for any K .

F Supplementary Online Appendix

Supplementary Online Appendix can be found online at *insert link*.

³⁸This, of course, can be split into two parts, $(1 - \lambda_2)\lambda_1\mathbf{M} = f_1/(f_0 + f_1 + f_2)\mathbf{M}$ represents the normalized play of agents of level-1 and $(1 - \lambda_2)(1 - \lambda_1)\mathbf{I} = f_0/(f_0 + f_1 + f_2)\mathbf{I}$ represents the normalized play of agents of level 0.