Planar Beauty Contests*

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

May 1, 2020

Abstract

We explore whether agents can learn the steady state of a two-dimensional, coupled linear system. We find that they can learn the steady state if the eigenvalues of the system are both stable or if the steady state is saddlepath stable with the one unstable eigenvalue being negative. They cannot learn the steady state if it is saddlepath stable with the one unstable eigenvalue being positive, or if there are two unstable eigenvalues. We show that our results cannot be explained by naïve or homogeneous level-k learning, but *can* be explained by adaptive learning or a mixed cognitive levels model.

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

JEL Classification: C30, C92, D83, D84.

*We thank Peter Bossaerts, Marcus Giamattei, Frank Heinemann, Cars Hommes, Rosemarie Nagel and Jan Tuinstra for useful discussions. The paper benefited from comments of participants in the 2018 CEF Conference in Milan, the 2018 ESA meeting in Berlin, the 2018 Behavioral and Experimental workshop in Singapore, the 2018 BEAM-ABEE workshop in Amsterdam, the 2019 Australasian Economic Theory Workshop (AETW), the 2019 NAMES conference in Seattle, the 2019 Australia New Zealand Workshop in Experimental Economics (ANZWEE) in Melbourne and seminars at the European University at St. Petersburg, Deakin University, Monash University, Macquarie University, and University of Sydney. 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

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). While the theoretical literature considers multi-dimensional systems, the experimental literature has mainly focused on learning in a single dimension.[1] In this paper, we provide experimental evidence on learning behavior in a coupled, linear *multivariate* setting. Such systems are theoretically tractable, and their dynamical properties can be expressed in terms of their eigenvalues. For example, in macroeconomics, it is common to use linearized multivariate dynamical systems where there are interactions in the determination of the different variables. The eigenvalues of such systems are important for deciding whether the solution is *determinate*, i.e., if there exists a unique non-explosive path, in which case comparative statics can be precisely described.[2]

To fix ideas, consider a linear (or linearized) economic model of the form

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} a_t^e \\ b_t^e \end{pmatrix} + \boldsymbol{d}\,, \tag{1}$$

where a_t^e and b_t^e denote the expected time t values of the variables a_t and b_t, and where the 2×2 matrix $\mathbf{M}$ and 2×1 vector $\boldsymbol{d}$ are exogenously given. There are a number of economic settings whose reduced forms map into this basic framework. Appendix A provides two examples, one from the industrial organization literature

[1] For instance, in many macroeconomic applications, agents learn how to forecast a single variable, inflation, or the price level alone (see, e.g., Hommes, 2011). In the game theory literature, subjects typically learn about a single action choice or the mixed strategy to play in a game (see, e.g., Camerer, 2003, Chapter 6).

[2] According to Blanchard and Kahn (1980), determinacy of the solution requires that the number of non-predetermined variables equals the number of unstable eigenvalues of the system. For instance, in the two-dimensional optimal growth model, there is one non-predetermined variable (consumption), and one predetermined variable (capital). Determinacy of the solution requires that the steady state exhibits the "saddle-path property", having one stable eigenvalue and one unstable eigenvalue. A strong rationality assumption is required for the agents to be exactly on the saddle-path.

and another from the macroeconomic literature on monetary policy. Under rational expectations, where $a_t^e = a_t$ and $b_t^e = b_t$, the system (1) has a unique equilibrium,[3] given by $(\mathbf{I} - \mathbf{M})^{-1}\boldsymbol{d}$. We refer to this as the rational expectation equilibrium (REE).

The main question we ask in this paper is *whether the eigenvalues of the reduced form, planar linear system* (1) *matter for the ability of human subjects to learn the REE.*

We address this question in the laboratory by proposing an extended, multi-dimensional version of the well-known Keynesian Beauty Contest game.[4] Specifically, we ask subjects to guess *two* numbers, instead of the usual one, in a coupled system (1), where the average of all subjects' guesses replace the expectations.[5] We refer to our game as the "Planar Beauty Contest" (PBC) game. The PBC game has a unique interior Nash equilibrium that corresponds to the REE of (1). Unlike in the univariate version, however, in the coupled planar system, guesses about the realization of one variable can matter for realizations of the other variable, and *vice versa*. Thus, the planar multivariate system makes the task of coordinating on equilibrium behavior even more complex than in the standard one-dimensional case, and indeed, the question of how agents learn in this more complex setting is our main focus.

It is difficult to know, a priori, how subjects will react to the additional complexity of the PBC. On the one hand, complexity might slow or even prevent learning of the REE, as subjects struggle to understand how their guesses for one variable should

[3] $\mathbf{I}$ denotes the identity matrix. Throughout the paper, the matrix $\mathbf{I} - \mathbf{M}$ is assumed to be invertible.

[4] Mauersberger and Nagel (2018) provide a comprehensive, up-to-date review of the extensive literature that followed the original paper of Nagel (1995). The name "Beauty Contest game" first appeared in Duffy and Nagel (1997) and Ho et al. (1998), referring to Keynes' comparison of investment behavior with newspaper contests (of Keynes' era) where readers had to choose the face (among several) that they believed would be chosen by most other readers. As we discuss below, the game became popular in developing and testing level-k and cognitive hierarchy models of strategic sophistication. The main focus of this paper, however, is on the *dynamical* properties of planar systems and on *learning* models that can explain our experimental data.

[5] This is the usual practice in the large literature on "learning-to-forecast" experiments (Hommes, 2011). See Sonnemans and Tuinstra (2010) for a connection between the beauty contest and learning-to-forecast experiments.

impact on their guesses for the other. On the other hand, the greater complexity of the PBC might promote greater introspection; subjects might respond to the greater complexity by reasoning harder about the problem at hand, in the extreme perhaps even applying "fixed–point reasoning" to directly solve for the steady state.

To address the question of learning, we ask subjects to play the PBC game repeatedly and we examine the dynamics of both variables, a_t and b_t, over time. Note that, in the interest of simplicity and consistent with the experimental beauty contest literature, there are no explicit intertemporal dynamical linkages from one period to the next in (1), that is, our basic framework is a static one. However, *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 variable's previous period value (as under a naïve expectations assumption), then model (1) is effectively a two-dimensional, coupled *first-order dynamical* system. Of course, agents may be more sophisticated and use a variety of different learning rules. Indeed, we consider the fit of a number of such rules to our experimental data, essentially running a contest among these various learning models.[6] We wish to emphasize that the interdependencies in the two-dimensional system that we study make for a more challenging test of these learning models. An important consequence of the interdependence between variables is that the *eigenvalues* of matrix $\mathbf{M}$ can matter for the stability of the steady state of the system. These eigenvalues are a main focus of our analysis. We report on four treatments of the PBC game that vary the eigenvalues of this 2-dimensional system.

In answer to our main question, we find that experimental subjects *are* able to learn the equilibrium steady state of the system (corresponding to the REE) in the "sink" case, where both eigenvalues are stable (less than 1 in absolute value). More remarkably, they are also able to learn the equilibrium steady state of the so-called "saddlepath negative" case, which is a particular version of the saddlepath solution where one eigenvalue is stable and the other unstable eigenvalue has a negative sign.

[6]Recently Anufriev and Hommes (2012) and Anufriev et al. (2019) also used experimental data to compare various learning models in the univariate framework.

By contrast, we find that in the other two cases, the "source" case, where both eigenvalues are unstable, and the "saddlepath positive" case, where one eigenvalue is stable and the other is unstable and positive, subjects are unable to coordinate on the steady state. Our findings thus suggest that steady states exhibiting the saddlepath property *can* be learned by agents who do not begin a process of social interaction with rational expectations knowledge of the equilibrium. This result was surprising to us, and, we believe, it is a very important finding. Indeed, saddlepath steady states are a main focus of macroeconomic analysis,[7] but their empirical relevance can be questioned because they require substantial computational efforts as well as coordination between participants to be reached. At the same time, we found that such convergence is not a *general* property of saddlepath stable solutions. This result indicates that reliance on the magnitude of the eigenvalues alone does not suffice to determine whether the equilibrium is learnable or not.

Our experimental results led us to ask which learning model could possibly explain the dynamics observed across all four treatments of our experiment. We study the convergent properties of several models analytically, and we use the experimental data to estimate and compare the fit of these learning models.

Specifically, we generalize the level-k model of Nagel (1995), Stahl and Wilson (1994, 1995), Costa-Gomes et al. (2001), and Costa-Gomes and Crawford (2006) and the cognitive hierarchy model of Camerer et al. (2004) and Chong et al. (2016), to our dynamic, planar setting.[8] We then compare the fit of the naïve learning, past averaging, adaptive and level-k models to our experimental data. Importantly, we find that homogeneous level-k models cannot consistently explain the behavior of subjects in some versions of our planar beauty contest game. In particular, in the saddle negative treatment, the homogeneous level-k model predicts *nonconvergence*

[7]The saddlepath environment is particularly attractive, since, by contrast with the sink case, the reaction of the economy to shocks to fundamentals can be uniquely determined only in this environment. Indeed, when the steady state has the sink property, there are infinitely many paths by which the system can adjust following a perturbation.

[8]For recent surveys of level-k and cognitive hierarchy models and applications, see Crawford et al. (2013) and Mauersberger and Nagel (2018).

to the steady state while human subjects consistently converge upon the steady state in this treatment. This difference reveals the advantage of adding another dimension to the forecasting task, as it allows us to more clearly assess the predictions of various learning models and to differentiate them from one another in terms of their fit to the data. By contrast, we find that an adaptive learning model *can* explain the convergence/divergence outcomes observed in *all four* of our experimental planar beauty contest treatments. Further, we show how this adaptive learning model can be reinterpreted as a mixed cognitive levels model. Finally, we show that this adaptive learning model fares the best across all of the models that we consider according to the out-of-sample predictive mean squared error.

1.1 Related literature

The Keynesian Beauty Contest game provides a simple and well-studied framework for understanding the extent to which agents can learn to acquire rational expectations in a simple forecasting game. The original game, as first explored experimentally by Nagel (1995) and Stahl and Wilson (1995), can be viewed as a single equation system. The aim of the game is for each player $i = 1, \dots, N$ to make a guess, a_i, about some target value, $a^* = p\bar{a}$, where $\bar{a}$ is the average of all N players' guesses and $p < 1$. Each player's guess, a_i, is restricted to lie in the interval $[0, 100]$. Since $p < 1$, the unique, dominance solvable Nash equilibrium prediction is for all to guess 0. However, this outcome is rarely observed in the first few rounds of play. Instead, as Nagel (1995), Stahl and Wilson (1994, 1995) and many others showed, subjects tend to apply only limited number of steps of iterated elimination of dominated strategies in the first round.[9] The experimental evidence suggests that there are sizeable fractions of level-0, 1, 2, and level-∞ types; the latter types simply solve the fixed point problem to

[9] Many subjects begin with some initial reference point for the average choice, a_0; typically $a_0 = 50$, though others may choose $a_0 = p \times 100$. Given this initial choice, "level-1" player types choose $p \times a_0$ as their guess. Slightly more sophisticated, "level-2" types presume that all other participants are level-1 types who will guess $p \times a_0$ and so these level-2 types best respond by choosing $p \times (p \times a_0)$ as their guess. Generalizing, a "level-k" type guesses $p^k \times a_0$.

compute the Nash equilibrium. In repeated play, agents use the history of outcomes and continue to apply level-k rules to the winning numbers and converge to the Nash equilibrium.

Subsequent experimental research by Güth et al. (2002), Sutan and Willinger (2009), and Lambsdorff et al. (2013) examine affine univariate systems where the target value, $a^* = p\bar{a} + d$, and $|p| < 1$. This setup yields a non-corner or "interior" equilibrium solution, with negative or positive feedback depending on the sign of p. Positive (negative) feedback systems are related to strategic complements (substitutes) as first noted by Haltiwanger and Waldman (1985). In a repeated experimental setting, univariate systems with negative feedback have been found to more reliably converge to the steady state as compared with positive feedback systems where convergence is slower or not observed over the time horizon of the experiment – see, e.g., Fehr and Tyran (2008), Heemeijer et al. (2009) and Hommes (2013). This observed difference for univariate and stable systems motivates our consideration of two different versions for the saddle-path solution, one where the unstable eigenvalue is positive and one where it is negative. As we will show, this distinction can also matter for whether adaptive learning behavior converges to the steady state of multivariate (planar) coupled systems.[10]

There is some experimental research that seeks to understand how human subjects form expectations of *two* interrelated endogenous variables in a coupled dynamical system, arising out of a reduced form of the "New Keynesian" model of monetary policy. See in particular, Adam (2007), Assenza et al. (2020), Pfajfar and Žakelj (2016) among others. These studies consider specific parameterizations of a forward-looking, first-order planar system of the form $\boldsymbol{x}_t = \mathbf{M}\boldsymbol{x}^e_{t+1} + \boldsymbol{d}$ and they consider the extent to which human subjects can learn to form forecasts in a manner that

[10]More recently, Benhabib et al. (2019) study affine univariate systems where the target value $a^* = p\bar{a} + d_i$, under both negative and positive feedback ($-1 < p < 1$), and where d_i is a private signal for each player i; in one of their treatments, d_i is a random draw from a mean zero distribution, and in that case, subjects find it easy to play the equilibrium strategy of using their private signal as their guess.

is consistent with the rational expectations equilibrium predictions (Adam, 2007) and/or whether monetary policy rules can move the system from an indeterminate (unstable) steady state to a determinate (locally stable) one (Assenza et al., 2020, Pfajfar and Žakelj, 2016). Our paper is related to this work in that we also consider learning in two dimensional systems. However, we are not interested in considering a specific model (e.g., the New Keynesian model) or the role of monetary policy; rather we are interested in how the eigenvalues of planar dynamical systems impact on the learning of the rational expectations equilibrium. Toward that end, we do not have an explicitly dynamical framework; subjects in our experiment forecast current (and not future) realizations of target variables as in the standard beauty contest literature. However, as noted above, if agents are adaptive learners, such behavior effectively makes our system dynamical. Further, unlike in the New Keynesian model experiments which do not provide subjects with much information about the data generating process and simply ask subjects to forecast inflation and the output gap, we provide our subjects with *full* information about the data generating process for the two endogenous variables, i.e., they know both the matrix $\mathbf{M}$ and the vector $\boldsymbol{d}$, just as in the experimental beauty contest literature. Thus our design enables subjects to immediately solve or "educe" the steady state of the model, though they must also consider the strategic uncertainty they face regarding the expectations formed by other subjects as explained below.[11] Further, we vary the stability of the system we study by changing the parameterization of the system in such a way that the steady state does not change across all of our different experimental treatments; this makes the analysis of convergence/divergence under the different stability conditions much clearer.

There is some 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

[11] Bao and Duffy (2016) also study learning under complete information but for a *univariate* negative feedback system.

saddlepath property can be learned under certain conditions, specifically, if agents have forecast rules of the same form as the equilibrium saddle-path relationship between the two variables. By contrast, in this paper, we do not endow subjects with such knowledge, though as noted, we do present them with the data generating equations of the system. Further, we show via simulations, that simple adaptive learning dynamics, initialized according to an uninformative prior belief that the means of the two variables, a and b, will begin at the midpoint of the guessing interval, closely track the behavior of the subjects in our experiment.

To summarize, our paper contributes to the literature in four ways. First, we design a novel experiment, the PBC game, which is a stylized description of many economic environments where the agents confront multiple interrelated decisions. Second, we further our understanding of the empirical relevance of saddlepath stable solutions in planar systems, by showing both that they can be learned, but also that this is not a general property. Third, we run a competition across several learning models using our experimental data. Our results suggest that adaptive expectations are the best approach to characterizing the experimental outcomes. Finally, we build a bridge between the adaptive learning literature of macroeconomics and the level-k learning models widely used in the game theoretical literature.

The remainder of the paper is organized as follows. Section 2 describes our PBC game and the four treatments of our experiment. The experimental results are presented in Section 3. Section 4 introduces a number of different learning models and asks whether their dynamics are consistent with features of experimental data. In Section 5 we estimate these models using our experimental data and compare their fit. Section 6 concludes with a summary and some suggestions for future research. There are also a number of Appendices, which contain the experimental instructions, proofs, additional figures, and robustness checks.

2 Experimental design

An experimental session consists of a group of N individuals who participate in T repetitions of the planar beauty contest game. In each period, each player i submits a pair of numbers, $(a_{i,t}, b_{i,t})$. Based on these "guesses", the group averages

$$\bar{a}_t = \frac{1}{N}\sum_{i=1}^{N} a_{i,t} \quad \text{and} \quad \bar{b}_t = \frac{1}{N}\sum_{i=1}^{N} b_{i,t}$$

are computed. Finally, the "target" values a_t^* and b_t^* are defined as

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

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

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

points. Thus, participants are motivated to submit their guesses as close as possible to the target values. Guessing both targets exactly would bring a maximum reward of 100 points. Note that deviations from the target values decrease the participant's payoffs by an equal amount.[12]

We study the behavior of experimental subjects in 4 treatments that differ in the matrix $\mathbf{M}$ and the vector $\boldsymbol{d}$. The treatments and the corresponding parameters are presented in the first two columns of Table 1. 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.

[12]The hyperbolic function in (3) strongly penalizes the smallest deviations from the target values, giving participants robust incentives to guess the targets precisely. The hyperbolic payoff structure has been used in Adam (2007) and Assenza et al. (2020).

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

Table 1: Four Experimental Treatments.

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). In the first three treatments, we set $m_{11} = 2/3$, so that the first eigenvalue is stable. This specific choice for the parameter of the uncoupled equation is one that is commonly used in the univariate beauty contest game; indeed the first equation of our planar system can be thought of as a version of the classic, 2/3 of the average, beauty contest game (albeit with an interior solution). Depending on the second eigenvalue m_{22} we have then three treatments. In the **Sink** treatment $m_{22} = -1/2$, and so both eigenvalues are stable. In the "saddle with negative feedback" (**SaddleNeg**) treatment, $m_{22} = -3/2$. Thus, the second eigenvalue is unstable (leading to the saddlepath property) and negative. In the "saddle with positive feedback" (**SaddlePos**) treatment, $m_{22} = 3/2$, and so the second eigenvalue is again unstable but now it is positive. For our last **Source** treatment, we sought to have two unstable eigenvalues and so we set $m_{11} = 3/2$ and $m_{22} = -3/2$.

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

$$\begin{pmatrix} a^E \\ b^E \end{pmatrix} = (\mathbf{I} - \mathbf{M})^{-1} \boldsymbol{d}\,, \tag{4}$$

both targets a^* and b^* coincide with the corresponding guesses. Thus, every participant gets the maximum possible payoff, which proves that (4) is the Nash equilibrium.

Referring to specification (1) that motivated our study, we call this the rational expectations equilibrium (REE). To make a comparison across treatments as clean as possible, we have chosen vectors $\boldsymbol{d}$ in different treatments so that they have *the same* equilibrium, $(a^E, b^E) = (90, 20)$.

As is typical in beauty contest experiments, we restrict the range of guesses submitted by participants in order to eliminate an effect of extreme guesses (mistakes or outliers).[13] That is, we inform participants that both a_i and b_i must belong to the interval $[0, 100]$.[14] Depending on the treatment, this boundary restriction 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. Note that in any boundary equilibrium, the target value for at least one variable will be outside of the $[0, 100]$ range. Therefore, none of these boundary equilibria is a REE. Furthermore, given the payoff function (3), payoffs in these boundary equilibria are always lower than in the unique interior equilibrium, $(90, 20)$.

Note that the only difference between the **SaddleNeg** and **SaddlePos** treatments is in the *sign* of the unstable eigenvalue. In particular, the eigenvalues in these two saddle treatments have the same absolute values. Making the eigenvalue negative introduces strategic substitutability or "negative feedback" to the b-number strategy in the **SaddleNeg** treatment: higher guesses imply a lower target. Instead, when the eigenvalue is positive, the b-number guess generates a strategic complementarity or "positive feedback" with higher guesses implying a larger target.

The experimental sessions were conducted in the Experimental Social Science Laboratory at the University of California, Irvine. We conducted four sessions for

[13] We think that our results would continue to hold without the bounded guessing interval, though learning might take longer and be noisier. Keeping the bounded guessing interval, however, is useful for thinking about initial conditions in our level-k analysis, and facilitates comparisons with the level-k literature.

[14] In fact, these restrictions motivated us to select the particular Nash equilibrium of $a^E = 90$ 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.

each of the four treatments. In all sessions, we asked a group of $N = 10$ participants to play the same PBC game for $T = 15$ consecutive periods. As our focus is on the convergence properties in the dynamic environment, we did not rematch participants so that the same 10 participants played all 15 periods.[15] 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). Every subject participated in one session only and thus we report data from $4 \times 4 \times 10 = 160$ subjects. The subjects were all undergraduate students at the University of California, Irvine.

At the start of each session, the Instructions (see Appendix C) were read loud 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, see 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 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.

After all the subjects submitted their guesses, the computer program calculated the average of all 10 submitted a-numbers and all 10 submitted b-numbers and determined the target values for the period (according to the treatment conditions for the matrix $\mathbf{M}$ and vector $\boldsymbol{d}$). Given the target values, subjects' payoffs in points were determined according to the payoff function (3). Period t ended with a second

[15]Rematching would make past information less relevant and thus interfere with learning. This approach is standard in this literature, e.g., Nagel (1995).

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 (based on the period averages), 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 in the lower part of the main decision screen for each period. This table showed, for each past period, their chosen numbers, the averages for both numbers, the computed target values, and the points they earned for those periods according to the payoff function (3). 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

We begin with an overview of the main results. The four panels of Fig. 1 show representative examples of the evolution of the averages, $\bar{a}$ and $\bar{b}$, across all four treatments, **Sink**, **SaddleNeg**, **SaddlePos**, and **Source**. The time evolution of $\bar{a}$ (the thick red line) and of $\bar{b}$ (the thin blue line) are graphed in relation to the dashed lines indicating the equilibrium levels, $a^E = 90$ and $b^E = 20$.[16] We observe the convergence of both the a and b numbers to those levels in the **Sink** and **SaddleNeg** treatments and we do not observe convergence in the **SaddlePos** and **Source** treatments.[17]

The most remarkable finding of the experiment is the contrast in dynamics be-

[16] We chose session 1 of each treatment as a representative session. Overall, the dynamics in the four sessions of each of the four treatments are very similar to those of session 1 as depicted in Fig. 1. See Figs. 7 to 10 in the Online Appendix F where we show the dynamics of the average guesses for each session individually. Figs. 11 to 14 in the Online Appendix F show the dynamics of the target values, a^* and b^*, for each session as well.

[17] They approach payoff-dominated, boundary Nash equilibria in these two treatments, see Table 2.

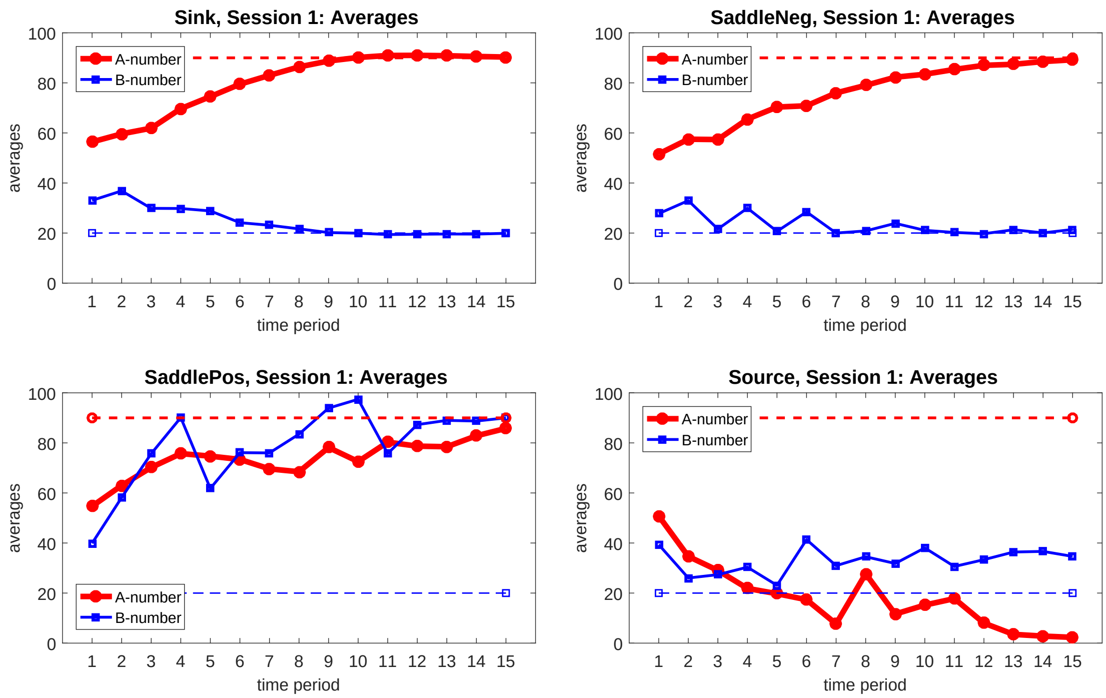


Figure 1: Typical dynamics of the average values, $\bar{a}$ and $\bar{b}$, observed in four different treatments of the experiment. The dashed lines indicate the REE.

tween two saddle treatments. Participants are able to learn a steady state with the saddlepath property under negative feedback, quickly converging to the REE. Participants *are not* able to learn a steady state with the saddlepath property under positive feedback. The outcome in these two saddle treatments are different, although these treatments differ only in the sign of the target for b-number as dependent on the average of the b-guesses.

Table 2 shows the average guesses of the a- and b-numbers in the first and the last periods for 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 REE (the last row of the table).

From this table we can make the following observations. 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 Figure 1. Specifically, in

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
REE		90	(0)	90	(0)	20	(0)	20	(0)	100.00	100.00	1500.00

Table 2: Statistics of average guesses and their deviations from the REE $(90, 20)$ for a and b numbers and of subjects' payoffs for each session of the corresponding treatment.

the first period the a- and b-guesses in all treatments are significantly different from the REE.[18] We observe convergence to the REE $(90, 20)$ in the **Sink** and **SaddleNeg** treatments and we do not observe convergence to this equilibrium in the **SaddlePos** and **Source** treatments. Indeed, using the four sessions level observations for the **Sink** and **SaddleNeg** treatments, we cannot reject the null that the average a- and b-guesses in period 15 are different from 90 and 20, respectively. However, we can reject this same hypothesis for period 15 for a- or b-guesses of the **SaddlePos** and **Source** treatments.[19] In the **SaddlePos** treatment, the b-guesses end up even further

[18]For each treatment, using $4 \times 10 = 40$ individual a-guesses, we reject the null hypothesis that the mean of them is 90. We also reject the null hypothesis that the mean of the initial b-guesses is 20 for each treatment. All p-values are less than 0.0025.

[19]We do not use individual level observations in the test because those are not independent within each session. Instead we use session level averages which are independent. The p-values for **Sink**, **SaddleNeg**, **SaddlePos** and **Source** for a-numbers are 0.076, 0.488, 0.034, 0.0001, respectively,

from the equilibrium than they were in the initial period. Similarly, the a-guesses in the **Source** treatment move away from the equilibrium. Finally, the b-guesses in the **Source** treatment only slightly move towards the REE.

Second, the heterogeneity in guesses (as measured by the standard deviations) is similar in all treatments for the first period choices and typically decreases to a very small number in the 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.[20] 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 two other 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 in the **SaddleNeg** treatment was apparently quicker which allowed subjects to earn the largest total payoff across all four treatments.

3.1 Choices in the first period

In the first period of our experiment, the participants had the same information as in the standard, one-shot beauty contest game. Specifically, since the first equation is

and for b-numbers are 0.106, 0.457, 0.001, 0.0001, respectively.

[20]One 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.

decoupled from the second, guessing the a-number is exactly equivalent to playing the standard game with the target given by $m_{11}\bar{a} + d_1$. For this reason, it is interesting to look at their choices in the first period alone.

Figure 2 presents histograms of individual guesses for the a-number in period 1. The left panel combines all groups from treatments **Sink**, **SaddleNeg** and **SaddlePos**, and the right panel shows the histogram for the **Source** treatment.[21] We observe that first period guesses are heterogeneous with a large spike around the middle of the guessing interval, 50. The remaining choices are concentrated in the right half of the interval for treatments **Sink**, **SaddleNeg** and **SaddlePos**, and they are concentrated in the left half of the interval for treatment **Source**. Notice that a few participants in the **Sink**, **SaddleNeg** and **SaddlePos** treatments (but not the Source treatment) submitted the REE quantity for the a-number, 90, as represented by the thick vertical line.

A standard approach for classifying choices in this game employs level-k reasoning. According to this classification (see, e.g., Nagel, 1995), level-0 subjects submit guesses distributed uniformly over the guessing interval $[0, 100]$, having a mean of $(0 + 100)/2 = 50$.[22] Subjects of level-1 best respond to this level-0 behavior and submit $50m_{11} + d_1$. Subjects of level-2 best respond to level-1 choices, and so on. When the best responses fall outside of the interval of strategies $[0, 100]$, they are truncated to the closest boundary. To illustrate this approach, we superimpose on the histograms shown in Figure 2 three vertical lines, corresponding to the levels of 0, 1, 2 (the legend specifies the corresponding values for these different levels and level 3 as well). Note that in the treatments with $m_{11} = 2/3$, the sequence of levels monotonically increases and converges to $a^E = 90$ in an infinite number of steps.

[21] The first three treatments have an identical equation for the a-number, with the target given by $(2/3)\bar{a} + 30$. In the **Source** treatment, the target for the a-number is $(3/2)\bar{a} - 45$. The test for differences between group means rejects the null of equal means of a-numbers between the first three treatments and the **Source** (p-value 0.002), but fails to reject the null of equal means of a-numbers across the first three treatments (p-value 0.871).

[22] An alternative specification for level-0 types has them guessing $m_{11} \times 100 + d$, as 100 is the upper bound of the guessing interval.

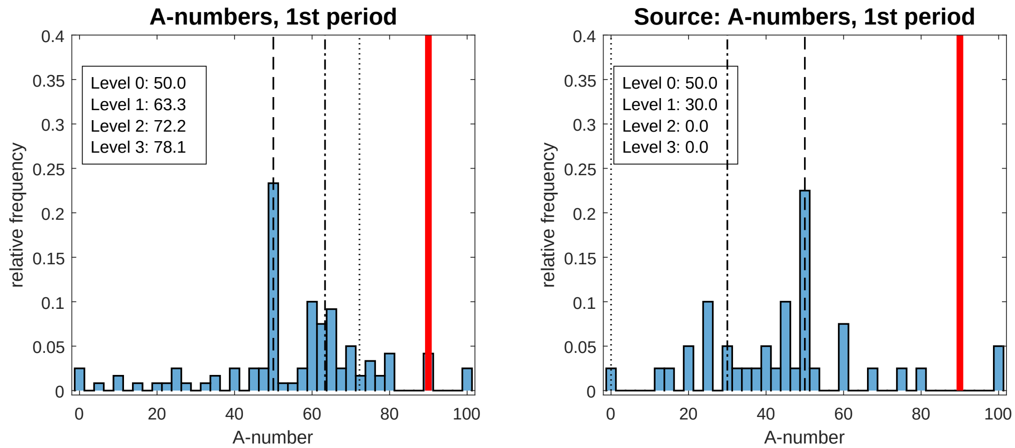


Figure 2: Frequencies of Period 1 a-choices and levels of reasoning (different dashed lines) and the REE (solid red line). *Left:* combined choices from **Sink**, **SaddleNeg** and **SaddlePos** treatments. *Right:* **Source** treatment.

Instead, in the case of the **Source** treatment, when $m_{11} = 3/2$ the sequence of level-k choices quickly decreases to 0 and stays there for any $k \geq 2$. As discussed above, this is one of the payoff-dominated boundary Nash equilibrium of our game with a bounded guessing interval.[23]

The target for the b-number is affected by the average guesses for *both* numbers, making this guessing task more complex. The four panels in Figure 3 show histograms of b-number choices for the first periods of the four different treatments, with a thick vertical line indicating the REE choice for b of 20. As in the case of the a-numbers, there is large heterogeneity in guesses, with spikes around 50 in all treatments, and skewness of the remaining choices to the left, except for the **SaddlePos** treatment where there are also many choices to the right of the mid-point of the guessing interval.[24]

To systematize these choices, we extend the level-k theory to the second dimension. As before, the subjects with level-zero are assumed to submit guesses distributed uniformly over the interval $[0, 100]$ with a mean of $(0 + 100)/2 = 50$. However, to

[23] If the average a-guess in the **Source** treatment is 0, the target will be -45. The closest possible guess to this target, i.e., the best response, is $a_i = 0$.

[24] On the other hand, in the **SaddlePos** treatment, we have more choices concentrated around the equilibrium value of 20.

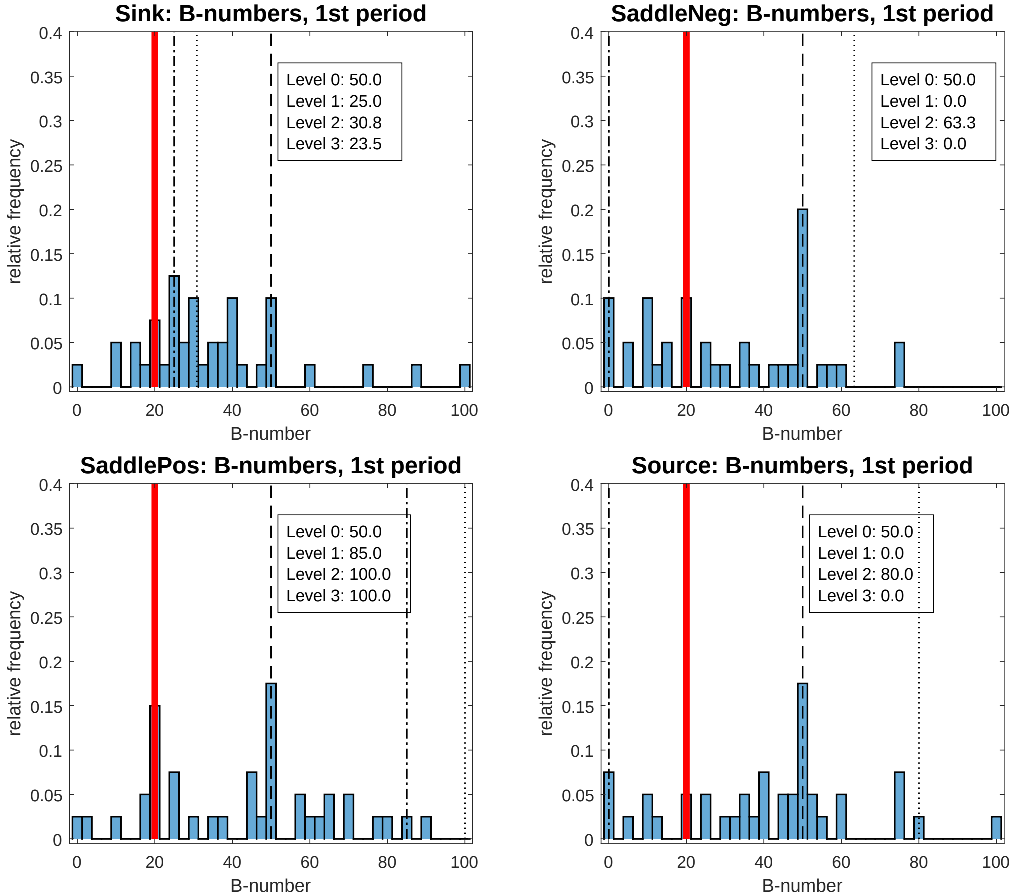


Figure 3: Frequencies of Period 1 b-choices and levels of reasoning (different dashed lines) and the REE (red line) for four treatments.

further define levels for the b-number, we must make assumptions about levels of rationality employed for *both* the a and b numbers. We will make the following "lock-step" assumption, that subjects at level k play a best response to others at *the same* level $k-1$ for both the a and b numbers. This means that subjects of level 1 best respond to a choice of 50 for both the a and b numbers and thus submit for their b-number guess $m_{21}50+m_{22}50+d_2$; subjects of level 2 best respond to level-1 guesses for both a and b numbers and so on. When the levels are outside $[0, 100]$ interval, the numbers are truncated to the closest boundary.

The step levels for the b-numbers following this procedure are shown by the vertical lines in the four panels of Figure 3 (the precise levels are again indicated in the legend).

Notice that in the **Sink** case, the levels oscillate and converge to 20. Oscillations make it difficult to identify the actual levels played by subjects, but we do observe spikes around level-1 and level-2 predictions in first period choices. In the **SaddleNeg** treatment, the level-1 choice is 0 (after truncation), where we also see a spike in our data. The level-2 choice is above 50, and the level-3 choice is 0 again, and so on.[25] In the **SaddlePos** treatment, the levels increase monotonically to 100, whereas in the **Source** treatment, similarly to **SaddleNeg**, the odd levels' choice is 0 and the process converges via oscillations to a two-cycle between 0 and 95.

Comparing this level-k model with the data of b-choices we conclude that the presence of a coupled variable in the PBC game leads to an even further decrease in the level of rationality for the b-number.[26] Looking back at statistics for the first period choices in Table 2, note that in 14 out of 16 sessions, the average a-guess is shifted from the middle point 50 towards the first level of rationality, whereas for b-guesses this happens in 11 cases out of 16.[27]

3.2 Dynamics and Speed of Convergence

We have already seen that the dynamics converged in both the **Sink** and **SaddleNeg** treatments and, moreover, a-guesses converged in **SaddlePos** as well. There are however some further dynamic features that we observe in our data. First, there is a clear difference in the type of convergence. The a-number converges almost monotonically from below in all three treatments. The b-number, however, converges almost monotonically in the **Sink** treatment and through oscillations in the **SaddleNeg**

[25] This process will converge to a two-cycle between 0 and 50.

[26] In other words, the aggregate data suggest that participants may have different levels of rationality in their a and b choices. Fig. 6 in the Online Appendix F illustrates this at the individual level.

[27] The only two exceptions for the a-number are Session 3 of the **Sink** treatment, where the average, 49.8, is just to the left of 50 and Session 1 of the **Source** treatment, where the average, 50.8, is just to the right from 50. For the b-number, the exceptions are all four sessions of the **SaddlePos** treatment and session 2 of **Source** treatment.

ε	**Sink**				**SaddleNeg**			
	Sess. 1	Sess. 2	Sess. 3	Sess. 4	Sess. 1	Sess. 2	Sess. 3	Sess. 4
20	5	8	4	4	5	2	4	3
10	7	11	6	11	9	5	5	5
5	8	-	8	-	11	6	8	7
1	10	-	14	-	-	12	12	9
0.5	10	-	-	-	-	14	13	9

Table 3: The first period when trajectory enters the ε-neighborhood of equilibrium in the experiment.

treatments.[28]

Second, there is a difference in the *speed* of convergence. We illustrate this difference in Table 3 by comparing the "first hit time", i.e., the first instance in which the trajectories for the average a and b numbers enter some ε-neighborhood of the equilibrium, in each of the four sessions of the **Sink** and **SaddleNeg** treatments. Let us fix $\varepsilon > 0$ and define the neighborhood as an open square around the equilibrium,

$$U_\varepsilon = \left\{(a,b) : |a - a^E| < \varepsilon \quad \text{and} \quad |b - b^E| < \varepsilon\right\}.$$

Let $t(\varepsilon)$ denote the period when the trajectory for the average values of the a and b numbers belong to the ε-neighborhood of the equilibrium for the first time. Formally, $t(\varepsilon)$ is such that $(\bar{a}_{t(\varepsilon)}, \bar{b}_{t(\varepsilon)}) \in U_\varepsilon$ and $(\bar{a}_t, \bar{b}_t) \notin U_\varepsilon$ for any $t < t(\varepsilon)$. Table 3 shows the first periods defined in this way for all sessions of the **Sink** and **SaddleNeg** treatments[29] for five different values of ε. Cases where the trajectory never reached the neighborhood during the experiment are indicated by the – symbol.

Table 3 suggests that the quickest convergence was in the **SaddleNeg** treatment. Indeed, for any ε, the values of the first hit times over 4 sessions are smaller for this

[28]Looking at dynamics in all sessions in Online Appendix F, we note that there is one exception, Session 4, of the **Sink** treatment. The oscillations in both numbers of that session are due to a single subject.

[29]In the four sessions of **SaddlePos** and **Source** treatments, the trajectories for the average a and b numbers *never* entered the U_ε neighborhoods for $\varepsilon \leq 20$.

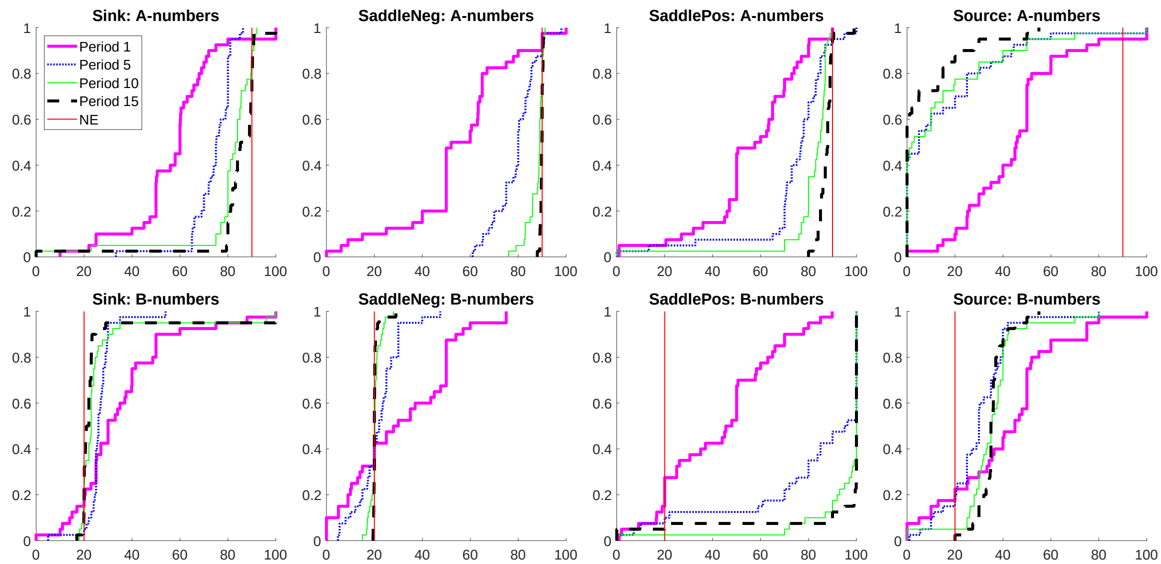


Figure 4: Cumulative frequencies of individual choices in periods 1, 5, 10, and 15 for **Sink**, **SaddleNeg**, **SaddlePos**, and **Source** treatments (from left to right) for a-number (upper panels) and b-number (lower panels). Vertical lines indicate the REE.

treatment than for the other convergent **Sink** treatment. For instance, two sessions of the **Sink** treatment did not converge to a 5-neighborhood of equilibrium in 15 periods, whereas in all four sessions of the **SaddleNeg** treatment such convergence occurred (within 8 periods, on average). Only one session of the **Sink** treatment converged to a 0.5-neighborhood while 3 of the 4 **SaddleNeg** treatment sessions achieved this convergence criterion.[30]

Finally, we illustrate the dynamics of individual guesses for all four treatments in Figure 4. Each of 8 panels shows 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 guesses are shown in the top panels and the corresponding b-number guesses are shown in the bottom panels. The steady state values are indicated by vertical (red) lines. Notice that the variance of individual choices decreases over time, with the greatest reduction

[30] As sometimes trajectories exits the ε-neighborhood, we also compared the *last* periods for which the trajectory *did not* belong to the ε-neighborhood of the equilibrium. The conclusions from this exercise are similar: the trajectories in the **Sink** treatment would stay around the equilibrium for a shorter amount of time than in the **SaddleNeg** treatment. See Appendix D.

occurring during the first 5 periods. 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.[31]

4 Behavioral Models

As we have seen, the ability of participants to converge to the REE depends on the treatment. We summarize the main features of the experimental data as follows:

1. The dynamics converge to the REE in the **Sink** and **SaddleNeg** treatments and do not converge to that equilibrium in the **SaddlePos**, and **Source** treatments. Convergence is the fastest in the **SaddleNeg** treatment.

2. The a-number converges to its REE level almost monotonically in the **Sink**, **SaddleNeg** and **SaddlePos** treatments. Convergence is quickest in **SaddleNeg** treatment and slowest in the **SaddlePos** treatment.[32]

3. The b-number converges to its REE level almost monotonically in the **Sink** treatment and via an oscillatory path in the **SaddleNeg** treatment.

4. The b-number does not converge to its REE level in **SaddlePos** and **Source** treatments. In both cases, however, the dynamics are rather stable, near some constant level.[33]

[31] Note that in the **SaddlePos** treatment many subjects submit a b-number guess of 100, but some submit a guess of 0 for the b-number.

[32] This conclusion follows from applying the "first hit time" convergence criterion to the trajectory of a-number only, see the details in Appendix D.

[33] Analysis of the boundary equilibria as reported in Table 1 reveals that in the **SaddlePos** treatment, the dynamics converge to a neighborhood of the boundary Nash equilibrium $(90, 100)$ and in the **Source** treatment the dynamics converge to a neighborhood of the boundary equilibrium $(0, 38)$. Recall that these equilibria are payoff-dominated and neither of them is the rational expectation equilibrium.

In this section we address the important question: *Can a single model of behavior explain these four features of our experimental data?* Then in the next section, we will estimate the various learning models we consider in this section and compare their fit to our data.[34]

We start with the simplest *naïve* learning model. Generalizing this model in different directions, we will discuss homogeneous level-k, adaptive, average, and cognitive hierarchy models.

4.1 Naïve Model

One of the simplest models is the so-called *naïve* model, where all participants' guesses, and thus the average guesses are set equal to the previous period's target values:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} \equiv \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} . \tag{5}$$

Using (2), we can write the naïve model as a two-dimensional dynamical system in terms of variables $\{\bar{a}, \bar{b}\}$:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + \boldsymbol{d} . \tag{6}$$

This system has a unique steady state given by the REE. The naïve model is linear and, therefore, its dynamic properties can be understood from the eigensystem of matrix $\mathbf{M}$, see Table 4. Since in all four treatments, the matrix $\mathbf{M}$ is triangular, the eigenvalues of this dynamical system, μ_1 and μ_2, are the diagonal elements of $\mathbf{M}$ that we denoted as m_{11} and m_{22}. The corresponding eigenvectors, $\mathbf{v}_1$ and $\mathbf{v}_2$, are also reported in Table 4 to provide more information about the dynamics. Note that in

[34] Thus we first focus on the *simulated paths* for each learning model over the 15 periods as in our experiment. These simulated paths may depend on initial conditions and parameters, but the learning model predictions do not use any information from the experimental data. In contrast, in econometric analysis of Section 5 we use the experimental data to "inform" the models at each time step about the most recent (aggregate) data that were available to the participants.

Treatment	**Eigen System** μ_1, $\mathbf{v}_1$	μ_2, $\mathbf{v}_2$	**Converges?** for a	for b	**Attractor in Simulations** (a, b)
Sink	$\frac{2}{3}, \begin{pmatrix}-7\\3\end{pmatrix}$	$-\frac{1}{2}, \begin{pmatrix}0\\1\end{pmatrix}$	Yes	Yes	$(90, 20)$
SaddleNeg	$\frac{2}{3}, \begin{pmatrix}-13\\3\end{pmatrix}$	$-\frac{3}{2}, \begin{pmatrix}0\\1\end{pmatrix}$	Yes	**No**	2-cycle $\{(90, 0), (90, 50)\}$
SaddlePos	$\frac{2}{3}, \begin{pmatrix}5\\3\end{pmatrix}$	$\frac{3}{2}, \begin{pmatrix}0\\1\end{pmatrix}$	Yes	No	$(90, 100)$ or $(90, 0)$ depending on init cond's
Source	$\frac{3}{2}, \begin{pmatrix}-6\\1\end{pmatrix}$	$-\frac{3}{2}, \begin{pmatrix}0\\1\end{pmatrix}$	No	No	$\{(0, 0), (0, 95)\}$ or $\{(100, 0), (100, 45)\}$

Table 4: Properties of the naïve model. This table shows the eigensystem (two eigenvalues, μ_1 and μ_2, and their corresponding eigenvectors, $\mathbf{v_1}$ and $\mathbf{v_2}$). It also shows the predictions of the model in terms of convergence for a generic initial point. The last column shows the attractor for simulations using truncation of guesses at boundary points as in (7).

all treatments, the eigenvector $\mathbf{v}_2$ is the same and is given by the vertical line.[35]

We remind the reader of some basic facts about dynamical systems. When an initial point of trajectory is exactly on the eigenvector of a linear system, the trajectory will stay on this vector forever. If μ denotes the corresponding eigenvalue, the trajectories move along the eigenvector by a factor μ per period with respect to the steady state. It follows that when $|\mu| > 1$, the trajectories starting on the eigenvector will diverge from the steady-state. Instead, when $|\mu| < 1$, the trajectories starting on the eigenvector will converge. When $0 < \mu < 1$, the trajectories converge monotonically, and when $-1 < \mu < 0$, trajectories converge by "jumping" through the steady state. Convergence is quicker when $|\mu|$ is small. Our system has two different eigenvectors, $\mathbf{v}_1$ and $\mathbf{v}_2$, that form a basis in $\mathbb{R}^2$. The dynamics of any trajectory can be most easily understood when represented in this basis. It follows that the dynamical system, initialized at any feasible point, will only converge (is globally stable) when both eigenvalues are less than 1 in absolute value. Generally, the speed of convergence is determined by the slowest dimension, i.e., by the largest (in absolute value) eigenvalue of the system.

[35] This is a simple consequence of the fact that the dynamics for the a-number are independent of the b-guesses in all of our treatments (i.e., matrix $\mathbf{M}$ is lower triangular).

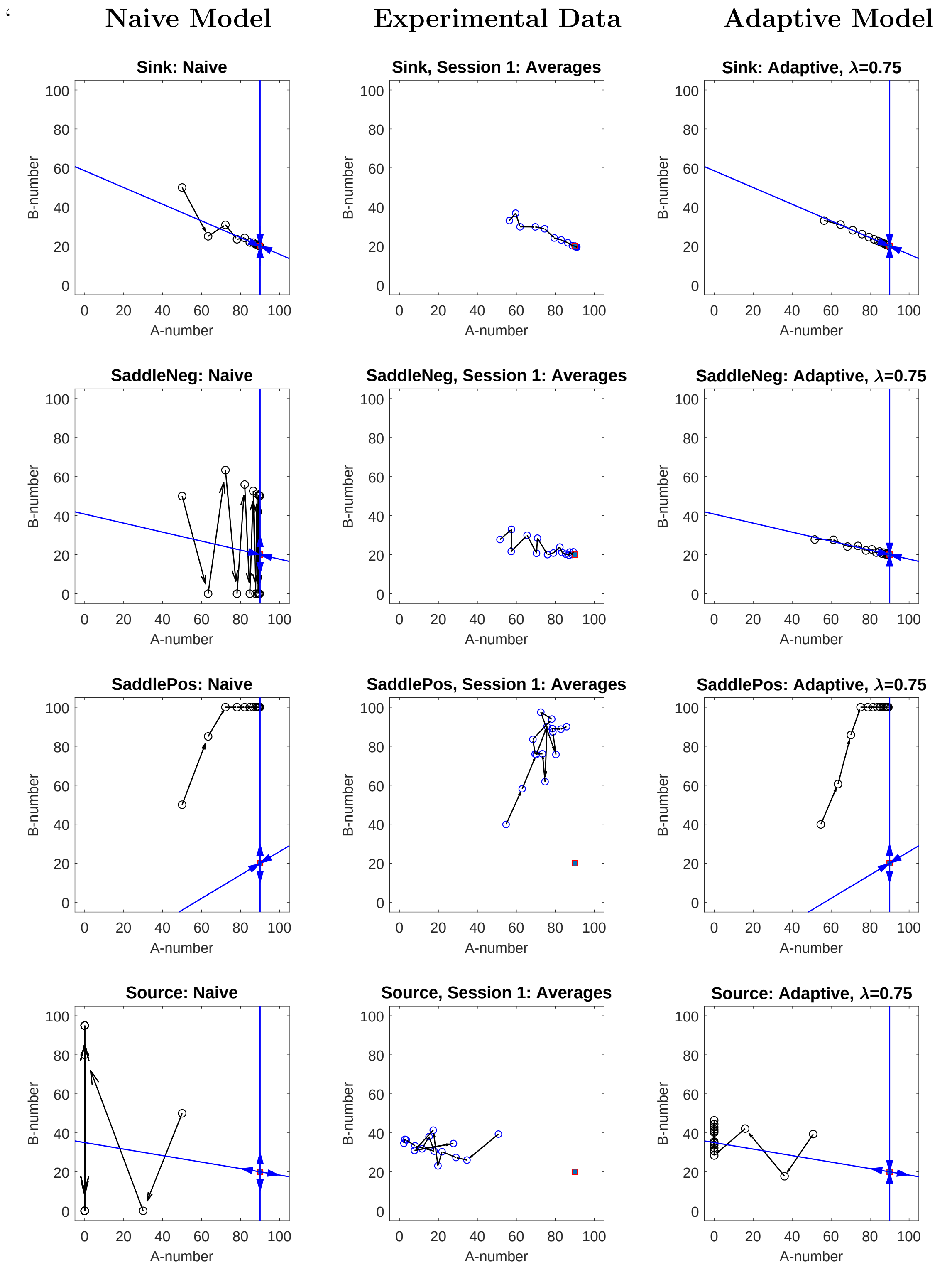


Figure 5: Dynamics of averages in an experimental group (*middle panels*) as compared with the dynamics of naïve model (6) (*left panels*) and adaptive model (11) with $\lambda = 0.75$ (*right panels*). For both learning models, the first 15 periods are simulated with the same initial point as in the experimental group. The two blue lines are the eigenvectors with the arrows indicating whether the direction is stable or not.

The dynamics of the naïve model for four treatments are illustrated in the left panels of Fig. 5. The two straight lines represent the two eigenvectors intersecting at the steady state (90,20). The arrows indicate the directions of trajectories along these vectors, whether they converge to or diverge from the steady state. The four left panels show the first 15 periods of trajectories generated by (6) for each of the four treatments. The trajectory in each treatment starts at the same initial point as in the first experimental session for this treatment.[36] To illustrate how similar dynamics of the naïve model are to the experimental dynamics, the middle panels of Fig. 5 show trajectories of average guesses in a representative experimental session (session number 1) of the same treatment.

From the second column of Table 4 we observe that in the **Sink** treatment, both eigenvalues are inside the unit circle and therefore the dynamics are globally stable. Indeed, both eigenvectors in the top left panel of Fig. 5 show convergence. Note that the trajectory of the naïve model jumps around the eigenvector $\mathbf{v}_1$; this is because the second eigenvalue is $-1/2$, i.e, negative. It is remarkable that the trajectory in the experimental session (middle panel) is also "jumping" along the dimension of b-numbers. It seems that the "naïve" model captures this dynamic feature quite well. The convergence of the model is slower in the direction of a-number. This is because the corresponding eigenvalue is $2/3 > |-1/2|$; this larger eigenvalue effectively determines the speed of convergence.

In the **SaddleNeg** and **SaddlePos** treatments, one of the eigenvalues is outside of the unit circle, see Table 4. Therefore, the steady state is locally unstable. However, since another eigenvalue is inside the unit circle, there is one stable direction. This direction corresponds to the dynamics for the a-number. Thus, a-numbers converge in the model, but b-numbers do not converge generically, i.e., unless the initial point is *exactly* on the stable eigenvector $\mathbf{v}_1$. The left panels of the second and third rows of Fig. 5 illustrate the dynamics of the naïve model in the two saddle treatments. When

[36] In the Online Appendix G we illustrate the dynamics of the naïve model in all four treatments for the initial condition $(50, 50)$, i.e., assuming that all subjects are level-0 in period 1.

simulating this model (as well as any other model later on) we take into account the fact that in the experiment the submitted guesses could not be outside of the interval $[0, 100]$. Therefore, in our simulations we modify (5) and simulate, instead, the following dynamics

$$\bar{a}_t = \max\left\{0, \min\left\{100, a^*_{t-1}\right\}\right\} \quad \text{and} \quad \bar{b}_t = \max\left\{0, \min\left\{100, b^*_{t-1}\right\}\right\}. \qquad (7)$$

As a result of this modification, the long-term dynamics of the naïve learning model change when they diverge, as happens in both of the saddle treatments. In the **SaddleNeg** treatment, the dynamic path for the b-number hits 0 at some point and eventually converges to a 2-cycle, cycling between 0 and 50.[37] In the **SaddlePos** treatment, the dynamics for the b-number under the naïve learning model (for the initial conditions in Fig. 5) eventually hits 100 and stays there so that the trajectory converges to $(90, 100)$. We report the attractors under simulations in the last column of Table 4. Comparing the simulated dynamics for the two saddlepath treatments with the dynamics in the experiment as in the middle panels, we conclude that in the **SaddleNeg** treatment the naïve learning model generates diverging dynamics for the b-number in contrast to the experiment. While in the **SaddlePos** treatment the naïve model captures convergence of a-number and divergence of b-number consistently with the experiment, the speed at which b-numbers diverge is larger in the model than in the experiment.[38]

Finally, in the **Source** treatment both eigenvalues are outside of the unit circle and the dynamics diverge away from the REE except in the special case where the system starts exactly at that equilibrium, (a^E, b^E). Again, we illustrate the dynamics in the left panel of Fig. 5, in the bottom row. While in both the experimental data

[37] When the a-number converges to 90, a zero b-guess leads to $-90/2 + 95 = 50$ as the target for the b-number. When 50 is submitted as a new guess (according to the naïve model), the target b-number value becomes negative, resulting in a choice of 0 after truncation.

[38] Moreover, when participants' b-guesses hit the upper bound of 100 implying much larger and unfeasible targets, participants' earnings are very low and they behave much more randomly than our truncated specification (7) assumes.

and the naïve model simulations, the dynamics diverge from the REE, there is a large discrepancy between the experimental data and the predictions of the naïve model; in the latter, the a-number converges to 0 while the b-number eventually oscillates between 0 and 95. By contrast, the experimental data converges to a boundary Nash equilibrium of (0,38) with no oscillation in the b-number.

We summarize this discussion as follows.

Result 1. *The naïve model cannot reproduce all four features of our experimental data. In particular, it does not predict convergence to the REE in the* ***SaddleNeg*** *treatment and there are discrepancies in other treatments as well, including very different dynamics in the* ***Source*** *treatment.*

4.2 Homogeneous level-k learning model

In Section 3.1 we analyzed first period choices in light of the level-k model. A simple generalization of this model can be made for dynamic choices. We will follow Nagel (1995) and define level-0 in period $t > 1$ as guessing the average of the a and b-numbers from the previous period,[39] i.e.,

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}. \tag{8}$$

Agents who follow the level-0 choice in each time period can also be labeled as "stubborn" agents, since their guesses do not change throughout all rounds of the experiment. For example, if all agents are level-0, i.e., if we employ a *homogeneous*-level 0 model, and in the first period the average guess is 50 for both numbers, then all

[39] One 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. The individual data suggests that both pieces of information were used. The definition of level-0 as guessing the past average, 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.

agents would continue to guess 50 for both numbers in all subsequent periods of the experiment. Obviously, such a homogeneous level-0 model generates trajectories that are very different from those that we observed in our experiment.

Following the same logic as for the choices in the first period, 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 would submit guesses that are equal to the previous period's targets. The dynamical model where all agents are of level 1, i.e., the homogeneous level-1 model, thus coincides precisely with the naïve model (5) analyzed in the previous section. It follows that the homogeneous level-1 dynamical model is not able to explain the convergence we observed in the **SaddleNeg** treatment and yields other discrepancies with our experimental data, as in the dynamics for the b-number in the **Source** treatment.

Proceeding further, we define a homogeneous level-2 model for time t by guesses that best respond to the level-1 choice at time t, i.e.,

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} + \boldsymbol{d}\,. \tag{9}$$

More generally, the level-k choice at time t will be to choose numbers that best respond to the choice of level $k-1$ at time t.

The homogeneous level-k learning model assumes that all agents are of level k and behave in the same way, making a level-k choice for some k at every period in time. Thus for any $k > 1$ the level-k model simply iterates the level-1 (or naïve) model. Therefore, the dynamics of the level-k model are governed by a linear system with matrix $\mathbf{M}^k$. It follows that the homogeneous level-k model will have exactly the same conditions for convergence, in terms of the eigenvalues, as the homogeneous level-1 (i.e., the naïve) model,[40] see the level-2 model in Table 6 and compare the

[40] If μ is an eigenvalue of $\mathbf{M}$ with eigenvector $\boldsymbol{v}$, then μ^k is an eigenvalue of $\mathbf{M}^k$ with the same eigenvector. Moreover $|\mu| < 1$ if and only if $|\mu^k| < 1$.

convergence predictions for that model with those for the naïve model in Table 4.

To summarize this discussion, we conclude that neither the level-0 model nor the homogeneous level-k model with any $k \geq 1$ can explain the observed convergence in the **SaddleNeg** treatment. Based on this analysis we have:

Result 2. *The homogeneous level-k learning model for any finite k cannot reproduce all four features of our experimental data.*

Note that our results hold only if we restrict ourselves to the model with a *homogeneous* level of iterative thinking. Later, in section 4.4, we will investigate mixed cognitive level models, where agents with different levels of rationality coexist, and we will find that, in such cases, this negative finding of Result 2 can be reversed.

4.3 Adaptive Learning Models

The adaptive learning model that we consider next, generalizes the naïve model by assuming that subjects change their guesses in response to past errors.[41] According to this model, players' choices are adapted towards the targets in a constant proportion to the forecast error. This model implies the following evolution of average choices:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + \lambda \left[\begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} - \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} \right] = \lambda \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} + (1-\lambda) \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}, \quad (10)$$

with $\lambda \in (0, 1]$. The naïve model is a special case of the adaptive learning model when $\lambda = 1$.

The second part of expression (10) suggests the following interpretation of the adaptive model. Assume that the population consists of individuals with different

[41]The adaptive model is popular in macroeconomics where it is used to model expectations and is known as *adaptive expectations*; 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).

levels of rationality. If fraction $1-\lambda$ of the population uses the level-0 model while the remaining fraction λ of the population best responds to those using the level-0 model and thus behave according to the level-1 model, then the dynamics in the whole population will be described by equation (10). This interpretation provides a novel and useful bridge between the homogeneous level-k models that we introduced in Section 4.2 and the model of adaptive expectations that is frequently used in the macroeconomics literature. We formulate this connection as the following result.

Result 3. *Let level-0 agents be stubborn as in (8), and suppose level-1 agents best respond to level-0 agents. Then the adaptive model is equivalent to this mixed cognitive level model with fixed proportions $1-\lambda$ of level-0 agents and λ of level-1 agents in the population.*

More general mixed models, where agents have various levels of rationality and those with higher levels best respond to the behavior of agents with lower levels, are discussed further in the next section. Here we will continue with the analysis of the adaptive model. Substituting the target values from (2) into equation (10) we obtain a linear dynamical system for the averages:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = (\lambda \mathbf{M} + (1-\lambda)\mathbf{I}) \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + \lambda \boldsymbol{d}\,. \tag{11}$$

As in the case of the naïve and homogeneous level-k models, the only steady state of the dynamical system characterized by adaptive learning is the REE $(a^E, b^E) = (\mathbf{I} - \mathbf{M})^{-1}\boldsymbol{d}$. However, the linear dynamics are now governed by the matrix $\lambda\mathbf{M} + (1-\lambda)\mathbf{I}$. Given that, in all of our treatments, the matrix $\mathbf{M}$ is lower triangular, the eigenvalues of this matrix are easy to determine. They are:

$$\mu_1 = 1-\lambda+\lambda m_{11} \quad \text{and} \quad \mu_2 = 1-\lambda+\lambda m_{22}\,. \tag{12}$$

The eigenvalues of the adaptive model are thus the weighted averages of 1 (with

Treatment	Eigen System μ_1, $\mathbf{v}_1$	Eigen System μ_2, $\mathbf{v}_2$	Converges to a^E and b^E? for a	Converges to a^E and b^E? for b
Sink	$1-\lambda/3$, $\begin{pmatrix}-7\\3\end{pmatrix}$	$1-3\lambda/2$, $\begin{pmatrix}0\\1\end{pmatrix}$	Always monotone	Always; jumps when $\lambda > 2/3$
SaddleNeg	$1-\lambda/3$, $\begin{pmatrix}-13\\3\end{pmatrix}$	$1-5\lambda/2$, $\begin{pmatrix}0\\1\end{pmatrix}$	Always monotone	Only for $\lambda < 0.8$; jumps when $\lambda > 0.4$
SaddlePos	$1-\lambda/3$, $\begin{pmatrix}5\\3\end{pmatrix}$	$1+\lambda/2$, $\begin{pmatrix}0\\1\end{pmatrix}$	Always monotone	Never, diverges to 0 or 100
Source	$1+\lambda/2$, $\begin{pmatrix}-6\\1\end{pmatrix}$	$1-5\lambda/2$, $\begin{pmatrix}0\\1\end{pmatrix}$	Never, diverges to 0 or 100	Never, for $\lambda < 0.8$ converges to $\mathbf{v}_1$

Table 5: Properties of the adaptive model. This table shows the eigensystem (two eigenvalues, μ_1 and μ_2, and their corresponding eigenvectors, $\mathbf{v_1}$ and $\mathbf{v_2}$). It also shows the predictions of the adaptive model in terms of convergence for a generic initial point.

weight $1-\lambda$) and the eigenvalues of matrix $\mathbf{M}$, i.e., of the naïve model (with weight λ). Moreover, matrix $\lambda\mathbf{M}+(1-\lambda)\mathbf{I}$ has the same system of eigenvectors as the matrix $\mathbf{M}$.[42] This allows us to characterize the convergence properties of the adaptive model in different treatments for different values of λ, see Table 5.

We find that to obtain convergence to the REE of the **SaddleNeg** treatment under the adaptive learning model, the weight, λ, that is assigned to the previous target by the adaptive model cannot be too large. Specifically, when $\lambda < 0.8$, the dynamics for the b-number will converge. Moreover, the "jumping" behavior along the b-dimension, which is visible in the phase diagrams of the experimental data from the **Sink** and **SaddleNeg** treatments, and which translates to non-monotone convergence in the b-number, is also consistent with the adaptive learning model, but only when λ is not too small. In fact, when $\lambda > 2/3$ this feature of the experimental data is reproduced in both the **Sink** and the **SaddleNeg** treatments. The adaptive model is therefore capable of capturing the major dynamic patterns observed in all four of our experimental treatments. For instance, the median numbers of the final experimental period are the same as those predicted by the adaptive model for the

[42] Assume that $\mathbf{v}$ is the eigenvector of $\lambda\mathbf{M}+(1-\lambda)\mathbf{I}$, associated with the eigenvalue μ_1. Then $(\lambda\mathbf{M}+(1-\lambda)\mathbf{I})\mathbf{v} = \mu_1\mathbf{v} = (1-\lambda+\lambda m_{11})\mathbf{v}$. But then $\lambda\mathbf{M}\mathbf{v} = \lambda m_{11}\mathbf{v}$ and so $\mathbf{M}\mathbf{v} = m_{11}\mathbf{v}$, meaning that $\mathbf{v}$ is the eigenvector of the matrix $\mathbf{M}$ that is associated with the eigenvalue m_{11}.

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 the adaptive model, which is very close to the average value of 35 observed in the experiment.

We thus arrive at the following conclusion

Result 4. *The adaptive model with weight* $\lambda \in (\frac{2}{3}, \frac{4}{5})$ *can reproduce all four features of the experimental data.*

The right panels in Fig. 5 show the trajectories of the adaptive model where $\lambda = 0.75$, which belongs to the interval identified in Result 4.[43] These adaptive model dynamics can be compared with the naïve model (left panels) and the dynamics from the experimental data (middle panels). As in the case of the naïve model, we initialized the dynamics for the adaptive model (11), so that the first observation corresponds to that of the first experimental group (session 1) of the corresponding treatment (shown in the middle panel). The adaptive model reproduces the main features of the experimental dynamics much better than the naïve model. Most importantly, the dynamics of the adaptive 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 (7) 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 the simulated path does not use the experimental data, and thus does not account for discrepancies between the model's prediction and the noisier experimental dynamics at every time step.

[43] Online Appendix G provides a more general discussion of the adaptive learning model and illustrates the dynamics of that model using other values for λ.

By rewriting the adaptive model (10) recursively, we obtain for any $t \geq 2$:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \lambda \begin{pmatrix} a^*_{t-1} + (1-\lambda)a^*_{t-2} + \cdots + (1-\lambda)^{t-2}a^*_1 \\ b^*_{t-1} + (1-\lambda)b^*_{t-2} + \cdots + (1-\lambda)^{t-2}b^*_1 \end{pmatrix} + (1-\lambda)^{t-1} \begin{pmatrix} \bar{a}_1 \\ \bar{b}_1 \end{pmatrix}. \quad (13)$$

Thus, the adaptive model effectively takes into account all past target values. These values were available to the participants in our experiment, and this might be an important reason why Result 4 holds. Intuitively, in the **SaddleNeg** treatment, initially the dynamics for the b-number are oscillating around the Nash equilibrium. If agents use past targets for the b-number, these oscillations will be dampened and convergence will be obtained. By contrast, for the **SaddlePos** treatment, the initial dynamics for the b-number moves in a positive direction away from the Nash equilibrium so that past averaging is not very useful.

Note that in (13), past target values have exponentially declining weights. For this reason, the model (13) is often called the exponentially weighted moving average (EWMA) model. An alternative model averages the previous L target values with equal weights. We refer to this averaging model as the moving average model, MAve(L).[44] When $t > L$, the average guesses given by this model are:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \frac{1}{L} \begin{pmatrix} a^*_{t-1} + \cdots + a^*_{t-L} \\ b^*_{t-1} + \cdots + b^*_{t-L} \end{pmatrix}. \quad (14)$$

This model requires L initial target values to initialize. For instance, when $L = 2$, the model takes the target values of the first two periods, and then, starting from $t = 3$, it assumes that agents guess the averages of the two previous target values of

[44] In the so-called "econometric learning" approach in macroeconomics, as advocated by Evans and Honkapohja (2001) this model is presented as a less restrictive alternative to Rational Expectations, with agents learning parameters of their perceived model by means of statistical inference from past observations. Both the EWMA and the MAve(L) models belong to this literature. The EWMA model is known as a *constant* gain model because the weight attached to the latest available observation is the same for every time t. The MAve(L) model assigns smaller and smaller weights to the most recent observation, see Equation (14), and so this model is known as the model with *decreasing* gain, which approximates recursive least squares learning.

the corresponding variable.[45]

We study the properties of this average model and illustrate its dynamics in Online Appendix H. On the basis of that analysis we conclude

Result 5. *The convergence properties of MAve(L) model, with $L \geq 2$, converges (doe not converge) to the Nash equilibrium when the experimental data does (does not), but it is not fully consistent with all features of our experimental data. In particular, with larger L convergence becomes much slower in the model than in the experiment.*

4.4 Mixed Cognitive Models

The adaptive model suggests that the aggregate choices can be better explained by models that combine various cognitive levels, e.g., levels 0 and 1. The dynamic version of such a "mixed cognitive model" starts by assuming that there is a distribution of the levels of rationality in the population. Some agents, whose proportion is denoted by f_0, use level-0 thinking by submitting as their guess the average number from the previous period. Other agents, in proportion f_1, assume that the population consists only of level-0 agents and play a best reply to the choice of those agents, i.e., they guess the previous period's target values. In general, level-k agents whose proportion is f_k, assume that all other agents are of level $k-1$ and best respond to them.[46]

Given an exogenous distribution $\mathcal{F} = \{f_k\}_{k=0}^{K}$, where K is the largest level in population, we can write the dynamics generated by such a model (in deviations

[45] When $L = 1$, this model is identical to the naïve model.

[46] In the literature such mixed cognitive models are sometimes referred as the "level-k" model to be distinguished from the cognitive hierarchy model of Camerer et al. (2004), Chong et al. (2016), where agents best respond to the mixed population of lower levels of rationality and which we will discuss later on. For this reason, we emphasized the assumption of homogeneity in level types in our prior discussion of homogeneous level-k behavior and in Result 2 earlier on.

Treatment	Level-2				Mixed model of level-1 (naïve) and level-2			
	EVs		**Converges?**		**EVs**		**Converges to a^E and b^E?**	
	μ_1	μ_2	for a	for b	μ_1	μ_2	for a	for b
Sink	$\frac{4}{9}$	$\frac{1}{4}$	Yes	Yes	$f_1\frac{2}{3}+f_2\frac{4}{9}$	$-f_1\frac{1}{2}+f_2\frac{1}{4}$	Always monotone	Always; jumps for $f_1 > 1/3$
SaddleNeg	$\frac{4}{9}$	$\frac{9}{4}$	Yes	**No**	$f_1\frac{2}{3}+f_2\frac{4}{9}$	$-f_1\frac{3}{2}+f_2\frac{9}{4}$	Always monotone	for $1/3 < f_1 < 13/15$; jumps for $f_1 < 3/5$
SaddlePos	$\frac{4}{9}$	$\frac{9}{4}$	Yes	No	$f_1\frac{2}{3}+f_2\frac{4}{9}$	$f_1\frac{3}{2}+f_2\frac{9}{4}$	Always monotone	Never; goes to 0 or 100
Source	$\frac{9}{4}$	$\frac{9}{4}$	No	No	$f_1\frac{3}{2}+f_2\frac{9}{4}$	$-f_1\frac{3}{2}+f_2\frac{9}{4}$	Never, goes to 0 or 100	Never; for $1/3 < f_1 < 13/15$ goes to $\mathbf{v}_1$

Table 6: Properties of the level-2 and mixed models.

from the Nash equilibrium) as follows

$$\begin{pmatrix}\bar{a}_t - a^E \\ \bar{b}_t - b^E\end{pmatrix} = f_0\begin{pmatrix}\bar{a}_{t-1} - a^E \\ \bar{b}_{t-1} - b^E\end{pmatrix} + f_1\begin{pmatrix}a^*_{t-1} - a^E \\ b^*_{t-1} - b^E\end{pmatrix} + f_2\mathbf{M}\begin{pmatrix}a^*_{t-1} - a^E \\ b^*_{t-1} - b^E\end{pmatrix} + \cdots =$$
$$= \left(f_0\mathbf{I} + f_1\mathbf{M} + f_2\mathbf{M}^2 + \cdots + f_K\mathbf{M}^K\right)\begin{pmatrix}\bar{a}_{t-1} - a^E \\ \bar{b}_{t-1} - b^E\end{pmatrix}.$$

where $\sum_k f_k = 1$ and $f_k \in [0, 1)$. As we already recognized in Result 3, when $K = 1$, i.e., if there are only two levels, 0 and 1, this model is identical to the adaptive model.

Result 4 then suggests that mixed cognitive models may describe the experimental data better than any homogeneous cognitive model. To illustrate this point further, we consider another mixed cognitive model. In this model, only level-1 and level-2 agents are present in proportions f_1 and $f_2 = 1 - f_1$.[47] The matrix of the corresponding system is $f_1\mathbf{M} + f_2\mathbf{M}^2$. Its eigenvalues are the convex combinations of the eigenvalues of matrix $\mathbf{M}$ and its second power. We summarize the properties of this model in the last column of Table 6 and observe that for $1/3 < f_1 < 3/5$ the dynamics are similar to the experimental data.

Result 6. *Both the mixed cognitive model with a population consisting of level-0 and*

[47]If our level-0 seems too simplistic, this model can be more reasonable as a mixing model of low levels, see footnote 39. However, when our data are used to compare models quantitatively, this mixing model is not found to be better than the adaptive learning model, as shown in the next section.

level-1 agents and the mixed cognitive model with a population consisting of level-1 and level-2 agents can reproduce all features of our experimental data.

It then remains an empirical question as to which of these mixed cognitive models describe the data best. This question will be addressed later, in Section 5.

4.4.1 Cognitive Hierarchy Model

Camerer et al. (2004) introduced a variant of the level-k model where agents of level k do not simply best respond to agents of level $k-1$, but rather to a distribution over all lower levels, which they call the Cognitive Hierarchy (CH) model.[48] In this model, agents are also distributed according to $\mathcal{F}$, but the agents of every level-k are more sophisticated than in the model we have discussed so far, as they play a best response to the rest of the population assuming that *all* other agents' levels are lower than theirs and that those lower-level agents are present in the proportions given by the normalized distribution $\mathcal{F}$. Thus, every level-k agent first builds a perceived distribution of the levels of the other agents in the population, $\{g_\ell\}_{\ell=0}^{k-1}$, with

$$g_\ell = \frac{f_\ell}{f_0 + f_1 + \cdots + f_{k-1}} \,,$$

where g_ℓ is the fraction of level-ℓ agents perceived by the level-k agent. Then, the agent submits as his/her guess, a best response to the perceived behavior of the other agents which, in deviation from the Nash equilibrium, is given by

$$\mathbf{M}\big(g_0\mathbf{I} + g_1\mathbf{M} + \cdots + g_{k-1}\mathbf{M}^{k-1}\big) \begin{pmatrix} \bar{a}_{t-1} - a^E \\ \bar{b}_{t-1} - b^E \end{pmatrix} .$$

This guess should be 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

[48] Chong et al. (2016) provide a generalization of their original cognitive hierarchy model, directly connecting it with the level-k model.

levels of rationality that exist in the population (i.e., from 0 up to $K \leq \infty$), we obtain the dynamics of the CH model.

We note that if $K = 1$, i.e., if there are only two levels, 0 and 1, the CH model is identical to the mixed model with levels 0 and 1, and, hence, to the adaptive model. For larger values of K, i.e., for more levels of rationality, the CH model may seem quite cumbersome. The next result shows that the CH model can be viewed as the application of several adaptive models.

Proposition 4.1. *The dynamic version of the CH model with distribution $\mathcal{F}$ is described by the following linear system*

$$\begin{pmatrix} \bar{a}_t - a^E \\ \bar{b}_t - b^E \end{pmatrix} = \prod_{i=1}^{K} (\lambda_i \mathbf{M} + (1 - \lambda_i)\mathbf{I}) \begin{pmatrix} \bar{a}_{t-1} - a^E \\ \bar{b}_{t-1} - a^E \end{pmatrix} , \tag{15}$$

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

Proof. See Appendix E. □

Proposition 4.1 implies that the convergence properties of the dynamic CH model depend on the product of matrices from different *adaptive* models. The coefficients of the adaptive models are given by the perceived relative weight of the *largest* level of rationality below the agent's level with the remaining weight given to the lower levels. Since in all of our treatments, the matrices are lower triangular, their product is also a lower triangular matrix. The diagonal elements of these matrices are thus the eigenvalues of the dynamic CH model (15), and they are equal to the products of the eigenvalues of the corresponding adaptive models. Thus, the CH model converges if all of the adaptive models combined on the right-hand side of (15) converge.[49]

Camerer et al. (2004) focus on the special case where $\mathcal{F}$ is a Poisson distribution. In this Poisson CH model, the proportion of level-k types in the population is given

[49]On the other hand, even if some of the adaptive models do not converge, the CH model may still converge.

by $e^{-\tau}\frac{\tau^k}{k!}$. The only parameter of this distribution, $\tau > 0$, characterizes the average level of rationality in the population.[50] In the special case where the Poisson CH model is truncated to two levels, 0 and 1, the proportion of level-1 types is given by $\tau/(1+\tau)$. Applying Proposition 4.1 to the case of the Poisson distribution we can establish a one-to-one correspondence between parameter τ and the weight in the adaptive model.

Corollary 4.1. *The adaptive model* (10) *corresponds to the Poisson CH model from Camerer et al. (2004) with two levels of rationality,* 0 *and* 1*, and* $\tau = \lambda/(1-\lambda)$.

5 Estimation

In this section, we compare the models discussed in Section 4 in terms of their quantitative fit to the experimental data. For a given session and period, each learning model $\mathcal{L}$ generates a prediction by mapping the experimental data available to the participants in that session and period t (i.e., all information through period $t-1$, including previous targets and the averages of both numbers) to the average guesses for period t:[51]

$$\begin{pmatrix} \bar{a}_t^m \\ \bar{b}_t^m \end{pmatrix} = \mathcal{L}\left[\begin{pmatrix} a_{t-1}^* \\ b_{t-1}^* \end{pmatrix}, \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}, \begin{pmatrix} a_{t-2}^* \\ b_{t-2}^* \end{pmatrix}, \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}, \ldots; \boldsymbol{\theta}\right]. \tag{16}$$

Some models contain parameters, which we denote by $\boldsymbol{\theta}$, and so we also estimate these parameters. Thus, each learning model gives one-period ahead predictions, $\bar{a}_t^m$ and $\bar{b}_t^m$ for the corresponding session at time t. To avoid cumbersome notation, we

[50] If there is some maximum level of rationality, K, as we assume, then the Poisson distribution is truncated and normalized accordingly.

[51] By using the average guesses instead of pooled estimation of all the individual guesses, we mitigate a potential heterogeneity bias highlighted by Wilcox (2006). Still, there might be some heterogeneity across treatments. We perform robustness checks in Online Appendix I, where we estimate the models separately for each treatment. There we see that the performance ranking of the models remains virtually unchanged.

do not add a session-specific index to the values in the right-hand side of (16).[52]

We estimate parameters where applicable, and compare the following models:

- the REE prediction;
- the naïve model (5);
- the average models, including the moving average models (14) with 2 to 5 lags;[53] and the exponential weighted moving average (EWMA) model (13).[54]
- homogeneous level-k models with $k = 0$ (stubborn), $k = 1$ (naïve), and $k = 2$;
- two versions of adaptive model: the standard model as in (10), and its generalization with different parameters, λ_a and λ_b, for a and b numbers (see Online Appendix G);
- mixed cognitive models with various combinations of level-k models;
- the cognitive hierarchy (CH) model with a Poisson distribution over levels up to $k = 2$, $k = 3$, and $k = 4$.

We estimate the parameters of the models containing parameters by minimizing the sum of squared errors (SSE) between the model predictions and the actual data from the experiment.[55] Since estimating the adaptive model requires one lag, for

[52]The presence of the experimental data in the right-hand side of (16) distinguishes this *one-period ahead* prediction from the simulated paths analyzed in the previous section, which does not make use of the experimental data.

[53]For these models, if the window over which the average is to be computed is larger than the available data at a given time period, then we compute the average over a shorter window of all available observations at this time period.

[54]We stress that although the EWMA model (13) is equivalent analytically to the adaptive model (10), and the models generate the same simulated path, we estimate them differently. The parameter of the adaptive model is estimated as a linear regression of the current period averages on the averages and guesses from the last period, whereas the parameter of the EWMA model is estimated by minimizing the SSE between the actual averages and the averages generated using all past targets as in (13). The initial $\begin{pmatrix} \bar{a}_1 \\ \bar{b}_1 \end{pmatrix}$ on the right-hand side of this equation is set to $\begin{pmatrix} a_1^* \\ b_1^* \end{pmatrix}$.

[55]This procedure is equivalent to the maximum likelihood estimation under the assumption of normal residuals.

comparison purposes we estimate all models starting from period $t = 2$. The SSE for a given session is thus defined as

$$\sum_{t=2}^{15} \left((\bar{a}_t^m - \bar{a}_t)^2 + (\bar{b}_t^m - \bar{b}_t)^2 \right).$$

We seek universal parameters for all treatments and sessions. For this reason, in estimating the model's parameters $\boldsymbol{\theta}$, we minimize the sum of the SSEs over all treatments and sessions. Parameter estimates resulting from this estimation for different models are shown in column 2 of Table 7. For all parameter estimates, we also report the associated standard errors (in parentheses).

We compute standard errors using a nonparametric bootstrap method from 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. The number of bootstrap samples is 1000. We compute parameter estimates for each of the bootstrap samples and then take their standard deviation as standard errors.

After estimating each model's parameters, we can finally compare the performance of these models using their out-of-sample one-step-ahead prediction mean squared error (MSE). We compute the MSEs for every treatment separately and report the results in the next four columns of Table 7. In the last column of the table, we report the MSEs *averaged* over all treatments, i.e., over all 16 sessions. To compute the out-of-sample MSEs for the models with parameters we use a variant of cross-validation (Stone, 1974). In particular, we first re-estimate the model parameters with one of the sessions *left out*. Then, we compute the out-of-sample one-step-ahead MSE of the estimated model on the data of the left-out session.[56] We performed this procedure

[56]The procedure is called "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

Models		Parameter Estimates	Out-of-sample MSEs Sink	SadNeg	SadPos	Source	Overall
REE		-	221.16	97.53	4559.81	5685.28	2640.94
Naïve		-	44.03	25.00	212.42	318.09	149.88
Average	MAve(2)	-	35.76	23.15	199.29	269.72	131.98
	MAve(3)	-	36.57	24.95	187.91	263.17	128.15
	MAve(4)	-	36.30	29.37	183.30	252.84	125.45
	MAve(5)	-	39.02	34.04	182.05	255.46	127.64
	EWMA	0.42 (0.06)	37.07	29.77	185.18	250.64	125.66
Level-k	0 (Stubborn)	-	49.02	59.20	138.00	210.57	114.20
	1 (Naïve)	-	44.03	25.00	212.42	318.09	149.88
	2	-	89.36	153.34	397.90	683.08	330.92
Adaptive	$\lambda_a = \lambda_b$	0.44 (0.03)	**32.33**	**13.42**	84.68	**114.58**	**61.25**
	$\lambda_a \neq \lambda_b$	0.38 (0.03) 0.47 (0.04)	32.48	13.06	87.86	112.34	61.44
Mixed	0 1 (Ada)	0.56 (0.03) 0.44(0.03)	**32.33**	**13.42**	84.68	**114.58**	**61.25**
	0 1 2	0.56 (0.03) 0.44 (0.03) 0.00 (0.00)	32.33	13.42	84.68	114.58	61.25
	1 2	0.84 (0.03) 0.16 (0.03)	48.73	15.68	232.68	279.69	144.19
CH-Poisson	$\max k = 1$ (Ada)	0.78 (0.08)	**32.33**	**13.42**	84.68	**114.58**	**61.25**
	$\max k = 2$	0.58 (0.05)	32.50	15.79	83.41	116.76	62.11
	$\max k = 3$	0.56 (0.05)	32.53	16.06	**83.23**	117.00	62.21

Table 7: Estimation and performance of various learning models in terms of the out-of-sample, one-step-ahead MSE using the leave-one-out procedure, which is appropriate to compare models with a different number of parameters. MSEs are computed for periods 2 to 15. The smallest MSEs for each treatment and overall are emboldened. Standard errors for the parameter estimates are in parentheses.

for each session of each of our four treatments. That is, for each treatment, each session is left out once and we compute MSEs for each of the four left-out sessions. Finally, we average these MSEs over the four sessions of each treatment and then over all 16 sessions to compute the overall MSE. Consistent with the estimation, we evaluate MSEs starting from period 2 for all models.

Comparing the MSEs for all considered models, we find that the REE prediction is the least accurate among all models. The naïve (Level-1) model shows much improved performance, especially for the converging **Sink** and **SaddleNeg** treatments. The

penalty for the number of parameters, i.e., over-parametrized models perform better "in-sample", but show worse "out-of-sample" performance 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 and under certain conditions, it is asymptotically equivalent to AIC (Stone, 1977) and the obtained performance ranking is equivalent to the "out-of-sample" model likelihood ranking.

moving average models generally perform better than the naïve model. In particular, the model with a 4-lag moving window is the best in the class of average models, narrowly beating the EWMA model. However, the homogeneous level-0 (stubborn) model performs better in terms of the overall MSE than any average model. This finding is predominantly due to the better performance of the stubborn, level-0 model in the case of the non-convergent treatments.

We then observe that the adaptive models perform substantially better than all models considered in Table 7. The standard adaptive learning model with one parameter ($\lambda_a = \lambda_b = \lambda$), which is equivalent to the mixed cognitive model with levels 0 and 1 (Result 3) and the Poisson CH model with two levels of rationality (Corollary 4.1), performs better than the adaptive model with the two separate parameters for the a and b numbers. Note that the parameter estimates for the single-parameter adaptive model are close to the parameter estimates of the EWMA model, but the fit of the EWMA model (in terms of MSE) is comparatively worse because the EWMA model does not consider the actual realizations of the past averages when making predictions. The CH-Poisson class of models shows reasonably good performance, but its overall performance worsens with increasing maximum k.

To summarise, a simple, single-parameter adaptive model wins the out-of-sample MSE competition further supporting the conclusions of the theoretical analysis of the previous section. The best fitting model is a mixture of 66% level-0, stubborn agents and 44% level-1 agents who best respond to the stubborn agents.[57]

Finally, it is interesting to investigate how parameters of the best performing model evolve. For this purpose we estimate the weights of the mixed model with levels 0 and 1 for all periods 1–15, for period 1 only and then for periods 2–15. Table 8 shows the corresponding parameter estimates and their standard errors (in parentheses).

[57]The estimated parameter of the adaptive model, $\widehat{\lambda} = 0.44$, is outside of the range $\left(\frac{2}{3}, \frac{4}{5}\right)$ derived in Result 4. The deterministic path for this model with $\lambda = \widehat{\lambda}$ converges in the **SaddleNeg** treatment, but does it without "jumping" along the b-dimension. The econometric model, however, allows for the exogenous noise and hence can generate fluctuations even with such low value for λ.

Level	**0**	**1**
Periods 1–15	0.63 (0.02)	0.37 (0.02)
Period 1	0.76 (0.04)	0.24 (0.04)
Periods 2–15	0.56 (0.03)	0.44 (0.03)

Table 8: Parameter estimates of the mixed model with levels 0 and 1 (equivalent to the adaptive model) for all 15 periods and separately for period 1 and periods 2–15. Standard errors are in parentheses. All parameters and differences in parameters between period 1 and periods 2–15 are significant at 1% level.

The standard errors are estimated by the bootstrap method previously described. All parameter estimates are significant at 1% level. Moreover, the difference in these parameter estimates between period 1 and periods 2–15 is also significant at 1% level. We observe that level-0 has the highest weight overall and in all sub-periods. However, the weight attached to level-0 reduces significantly after period 1 suggesting some increase in the level of rationality.

6 Conclusion

We have designed and reported on an experiment examining learning behavior in linear, coupled, multivariate systems, which are widely used in 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, planar beauty contest game. For simplicity, we have focused on the simplest possible structure for the PBC such that one variable, a, is decoupled, but the other variable, b, depends on the average predictions of both a and b.

We conducted four treatments with various eigenvalues for the planar system. Convergence is achieved in the **Sink** treatment and the saddle case but only when there is negative feedback (the sign of the unstable eigenvalue is negative), in the so-called **SaddleNeg** treatment. Rather remarkably, participants are able to learn and follow the saddlepath in this case without initially being placed on the saddle path itself. Saddlepath dynamics often appear in theoretical macroeconomic models and it is typically assumed the agents operate on the saddle path. Our results suggest that for certain parameterizations of such models, this assumption may be reasonable, whereas, for other parameterizations of such models, it is not. Indeed, we observe that in the **SaddlePos** treatment (where the sign of the unstable eigenvalue is positive) and in the **Source** treatment, the dynamics reliably result in divergence away from the REE. Thus, as found in uni-variate models, such as the original beauty contest game, negative feedback also seems to play a role in disciplining convergence in multivariate, planar systems.

Finally, we have compared the performance of a variety of different learning models found in the literature on learning in games and learning in macroeconomics in terms of their convergence dynamics and the fit of the estimated model. Indeed, an important contribution of our paper is to 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 that we consider, we find that a mixed, cognitive levels model (with levels 0 and 1), which is analogous to the single-parameter adaptive learning model, outperforms all other considered learning models in terms of the out-of-sample mean squared error of its predictions.

There are several directions in which the evaluation of learning models in multivariate systems using experimental data might profitably proceed. For instance, one could consider *fully coupled* systems or systems with more than 2 dimensions or systems with a more explicit dynamic linkage between periods. While we think all of these extensions are interesting and worth pursuing, we leave such exercises to future research.

APPENDIX

A Motivating Examples

In this appendix, we provide two motivating examples of economic models that map into the basic framework described in 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 the 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 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. Note that we assume that there are neither economies, not dis-economies of scope.

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

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

where p_i^A is the price of the product, $\bar{p}_{-i}^A = \frac{1}{N-1}\sum_{j\neq i} p_j^A$ is the average price charged by all other firms in market A. Parameter γ^{AA} specifies the relationship between the products of various firms in this market: for $\gamma^{AA} > 0$ the goods are substitutes, whereas for $\gamma^{AA} < 0$ they are complements. Linear demand as in (A.1) can be derived from consumer's linear-quadratic utility function.

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

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

where p_i^B is the price of the product in market B and $\bar{p}_{-i}^B = \frac{1}{N-1}\sum_{j\neq i} p_j^B$ is the average price charged by all other firms in this market. Parameter γ^{BB} specifies the relationship between the products of various firms in market B in the same way as γ^{AA} did it for market A: for $\gamma^{BB} > 0$ the goods are substitutes, whereas for $\gamma^{BB} < 0$ 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 depend on the average price in market A.[58] When $\gamma^{AB} > 0$ the higher average price of products sold in market A lead to a higher demand of good 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

$$\pi_i(p_i^A, p_i^B; \bar{p}_{-i}^A, \bar{p}_{-i}^B) = \left(p_i^A - c\right) D_i^A(p_i^A, \bar{p}_{-i}^A) + \left(p_i^B - c\right) D_i^B(p_i^B, \bar{p}_{-i}^A, \bar{p}_{-i}^B)\,.$$

The first-order conditions are given by

$$\frac{\partial \pi_i}{\partial p_i^A} = \alpha^A - \beta^A p_i^A + \gamma^{AA}\bar{p}_{-i}^A - \beta^A(p_i^A - c) = 0$$
$$\frac{\partial \pi_i}{\partial p_i^B} = \alpha^B - \beta^B p_i^B + \gamma^{AB}\bar{p}_{-i}^A + \gamma^{BB}\bar{p}_{-i}^B - \beta^B(p_i^B - c) = 0$$

[58] Another 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).

Solving these equations we find optimal prices of firm i for markets A and B:

$$p_i^A = \frac{1}{2\beta^A}\left(\alpha^A + \gamma^{AA}\bar{p}_{-i}^A + c\beta^A\right) ,$$
$$p_i^B = \frac{1}{2\beta^B}\left(\alpha^B + \gamma^{AB}\bar{p}_{-i}^A + \gamma^{BB}\bar{p}_{-i}^B + c\beta^B\right) .$$

Thus the price-setting game with N firms is exactly equivalent to our planar beauty contest game as introduced in Section 2.

A.2 New Keynesian model example

For another motivating example of a planar system studied by macroeconomists that is consistent with our setup, consider a contemporaneous expectations (static) version of a New Keynesian 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:

$$\begin{aligned} \pi_t &= \pi_t^e + \kappa y_t \\ y_t &= y_t^e - \varphi(i_t - \pi_t^e - r) \\ i_t &= \lambda(\pi_t - \pi) \end{aligned}$$

In these equations, π_t^e denotes expected inflation, y_t^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.[59]

The system can be rewritten in the following matrix form:

$$\begin{pmatrix} \pi_t \\ y_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} \pi_t^e \\ y_t^e \end{pmatrix} + \boldsymbol{d}$$

[59] Modern versions of the model have forward looking expectations, i.e. π_{t+1}^e and y_{t+1}^e in place of π_t^e and y_t^e, but this timing difference only complicates the timing of payoff realizations in our experiment and is not our main question of interest, which concerns the ability of agents to learn the steady state of various parameterizations of coupled, planar system.

where

$$\mathbf{M} = \Omega \begin{bmatrix} 1 + \kappa\varphi & \kappa \\ 1 & \varphi(1-\lambda) \end{bmatrix},$$

$$\boldsymbol{d} = \Omega \begin{pmatrix} \kappa\varphi \\ \varphi \end{pmatrix} (\lambda\pi + r)$$

and

$$\Omega \equiv \frac{1}{1 + \kappa\varphi\lambda}$$

Note that this is an example of a multivariate, coupled system of equations where expectations of the two endogenous variables both matter for the realizations of those variables. In our experiment, we take π_t^e and y_t^e to be the *average* forecasts of the n agents (subjects) who have full knowledge of the data generating process.

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

B Equilibria of the Game

Consider the planar beauty contest game with matrix $\mathbf{M}$, such that $\mathbf{I}-\mathbf{M}$ is invertible matrix (in other words, 1 is not among the eigenvalues of $\mathbf{M}$). This game with individual strategy space $\mathbb{R}$, i.e., with no bounds on possible guesses, has a unique Nash equilibrium, and it is given by the strategies in (4).

To prove it note that when $a_i = a^E$ and $b_i = b^E$ for every i, the averages are $\bar{a} = a^E$, $\bar{b} = b^E$, and from (2) the target values are

$$\begin{pmatrix} a^* \\ b^* \end{pmatrix} = \mathbf{M}(\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d} + \boldsymbol{d} = \mathbf{M}(\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d} + (\mathbf{I}-\mathbf{M})(\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d} = (\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}\,.$$

Thus the target values coincide with individual guesses for both numbers, which leads to the maximum possible payoff for a participant according to (3). Any deviation from such profile will lead to lower payoffs.

Since in all our treatments point $(a^E, b^E) = (90, 20)$ lies in the interior of the strategy space $[0, 100] \times [0, 100]$, the same profile is always the Nash equilibrium in our experiment. However, other Nash equilibria are also possible due to the boundaries in the strategy space.

We will first show that the asymmetric equilibria are impossible in the game. Assume that there is a pair of different strategies in equilibrium profile. Without loss of generality we assume that the a-numbers differ, and 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 that $a_L < \bar{a} < a_R$. We will show now that if the target is to the left from $\bar{a}$, then strategy a_R can be replaced by another strategy (without changing the b-number denoted as b) a with larger payoff. (If the target is to the right from $\bar{a}$, a symmetric argument can be made with replacing a_L strategy.)

Indeed, let us decrease guess a_R by a small amount $\varepsilon > 0$. This will decrease the average of a-numbers by ε/N, and change the a-target to $a^* - \varepsilon m_{11}/N$. The distance between the new target and new strategy is $a_R - a^* - \varepsilon(N - m_{11})/N$. The target for the b-number has changed to $b^* - \varepsilon m_{21}/N$. Therefore, 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 as the strategy (a_R, b) was replaced by $(a_R - \varepsilon, b)$, there is an increase in performance as measured by a sum of absolute deviations of a- and

b-numbers by at least $\varepsilon(N - m_{11} - |m_{21}|)/N$ and this quantity is strictly positive for $N = 10$ as in our treatments.

Since only symmetric equilibria are possible, all participants submit the same number in any equilibria. But then the extra equilibria in the beauty contest game with bounds are possible only when participants submit one or both guesses on the bounds and receive the targets outside of the bounds. It is then straightforward to check the following:

- Evolution for a-number will give extra equilibria 0 and 100, but only in the **Source** treatment.
- Evolution for b number in the equilibrium with $a^E = 90$ will give extra equilibria 0 and 100 but only in the **SaddlePos** treatment.
- In the **Source** treatment: in the a-number equilibrium 0 the evolution for b number has a unique equilibrium 38; in the a-number equilibrium 100 the evolution for b number has a unique equilibrium 18.

This completes proofs of the statements of Section 2 summarised 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 $1, 2, \ldots, 15$ 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:

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

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

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”:

$$50.50 \qquad 100 \qquad 20.25$$

and “B-numbers”

$$10.79 \qquad 30 \qquad 0\,.$$

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:

Your payoff in points each period =

$$\frac{500}{5 + |\text{your "A-number"} - \text{A*}| + |\text{your "B-number"} - \text{B*}|)}$$

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:

$$\frac{500}{5 + |50.50 - 67.95| + |10.79 - 39.7|} = \frac{500}{5 + 17.45 + 28.95} = 9.73 \text{ "points"}$$

(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

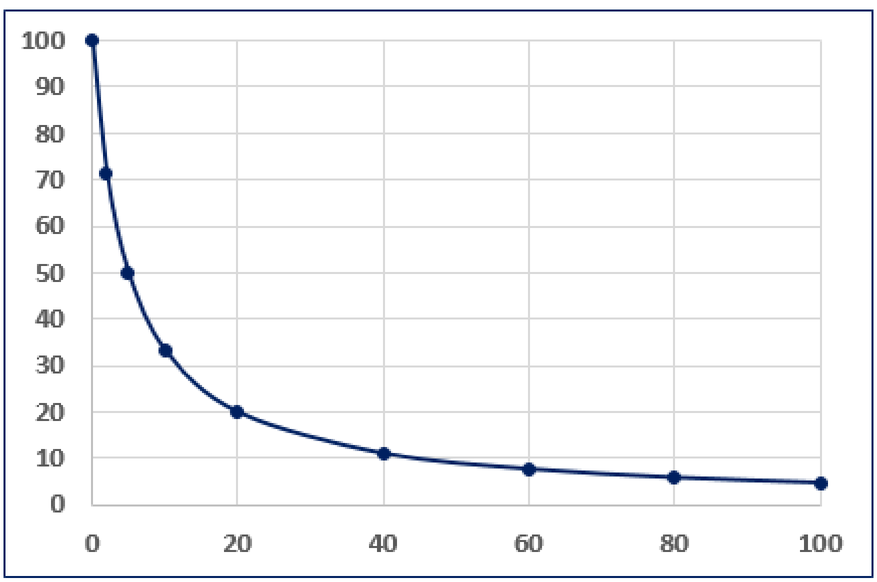
Points Table The table gives the number of points for a given discrepancy of “A-number” from the target value A* (the first column) and a given discrepancy of “B-number” from the target value B* (the first row).

The figure below shows the relation between the number of points you score (vertical axis) and the combined discrepancy

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

|B-B*| / |A-A*|	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

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 Speed of Convergence

In addition to the first hit time as reported in Table 3, we made an across treatment comparison of the latest experimental period when the trajectory entered the ε-neighborhood to stay there until the end of the experiment (so to say, the period of "irreversible entry"). This statistics is only relevant for the sessions when the trajectory entered the neighborhood at least once. Hence, we do not compute it for **SaddlePos** and **Source** treatments. Formally, given $\varepsilon > 0$, the latest time to irreversibly enter the ε-neighborhood, $\tau(\varepsilon)$, is the period such that $(\bar{a}_{\tau(\varepsilon)-1}, \bar{b}_{\tau(\varepsilon)-1}) \notin U_\varepsilon$ and $(\bar{a}_t, \bar{b}_t) \in U_\varepsilon$ for any $t \geq \tau(\varepsilon)$.

Results presented in Table 9 confirm the conclusion made in the main text for comparison between the **SaddleNeg** and **Sink** treatments. Not only convergence

	Sink				**SaddleNeg**			
ε	Sess. 1	Sess. 2	Sess. 3	Sess. 4	Sess. 1	Sess. 2	Sess. 3	Sess. 4
20	5	8	4	4	5	2	4	3
10	7	-	6	15	9	5	5	5
5	8	-	-	-	11	6	8	7
1	10	-	-	-	-	12	12	9
0.5	15	-	-	-	-	14	13	14

Table 9: The latest period when the trajectory enters the ε-neighborhood of equilibrium irreversibly, i.e., to stay there until the end of the experiment. The hyphen denotes the cases when the trajectory was outside of the ε-neighborhood in the last period of the experiment.

	Sink				**SaddleNeg**				**SaddlePos**			
ε	Sess. 1	Sess. 2	Sess. 3	Sess. 4	Sess. 1	Sess. 2	Sess. 3	Sess. 4	Sess. 1	Sess. 2	Sess. 3	Sess. 4
20	5	8	4	4	5	2	4	3	3 (9)	7	4	3
10	7	11 (14)	6	11 (15)	9	5	5	5	11 (14)	11	7	6
5	8	-	8 (-)	-	11	6	8	7	15	14	10	8 (11)
1	10	-	14 (-)	-	15	12	12	9	-	-	-	-
0.5	10 (15)	-	-	-	-	14	13	9 (14)	-	-	-	-

Table 10: The first hit times for a-number trajectories for different neighborhoods. The hyphen denotes the cases when the trajectory never entered the ε-neighborhood. The periods of irreversible entry are shown in parentheses when they differ from the first hit time.

happens faster in the **SaddleNeg** treatment, it is also the case that in this treatment the trajectories stay in the vicinity of the equilibrium for a longer time before the end of the experiment.

We will now make the same analysis for the trajectory of a-numbers only. For $\varepsilon > 0$, the neighborhood is defined as $U_\varepsilon = (a^E - \varepsilon, a^E + \varepsilon)$. Table 10 shows the first hit time (and the period of irreversible entry in parentheses) for all sessions of the **Sink**, **SaddleNeg**, and **SaddlePos** treatments for five different values of ε. (In the **Source** treatment, the trajectories never entered the U_ε neighborhood for $\varepsilon \leq 20$.

Table 10 suggests that the quickest convergence of a-number was in the **SaddleNeg** treatment. It also shows that the convergence in **SaddlePos** treatment was the slowest when the relatively small neighborhood around equilibrium is considered.

E Proof of Proposition 4.1

Let $\mathbf{H}(K)$ denote the operator that describes the dynamics of the CH model, i.e., the map from period $t-1$ to period t deviations of a and b numbers from the Nash equilibrium. We need to prove that $\mathbf{H}(K) = \prod_{i=1}^{K}(\lambda_i \mathbf{M} + (1-\lambda_i)\mathbf{I})$ with λ's defined as above.

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, we conclude that $\mathbf{H}(1) = f_1\mathbf{M} + (1-f_1)\mathbf{I}$, as required.

Assume now that 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 the level 2 is the highest level in population, we can complete the model, by weighting this best response with f_2 and putting the remaining weight $1-f_2$ to the average play of level-0 and 1 agents as defined above. This leads to the operator

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

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

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

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

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 puts the remaining weight to the average play of level-1 and 0 agents. (This, 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.) The level-3 agents' best response is then $\mathbf{M}\big(\lambda_2\mathbf{M} + (1-\lambda_2)\mathbf{I}\big)\big(\lambda_1\mathbf{M} + (1-\lambda_1)\mathbf{I}\big)$. If the level 3 is the highest level in population, we can complete the model, by weighting this best response with f_3 and putting the remaining weight $1 - f_3$ to the average play of levels 0, 1 and 2 agents. This leads to the operator

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

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

Continuing in the same way we can prove the statement for any K.

F Appendix (Not for publication). Additional Information on Experiment

In this Appendix we collect extra information about our experiment.

Figure 6 shows a scatter plot of individual choices in period 1. The four panels correspond to our four treatments. Every point correspond to one or more individuals submitting a-guess as shown on the horizontal axes together with the b-guess as shown on the vertical axes. The frequencies of the choices are indicated by the size of the circles: the larger the circle, the more individuals submitted the corresponding pair of guesses.

On top of this scatter plot (a sort of two-dimensional histogram) we superimpose the lines indicating various levels of rationality as defined in Section 3.1. The thick red lines correspond to the rational expectation choice (a^E, b^E). The other lines indicate various levels of rationality: level 0, i.e., $(50, 50)$ guesses, is shown by the thin solid lines, level 1 is shown by the dashed lines, level 2 is shown by the dashed-dotted line, and level 3 is shown by the dotted line.

The choice on the intersection of two lines corresponding to the same level of rationality would indicate an individual consistency in levels of rationality for a and b numbers. Inspection of Fig. 6 shows that even if the choices are quite dispersed, there are few clear cases of consistency when participants apply level 0 to their both choices or derive the REE. For other levels of rationality, we do not observe such consistency.

Figs. 7 to 10 show the dynamics of average guesses for each session in treatments **Sink**, **SaddleNeg**, **SaddlePos**, **Source**, respectively. The left panels show the time evolution of $\bar{a}$ (thick red line) and $\bar{b}$ (thin blue line). The dashed lines show the levels of the REE, $a^E = 90$ and $b^E = 20$. The middle and right panels show the same evolution as the phase diagram (the right panel shows the zoomed version of the middle panel).

Figs. 11 to 14 show the dynamics of *target* values for each session in treatments **Sink**, **SaddleNeg**, **SaddlePos**, **Source**, respectively. The left panels show the time evolution of a^* (thick red line) and b^* (thin blue line). The dashed lines show the REE levels, $a^E = 90$ and $b^E = 20$. In the only equilibrium where all participants guess the

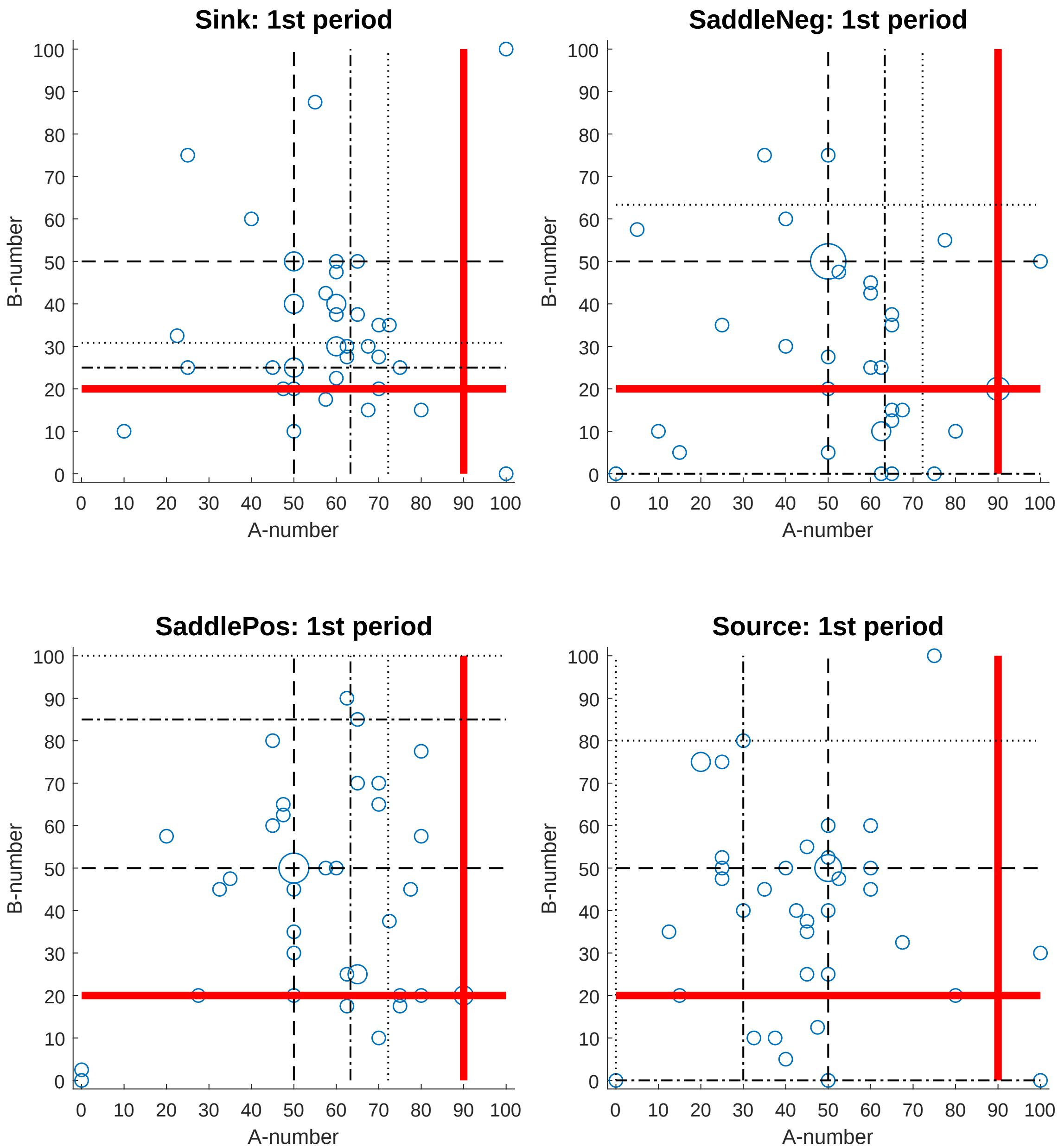


Figure 6: Frequencies of individual pair of guesses in period 1 and levels of reasoning (different dashed lines) and the REE (red thick line) for four experimental treatments.

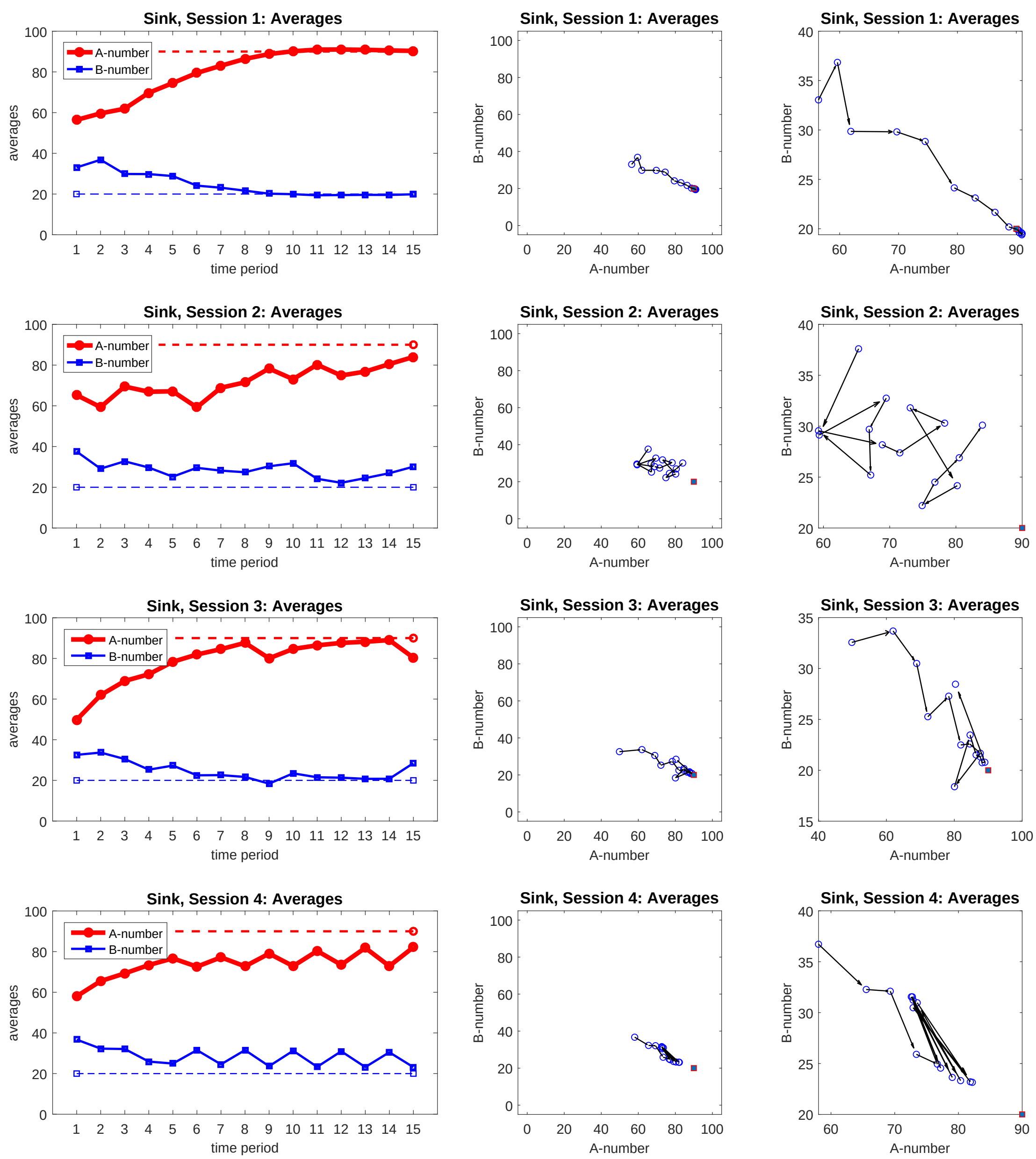


Figure 7: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **Sink** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

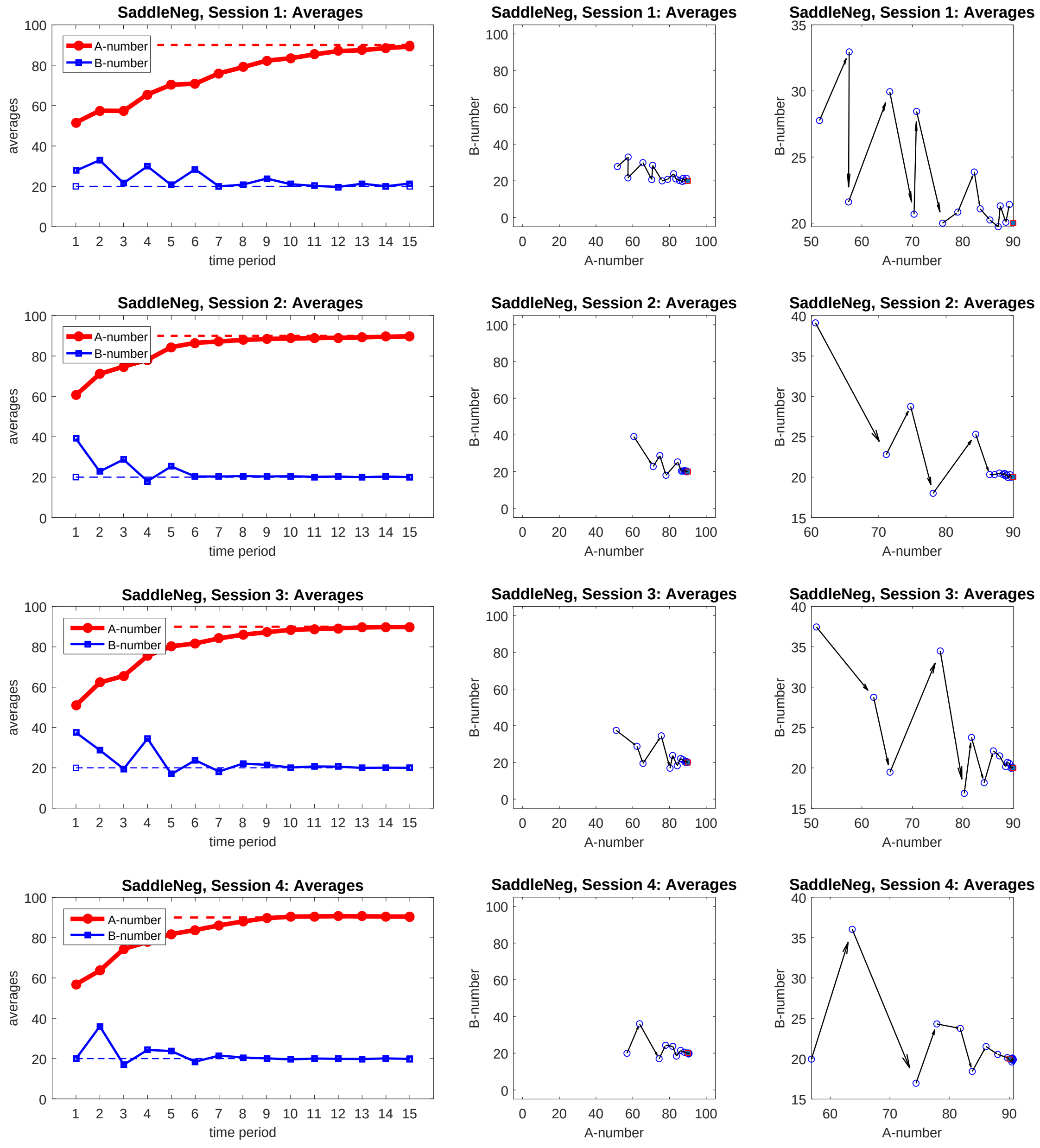


Figure 8: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **SaddleNeg** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

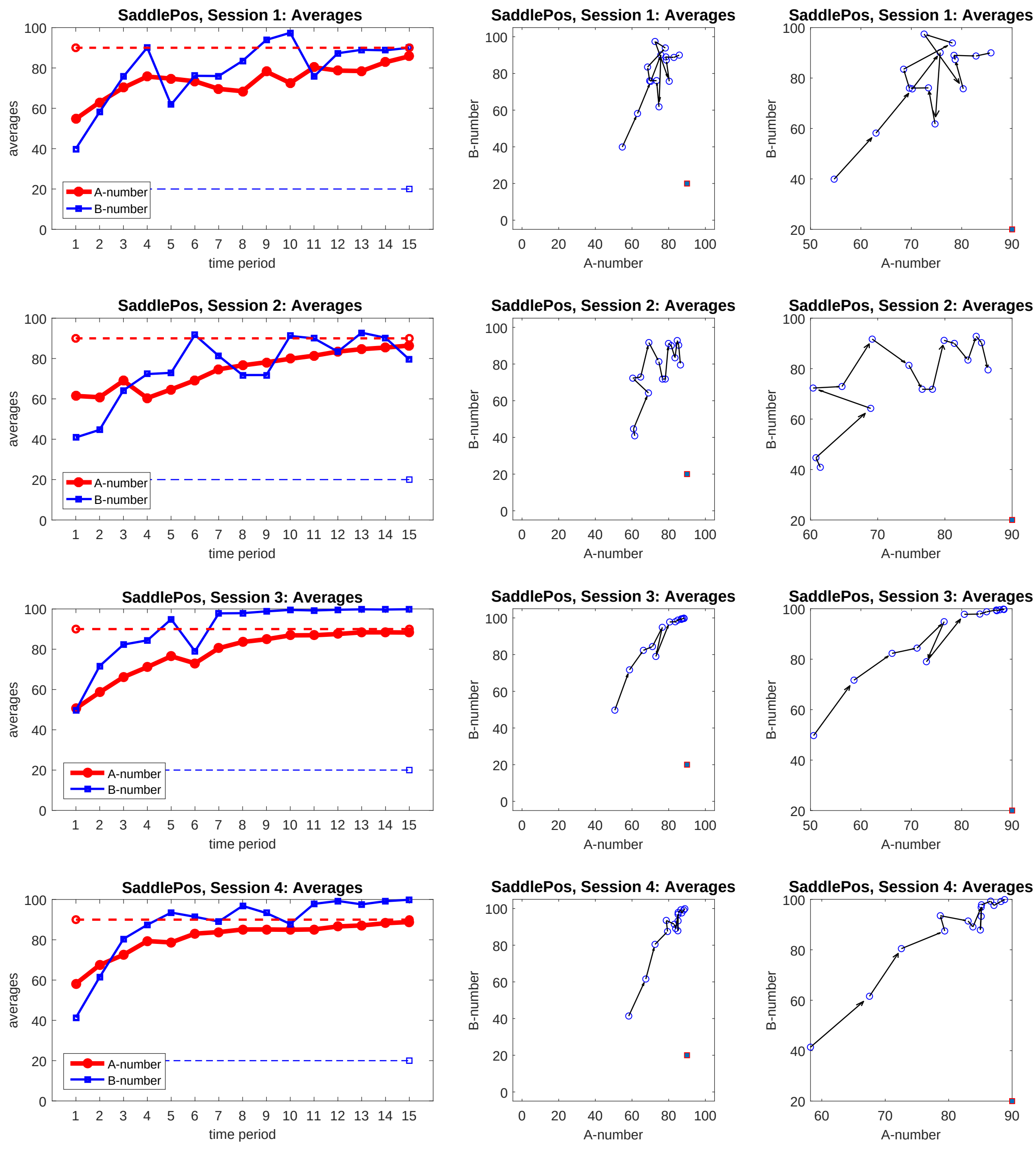


Figure 9: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **SaddlePos** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

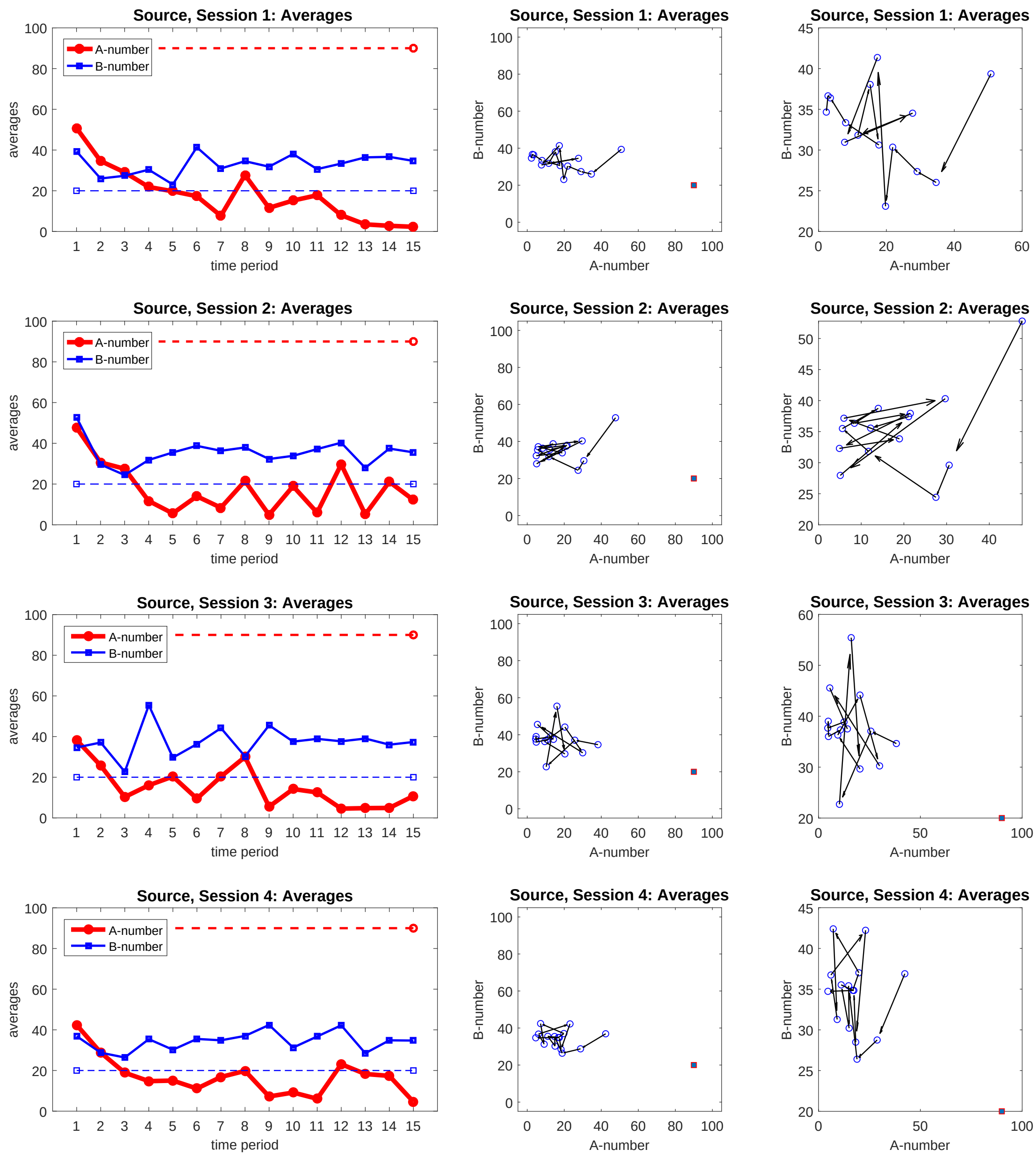


Figure 10: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **Source** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

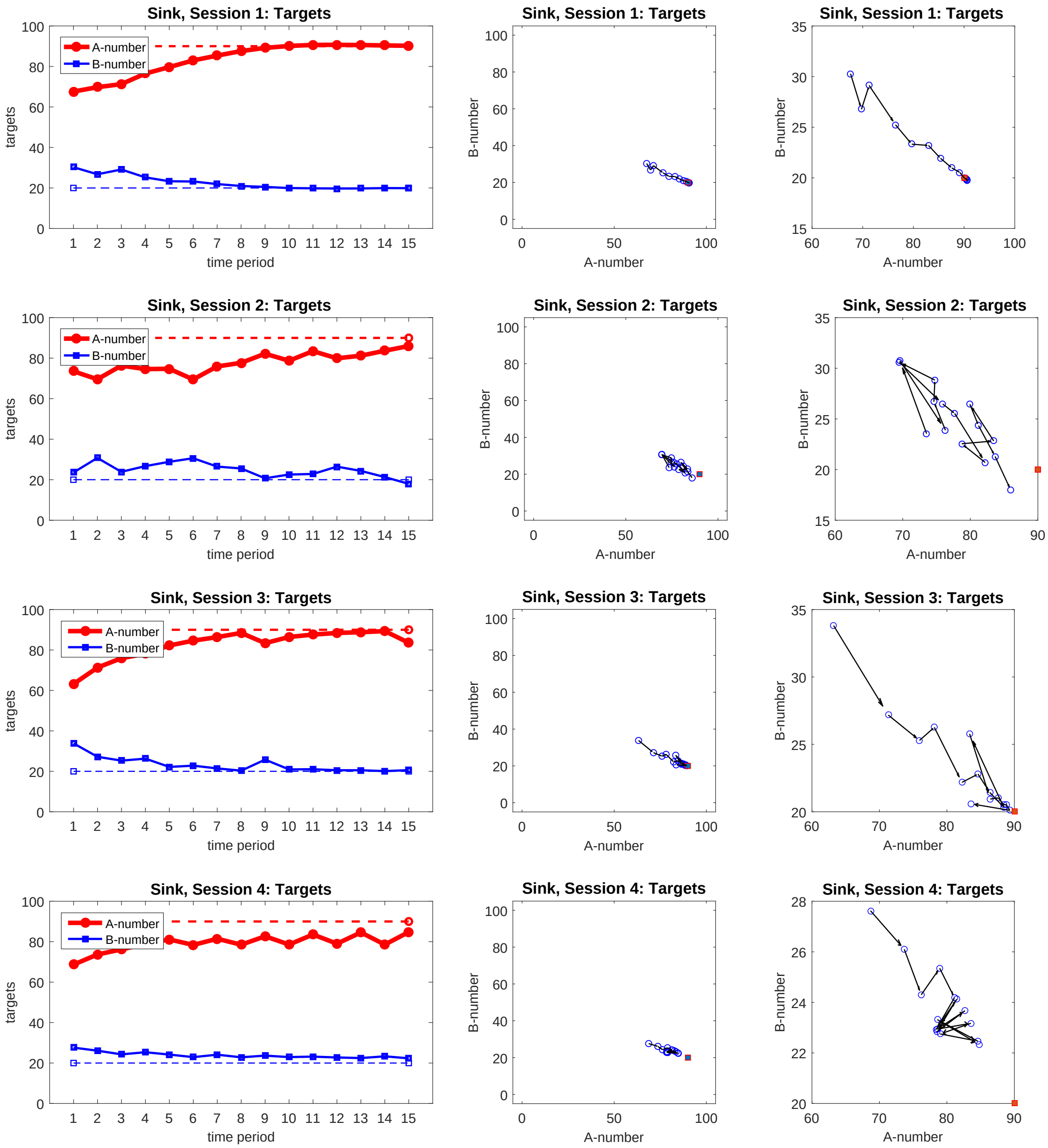


Figure 11: Dynamics of the target values, a^* and b^* in the **Sink** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

targets correctly, the targets would be on these levels. The middle and right panels show the same evolution as the phase diagram (the right panels shows the zoomed version of the middle panel). Note that targets may lie out of the range $[0, 100]$, and indeed this happens in treatments **SaddlePos** and **Source**.

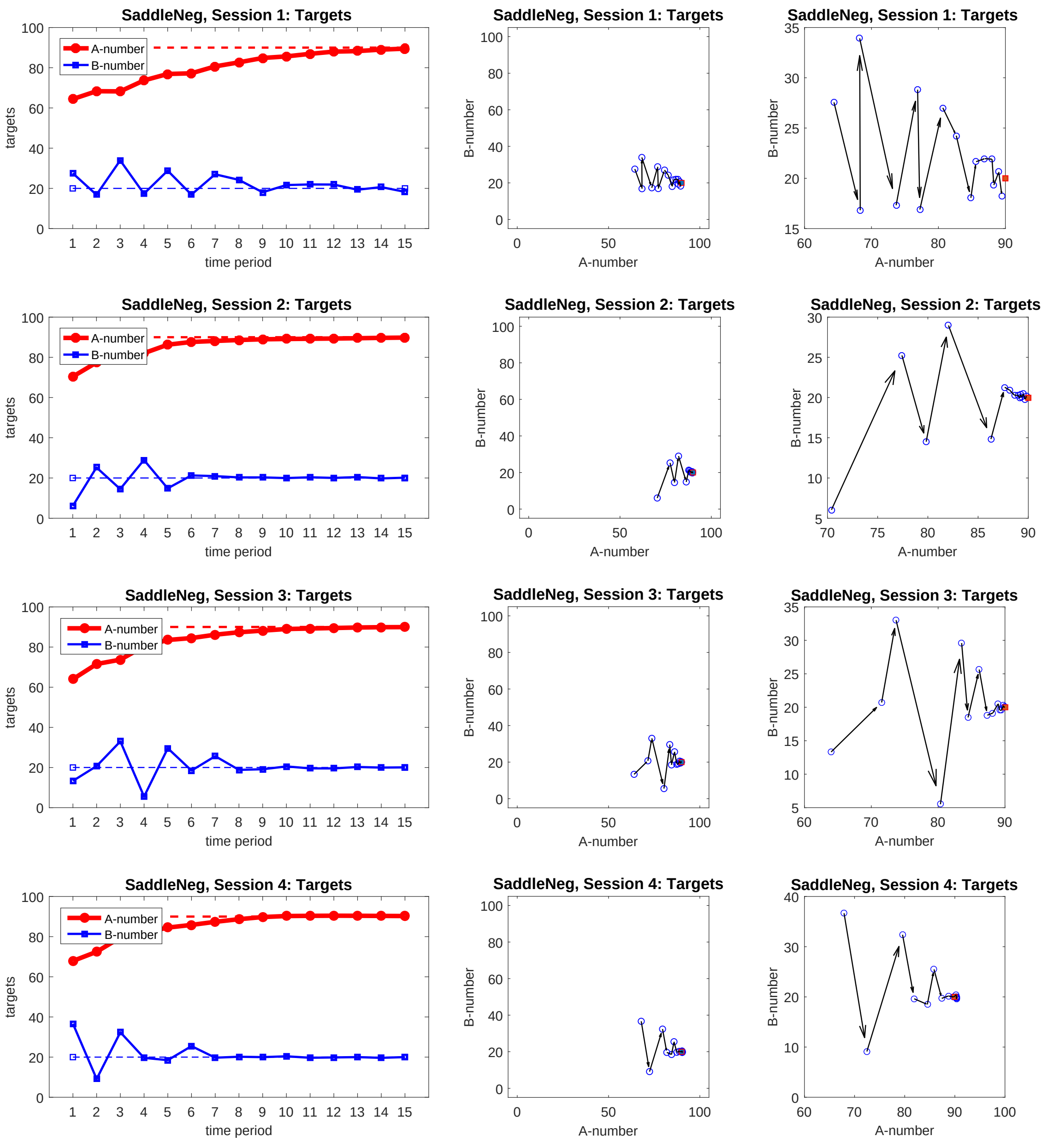


Figure 12: Dynamics of the target values, a^* and b^* in the **SaddleNeg** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

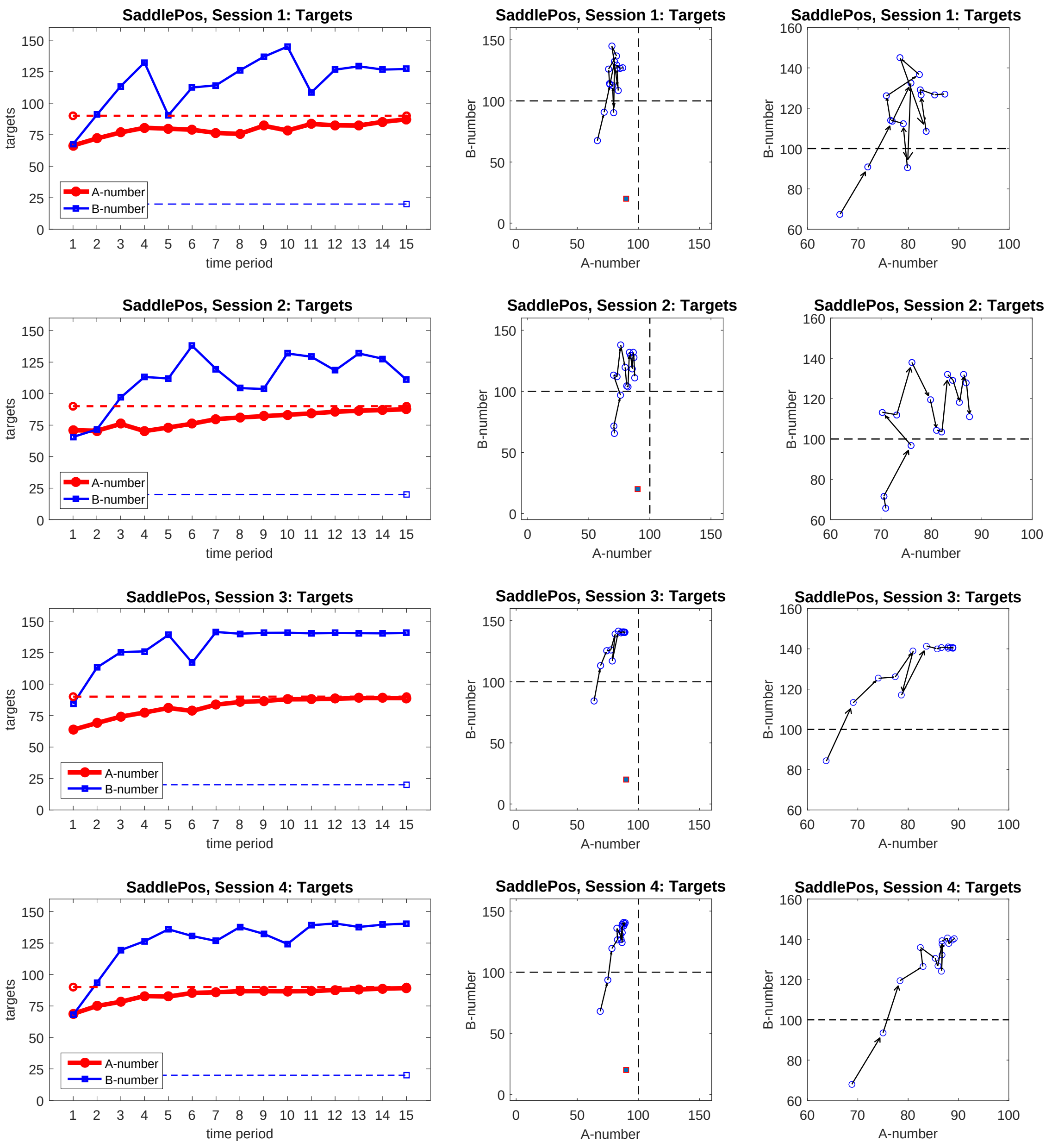


Figure 13: Dynamics of the target values, a^* and b^* in the **SaddlePos** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

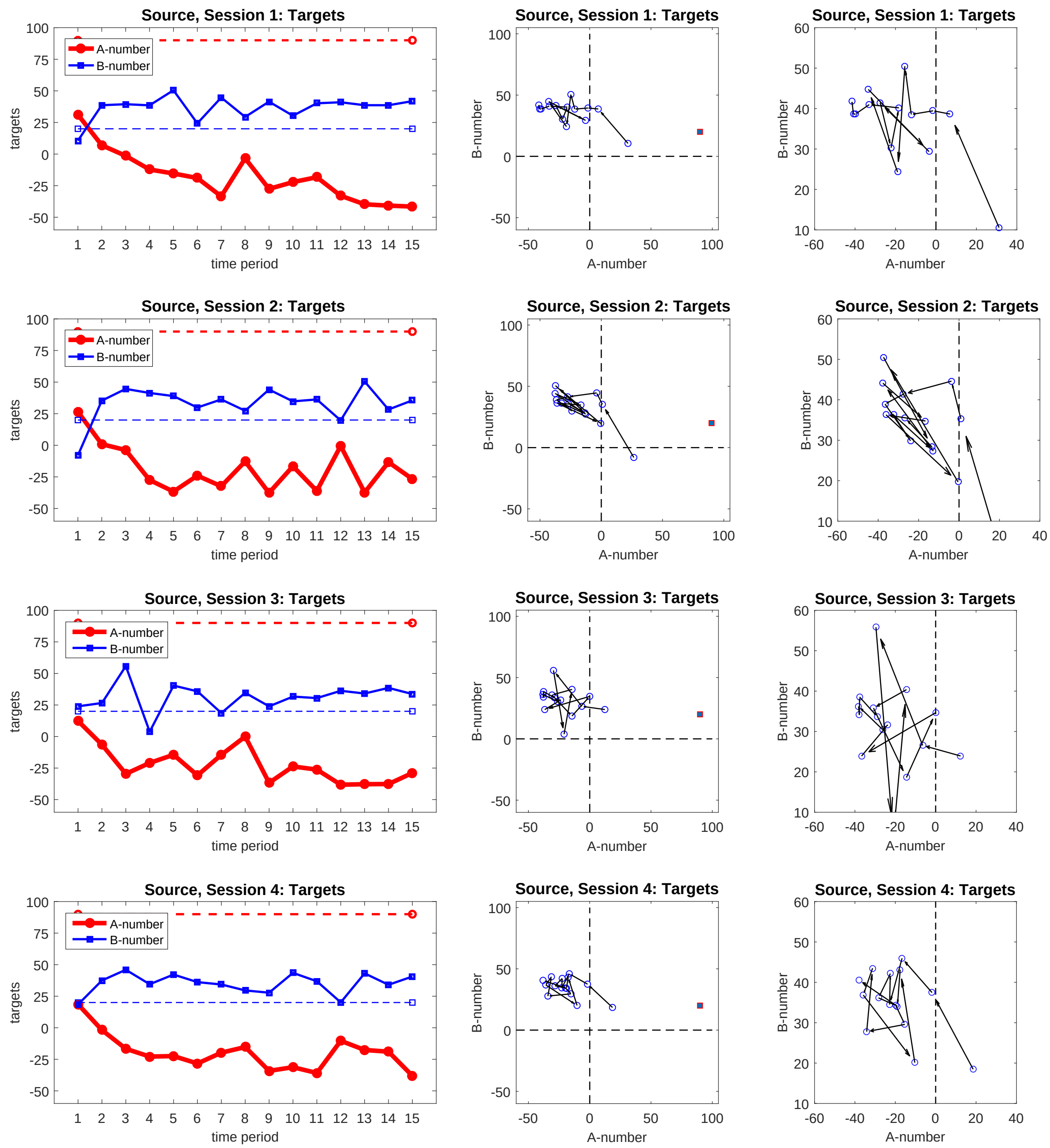


Figure 14: Dynamics of the target values, a^* and b^* in the **Source** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

G Online Appendix (Not for publication). Dynamics of the adaptive models

In Section 4.3 we introduced the adaptive model. In this section, we start with a more general version of this model when different adaptation rates are used for a and b numbers. We obtain formal results on the properties of this model that, in the special cases, lead to the results we used for the naïve and adaptive models in the main text. Then, we illustrate the dynamics of the adaptive model of Section 4.3 for four various values of parameter λ and discuss the properties of this model for our four treatments; see Table 5 of the main text.

In the matrix form, the general adaptive model can be written as

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \mathbf{\Lambda} \begin{pmatrix} a_t^* \\ b_t^* \end{pmatrix} + (\mathbf{I} - \mathbf{\Lambda}) \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} \tag{G.1}$$

where we introduced diagonal matrix:

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_a & 0 \\ 0 & \lambda_b \end{pmatrix} .$$

Note that in this model different adaptation rates, λ_a and λ_b, are used for a and b numbers. Such a model can be obtained as a mixture of the homogeneous level-0 and level-1 models, when it is assumed that distribution of agents across those levels of rationality is different for a and for b numbers. We estimate this model in Section 5.

Substituting into (G.1) equation (2) that defines the target value, we obtain

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = (\mathbf{I} - \mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{M}) \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + \mathbf{\Lambda}\boldsymbol{d} \tag{G.2}$$

Proposition G.1. *Let us assume that matrix $\mathbf{I} - \mathbf{M}$ is not invertible and $\lambda_a\lambda_b \neq 0$.*

Then there exists a unique steady state of dynamics (G.2) *given by* $(\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}$.

Proof. Denote the steady state of the system as $\begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix}$. Then at the steady state we have

$$\begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix} = (\mathbf{I}-\boldsymbol{\Lambda}+\boldsymbol{\Lambda}\mathbf{M})\begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix} + \boldsymbol{\Lambda}\boldsymbol{d} \quad \Leftrightarrow \quad \mathbf{0} = -\boldsymbol{\Lambda}(\mathbf{I}-\mathbf{M})\begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix} + \boldsymbol{\Lambda}\boldsymbol{d}$$

As neither of λ_1 and λ_2 is zero, matrix $\boldsymbol{\Lambda}$ is invertible. Therefore we can simplify to

$$(\mathbf{I}-\mathbf{M})\begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix} = \boldsymbol{d} \quad \Leftrightarrow \quad \begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix} = (\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}\,.$$

This proves the statement. □

Therefore, the dynamics (G.2) has a unique steady state coinciding with the Nash equilibrium of our game (4). Dynamics of this model then can be written in deviations from this steady state

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} - \begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix} = \widetilde{\mathbf{M}}\left(\begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} - \begin{pmatrix} \bar{a}^E \\ \bar{b}^E \end{pmatrix}\right)\,.$$

with matrix $\widetilde{\mathbf{M}} = \mathbf{I}-\boldsymbol{\Lambda}+\boldsymbol{\Lambda}\mathbf{M}$. The stability properties of the model can now be expressed in terms of the elements of matrix $\mathbf{M}$.

Proposition G.2. *Consider dynamics* (G.2) *and assume that matrix* $\mathbf{M}$ *is lower triangular. The steady state of the dynamics,* $(\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}$, *is globally stable if and only if the following conditions are satisfied*

$$-2 < \lambda_a(m_{11}-1) < 0 \quad \textit{and} \quad -2 < \lambda_b(m_{22}-1) < 0\,, \tag{G.3}$$

where m_{11} *and* m_{22} *are the diagonal elements of matrix* $\mathbf{M}$.

Proof. The stability of the steady state depends on the eigenvalues of matrix

$$\widetilde{\mathbf{M}} = \mathbf{I} - \mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{M} = \begin{pmatrix} 1 - \lambda_a + \lambda_a m_{11} & \lambda_b m_{12} \\ \lambda_a m_{21} & 1 - \lambda_b + \lambda_b m_{22} \end{pmatrix}.$$

Since matrix $\mathbf{M}$ is lower triangular, $m_{12} = 0$. Then, matrix above is also lower triangular and its eigenvalues are

$$\mu_1 = 1 - \lambda_a + \lambda_a m_{11} \qquad \text{and} \qquad \mu_2 = 1 - \lambda_b + \lambda_b m_{22}$$

The standard condition for local stability is that both eigenvalues are less than 1 in absolute value. As our system is linear, these conditions are also necessary and sufficient for the global stability of the steady state. This proves the statement. □

With these results we obtain the results in Section 4. First, the naïve model, as introduced in Section 4.1, is a special case of model (G.1) where $\lambda_a = \lambda_b = 1$. In this case matrix $\widetilde{\mathbf{M}}$ that governs the dynamics is simply matrix $\mathbf{M}$ that we used in the experiment, and the eigenvalues are m_{11} and m_{22}. This justifies Table 4.

Next, the adaptive model, as described in Section 4.3, is a special case of model (G.1) with common $\lambda = \lambda_a = \lambda_b$. This justifies our focus on eigenvalues as in (12) and results in Table 5. We will now discuss it in more details for each treatment.

The dynamics of the adaptive model is illustrated in Figs. 15 to 18 that compare dynamics in the experiment with dynamics of the adaptive model with $\lambda = 1$ (i.e., naïve model), $\lambda = 0.75$, $\lambda = 0.5$, and $\lambda = 0.25$, respectively. For each case the phase plots of experimental data in the first session of the corresponding treatment (left panel) are compared with the simulated dynamics. In the middle panels the initial conditions are chosen to be $(50 \quad 50)'$, i.e., they are the same in all treatments and models. In the right panels the initial conditions are selected to be equal to the average values in the first period in the corresponding experimental session.

We now discuss the results of dynamics of the adaptive model presented in Table 5 for our four treatments.

Sink. In this treatment matrix $\mathbf{M}$ has eigenvalues $m_{11} = 2/3$ and $m_{22} = -1/2$. As both are inside the unit circle, the eigenvalues of adaptive model (which are weighted averages of 1 and eigenvalues of $\mathbf{M}$, see (12)) are also inside the unit circle. Therefore, the adaptive model will always converge. Variable a converges monotonically, and when λ decreases from 1 (naïve model) to 0, the rate of convergence of a, given by $\mu_1 = 1 - \lambda/3$ will be higher (so the convergence will be slower). Dynamics of b will jump around the eigenvector $\mathbf{v}_1$ for $\lambda > 2/3$. See the upper panel of Figs. 15 and 16 for illustrations of such case. When $\lambda < 2/3$, both eigenvalues become positive and dynamics does not jump through the eigenvector. It will then be monotone both for a and b, with rate of both variables getting closer to 1 (i.e., slower convergence). Compare the upper panel of Fig. 17 where $\lambda = 0.5$ and the upper panel of Fig. 18 where $\lambda = 0.25$.

SaddleNeg. In this treatment matrix $\mathbf{M}$ has eigenvalues $m_{11} = 2/3$ and $m_{22} = -3/2$. Therefore, the first eigenvalue in the adaptive model will always be positive and inside the unit circle. Thus dynamics of a is always monotonically converging. Its convergence, whose rate is given by $1 - \lambda/3$, becomes slower with smaller λ. The second eigenvalue will enter the unit circle when λ will be low enough. Thus the steady state will gain stability with decrease of λ. Indeed, the dynamics become (globally) stable with

$$\mu_2 = 1 - \lambda - \frac{3}{2}\lambda = 1 - \frac{5}{2}\lambda > -1 \;\Leftrightarrow\; \lambda < 0.8\,.$$

Thus in all illustrations in Figs. 16-18 the equilibrium is globally stable. When $\mu_2 > 0$, i.e., when $\lambda < 0.4$, we will in addition have convergence without jumping through the eigenvector $\mathbf{v}_1$, implying monotone convergence of b. This is well visible in Fig. 18

when $\lambda = 0.25$. When $0.4 < \lambda < 0.8$, the dynamics of b converges to the equilibrium and jumps through the eigenvector, implying that it may be consistent with the experimental outcome.[60] When $\lambda > 0.8$ the dynamics of b variable diverges. Given that in our simulations we impose bounds as in (7), dynamics of b will be bounded and its attractor will depend on precise value of λ.

SaddlePos. In this treatment matrix $\mathbf{M}$ has eigenvalues $m_{11} = 2/3$ and $m_{22} = 3/2$. As in the previous two treatments, the first eigenvalue of the adaptive model is always between 2/3 and 1. It guarantees monotone convergence of a. The second eigenvalue is positive and outside of the unit circle. Therefore the eigenvalue μ_2 of the adaptive model will be outside of the unit circle for any λ. Dynamics of b will always be unstable, except for a non-generic case, when it starts exactly at vector $\mathbf{v}_1$. This is illustrated in the third row of Figs. 16-18. Note that rate of divergence is related to the value of $\mu_2 = 1 + \lambda/2$. It increases with λ implying that divergence is faster for larger λ. This also can be seen in illustrations. Finally, notice that since we impose the same bounds in our simulation as in the experiment, see Eq. 7, we can compute the final value of b. Indeed, since a converges to 90, dynamics of $\bar{b}_t$ for large enough t can be simplified to $\bar{b}_t = -10\lambda + (1 + \lambda/2)\bar{b}_{t-1}$. From here we can see that if $\bar{b}_{t-1} = 0$, it will stay at this lower bound for the next period, and, hence, forever. The same will happen with the higher bound. Thus, the attractors in the simulations are the same as for the naïve model (see the last column in Table 4).

Source. In this treatment matrix $\mathbf{M}$ has eigenvalues $m_{11} = 3/2$ and $m_{22} = -3/2$ which are both outside of the unit circle. The eigenvalue μ_1 associated with the dynamics of variable a will always be between 1 and 3/2, implying monotone divergence from the steady state. In fact, in all simulations shown in Fig. 16-18 dynamics of a

[60] When $\lambda > 0.4$, the dynamics will jump through the eigenvector $\mathbf{v}_1$. However, this does not necessarily mean that convergence will not be monotonic, as it depends on the initial conditions. For example in the case of $\lambda = 0.75$ as shown in Fig. 16, the dynamics is not monotone, when it starts at $(50, 50)$ (second row, middle panel), but it is monotone when it starts in $(27, 76, 21.41)$ as in the experiment (second row, right panel). The same will be the case when $\lambda = 0.5$, see Fig. 17.

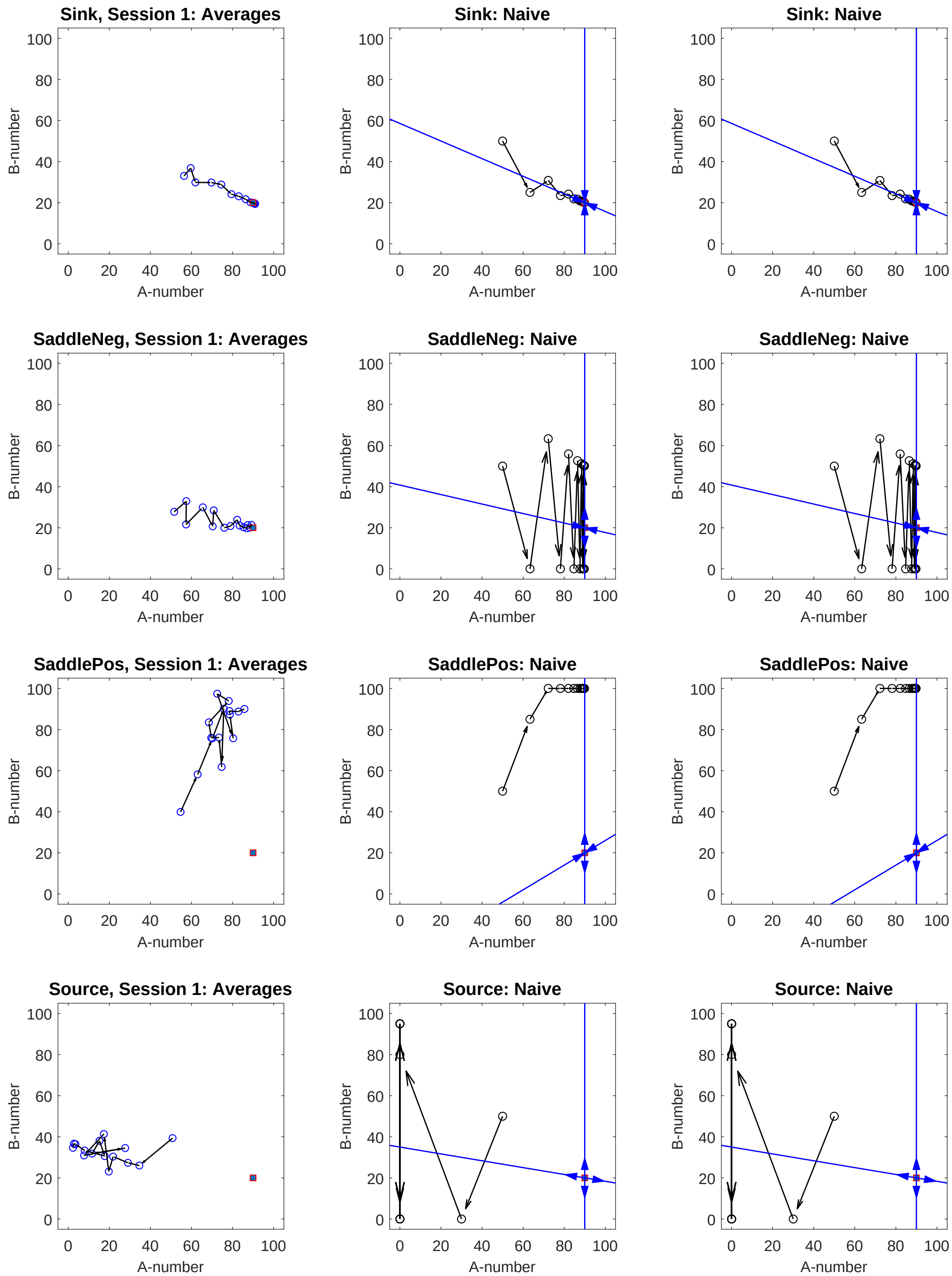


Figure 15: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the naïve model. First 15 periods are simulated using the naïve model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group (*right panels*). The two blue lines are the eigenvectors with the arrows indicating whether the direction is stable or not.

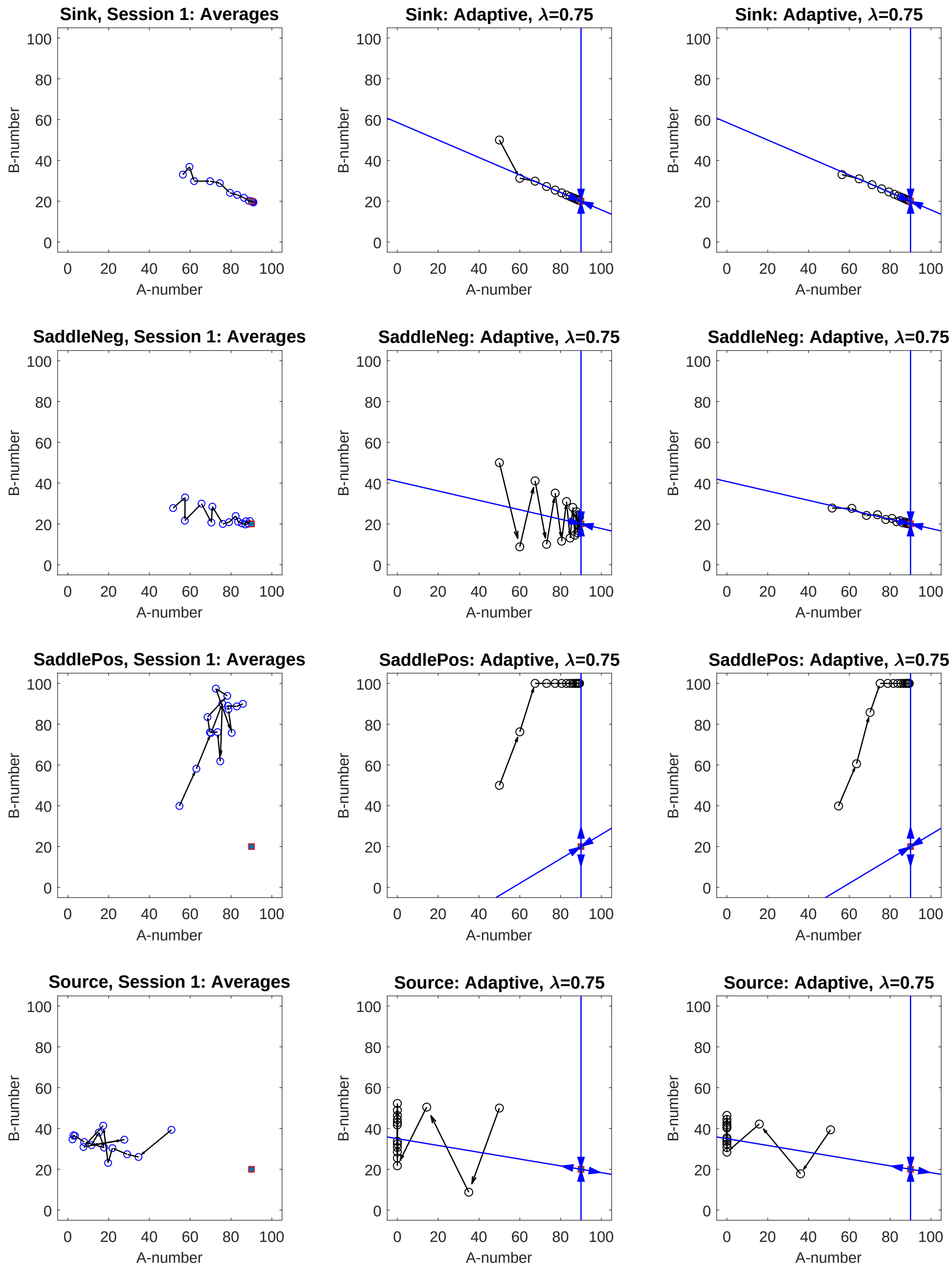


Figure 16: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the adaptive model with $\lambda = 0.75$. First 15 periods are simulated using the adaptive model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group (*right panels*). The two blue lines are the eigenvectors with the arrows indicating whether the direction is stable or not.

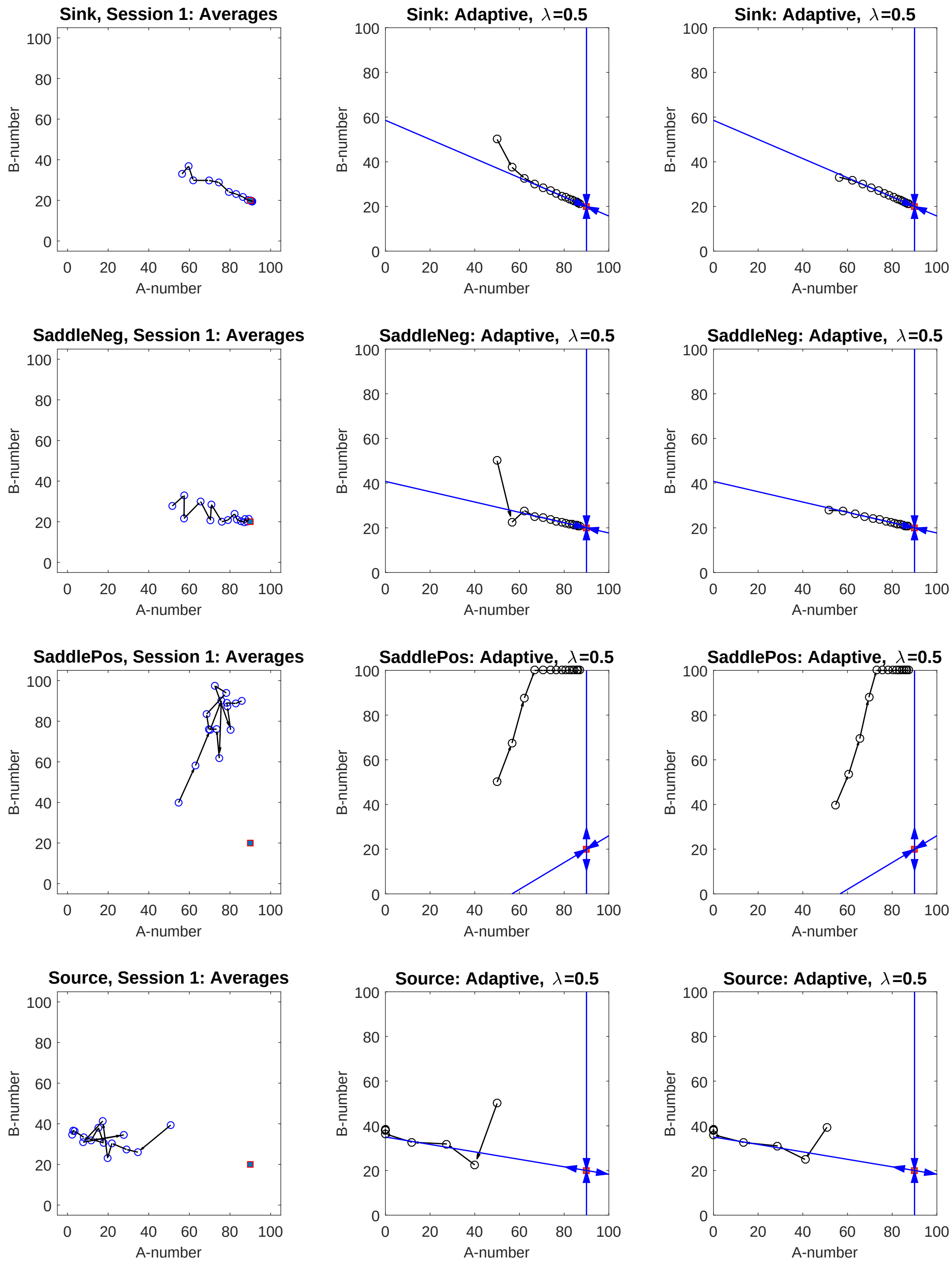


Figure 17: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the adaptive model with $\lambda = 0.5$. First 15 periods are simulated using the adaptive model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group (*right panels*). The two blue lines are the eigenvectors with the arrows indicating whether the direction is stable or not.

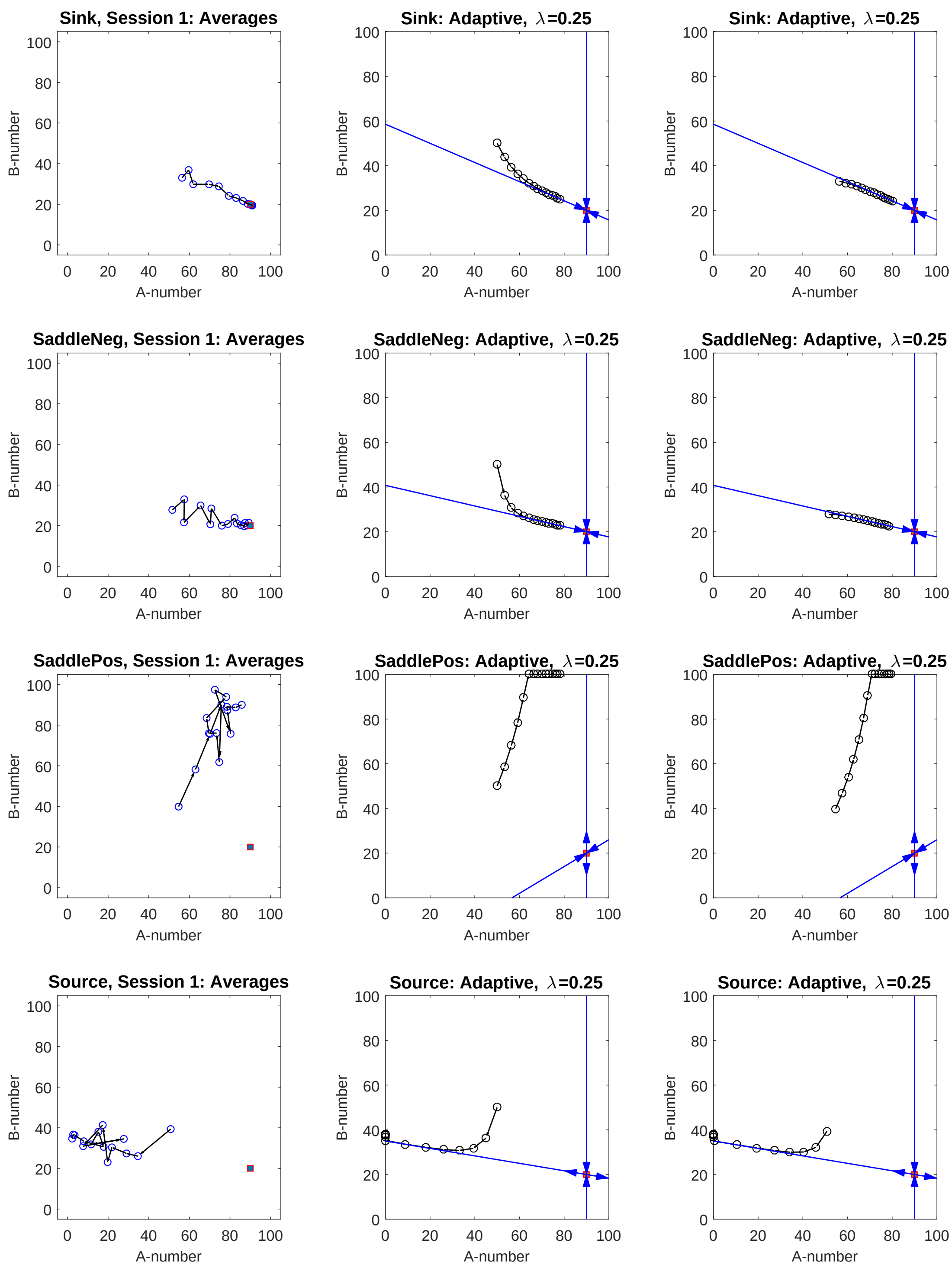


Figure 18: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the adaptive model with $\lambda = 0.25$. First 15 periods are simulated using the adaptive model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group (*right panels*). The two blue lines are the eigenvectors with the arrows indicating whether the direction is stable or not.

diverges to 0. The rate of divergence is $\mu_1 = 1 + \lambda/2$ and thus the smaller λ is, the slower dynamics diverge. The second eigenvalue of the adaptive model will be inside the unit circle when

$$\mu_2 = 1 - \lambda - \frac{3}{2}\lambda > -1 \;\Leftrightarrow\; \lambda < 0.8\,.$$

For this λ dynamics of b will be closer and closer to the eigenvector $\mathbf{v}_1$. The "convergence" towards this vector will be monotone (without jumping through it) for $\lambda < 0.4$, see the last row in Fig. 18 where $\lambda = 0.25$. It will not be monotone for $\lambda > 0.4$ (see the last rows in Figs. 16 and Figs. 17). The eigenvector $\mathbf{v}_1$ intersects $x = 0$ line in the point $y = 35$ and it intersects $x = 100$ line in the point $y = 18.333$. Thus if the dynamics of a diverges to 0, and b is already close to this vector at the moment when a hits the lower bound, the value of b will be close to 35.[61] However, when a will hit the lower bound of 0, it will stay there forever. Then the dynamics of b will be governed by equation $b_{t+1} = (1 - 5\lambda/2)b_t + 95\lambda$, whose steady state is $y = 38$. Instead, if the dynamics of a diverges to 100, then when b is close to this vector at the moment when a hits the upper bound, the value of b will be close to 18.333. But when a will hit the upper bound of 100 and will stay there forever, the dynamics of b will be governed by equation $b_{t+1} = 45\lambda + (1 - 5\lambda/2)b_t$, whose steady state is $y = 18$.

[61] Note that this is very close to the value of 15th period in the experiment.

H Online Appendix (Not for publication). Properties of moving average models

In Section 4.3 we introduced the moving average model with L lags, see Eq. (14). Table 11 summarizes properties of MAve(L) model. When $L = 1$, we have the naïve model, and so the left columns of Table 11 are the same as in Table 4. The next columns correspond to the case of $L = 2$. The system is described by a 4-dimensional system, whose eigenvalues we compute numerically to identify convergence conditions for the whole system.[62] The analytic results for the moving average model with larger L are limited, as the dimension of the system increases (with L) so we must rely on simulations. However, the case when $L \to \infty$, i.e., when participants in each period average *all* past target values, can be studied using the principle of E-stability.

Introduce vector $\boldsymbol{z}_t = (a_t \quad b_t)'$. In this notation the guesses of the moving average model are given by vector $\bar{\boldsymbol{z}}_t = \sum_{s=1}^{t} \boldsymbol{z}^*_{t-s}/t$, where with re-indexing the average is now taken for observations from initial 0 period until time $t-1$, so that $L \to \infty$ case corresponds to $t \to \infty$. We then obtain the following recursive relation

$$\bar{\boldsymbol{z}}_t = \bar{\boldsymbol{z}}_{t-1} + \frac{1}{t}(\boldsymbol{z}^*_{t-1} - \bar{\boldsymbol{z}}_{t-1}) = \bar{\boldsymbol{z}}_{t-1} + \frac{1}{t}\big((\mathbf{M} - \mathbf{I})\, \bar{\boldsymbol{z}}_{t-1} + \boldsymbol{d}\big)\,.$$

Asymptotically, when $t \to \infty$, this system can be approximated by a dynamics in a notional time of a continuous linear system of the ODE (see Ljung, 1977 and Evans and Honkapohja, 2001 for technical details):

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\bar{\boldsymbol{z}}_\tau = \boldsymbol{F}(\bar{\boldsymbol{z}}_\tau) = (\mathbf{M} - \mathbf{I})\, \bar{\boldsymbol{z}}_\tau + \boldsymbol{d}\,.$$

The local stability conditions of this system can be analyzed using the Jacobian matrix of map $\boldsymbol{F}$ at the fixed point (which coincide with the REE). This matrix is $\mathbf{M} - \mathbf{I}$

[62] As the dynamics of the variable a is independent of b, we can determine the converging properties of this variable and then infer the properties of b from the stability conditions of the whole system.

Treatment	$L = 1$ (naïve)				$L = 2$		$L = 3$		$L = \infty$ (LS learning)			
	EVs		**Converges?**		**Converges?**		**Converges?**		**EVs of T-map**		**Converges?**	
	μ_1	μ_2	for a	for b	for a	for b	μ_1	μ_2	for a	for b		
Sink	$\frac{2}{3}$	$-\frac{1}{2}$	Yes	Yes	Yes	Yes	Yes	Yes	$-\frac{1}{3}$	$-\frac{3}{2}$	Yes	Yes
SaddleNeg	$\frac{2}{3}$	$-\frac{3}{2}$	Yes	No	Yes	Yes	Yes	Yes	$-\frac{1}{3}$	$-\frac{5}{2}$	Yes	Yes
SaddlePos	$\frac{2}{3}$	$\frac{3}{2}$	Yes	No	Yes	No	Yes	No	$-\frac{1}{3}$	$\frac{1}{2}$	Yes	No
Source	$\frac{3}{2}$	$-\frac{3}{2}$	No	No	No	No	No	No	$\frac{1}{2}$	$-\frac{5}{2}$	No	No

Table 11: Properties of the moving average model MAve(L) with different L. In the last four columns the table reports the eigenvalues of the ODE system (so-called T-map) that define the E-stability of the model when $L \to \infty$. The E-stability condition is that the eigenvalues of the T-map are negative.

whose eigenvalues are $m_{11} - 1$ and $m_{22} - 1$. The ODE system is asymptotically stable if both these values are negative. This is the condition of asymptotic convergence of the average model.

The last columns of Table 11 summarize the results of this analysis. As Table 11 indicates, the major predictions of the MAve(L) model with $L \geq 2$ are consistent with our experimental data: the model generates convergence in exactly those treatments (**Sink** and **SaddleNeg**) where convergence was observed in the experiment and divergence in the other two treatments (**SaddlePos** and **Source**).

To simulate the moving average model we initialize the targets for the first period only. Then, equation (14) is used for $t > L$, whereas when $t \leq L$, all available observations are equally weighted.[63] Figs. 19 and 20 compare the experimental data (left panels) with simulations of the moving average model for $L = 2$ and $L = 3$ (which look similar). When the dynamics are initialized using the experimental data (the right panels) there is a visible difference in the dynamics between the data and the model. Convergence in the **Sink** and **SaddleNeg** treatments occurs in the model quicker than in the experiment (at least in the first 3-5 periods) and in a more orderly

[63]In Eq. (14) the factor on the right-hand side becomes $1/(t-1)$ and the last term in the sums are from period 1.

way (e.g., without the "jump" patterns for b-guesses). Divergence in the **SaddlePos** and **Source** cases is also somewhat quicker, i.e., the dynamics reach the boundaries in fewer steps in the model as compared with the data.

Fig. 21 shows simulations of the model when at each step *all available* targets are averaged. This is, in effect, the MAve(15) model whose property can be approximated by the $L = \infty$ case. From simulations we observe that as L gets larger, the convergence path generated by the model becomes even less similar to the experimental data. The dynamics start with a larger jump than in the experiment, indicating a quicker reaction and convergence, but eventually it does not reach equilibrium in 15 steps. The convergence turns out to be very slow, and this is, because this average model has a *decreasing* gain, i.e., the new observations get lower weights as time goes on (the weights are inversely proportional to the number of past periods). This delays the incorporation of new information.

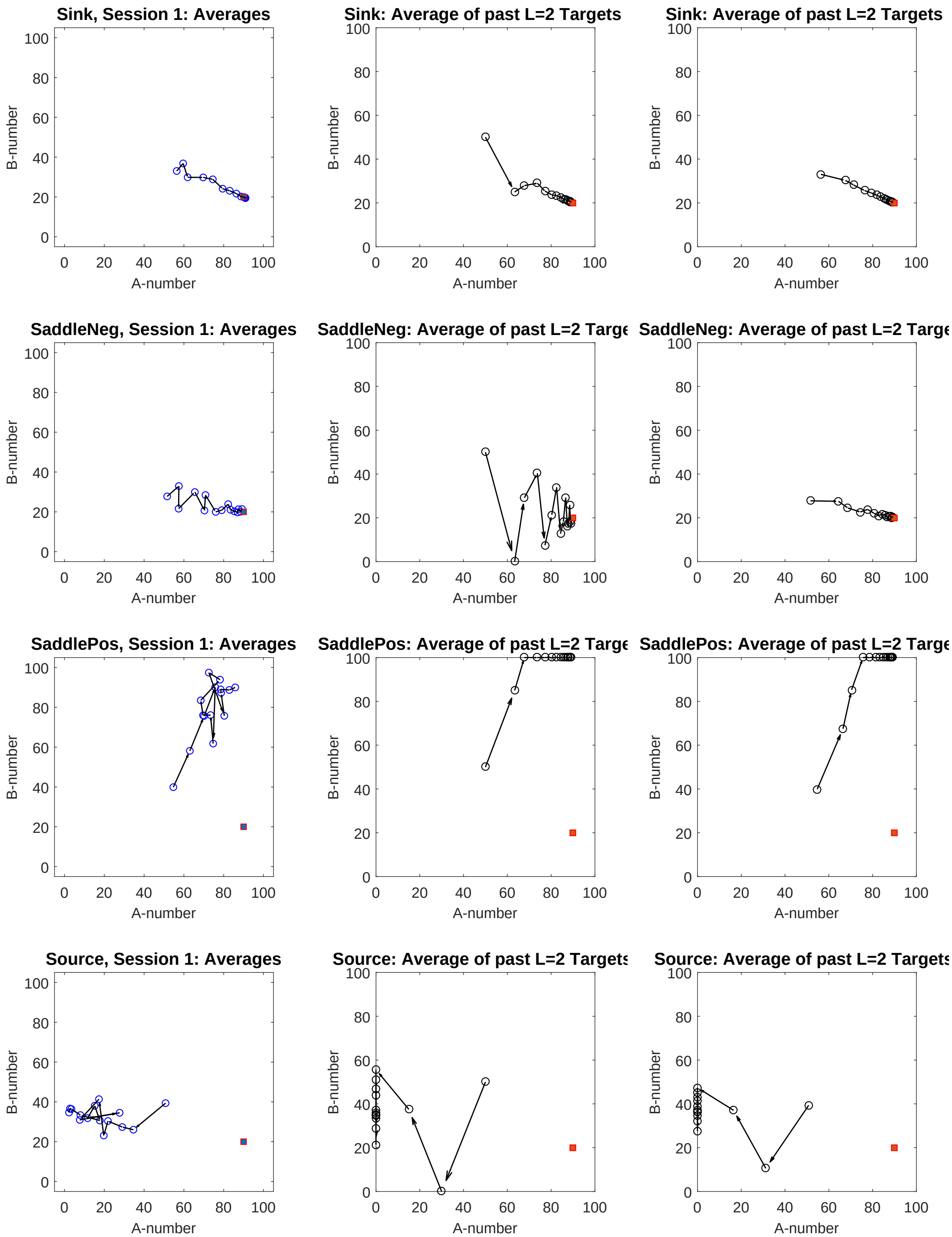


Figure 19: Dynamics of averages in the experimental group (*left panels*) as compared with the dynamics of the MAve(2) model. First 15 periods are simulated using the moving average model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experiment (*right panels*).

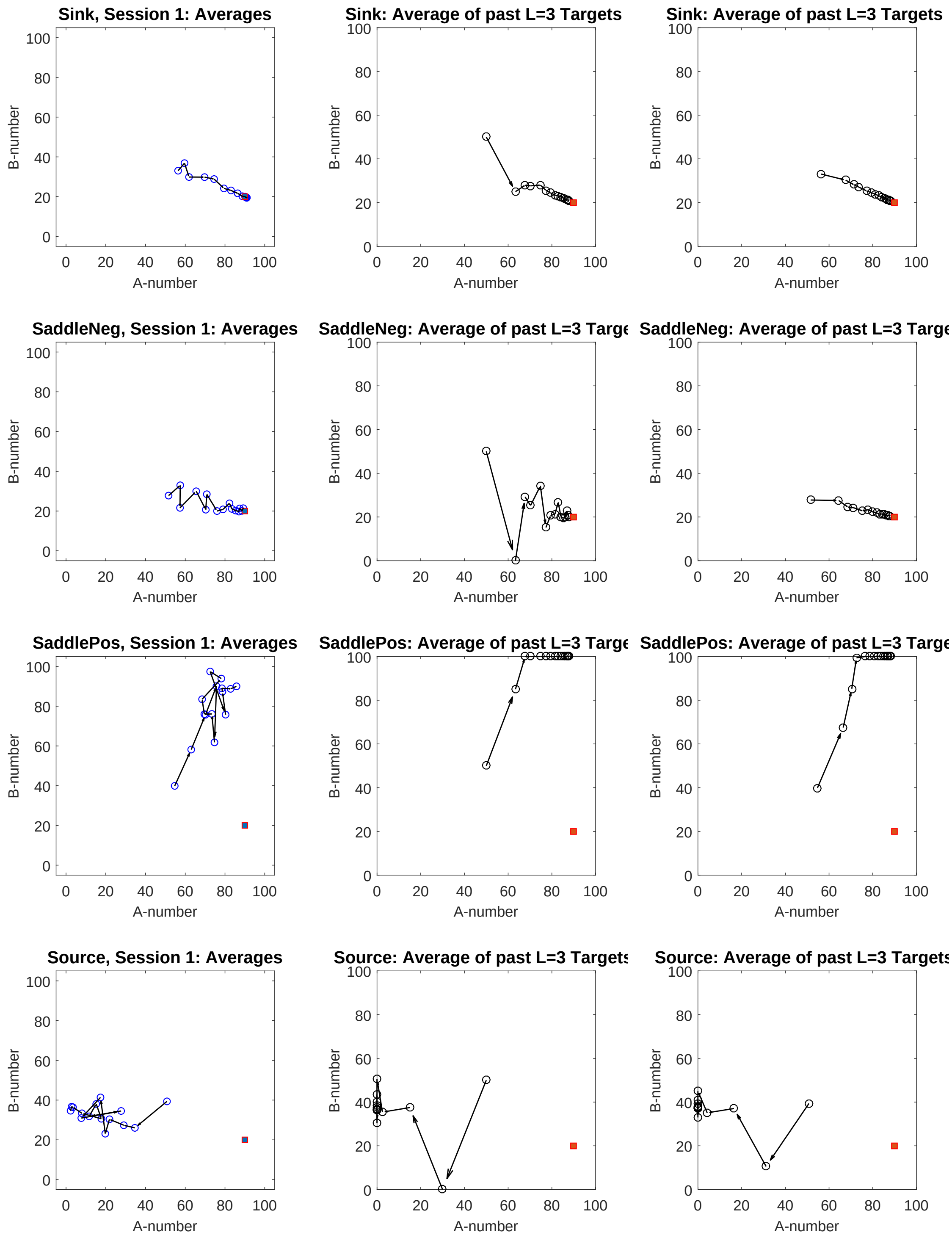


Figure 20: Dynamics of averages in the experiment (*left panels*) as compared with the dynamics of the MAve(3) model. First 15 periods are simulated using the averaged model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experiment (*right panels*).

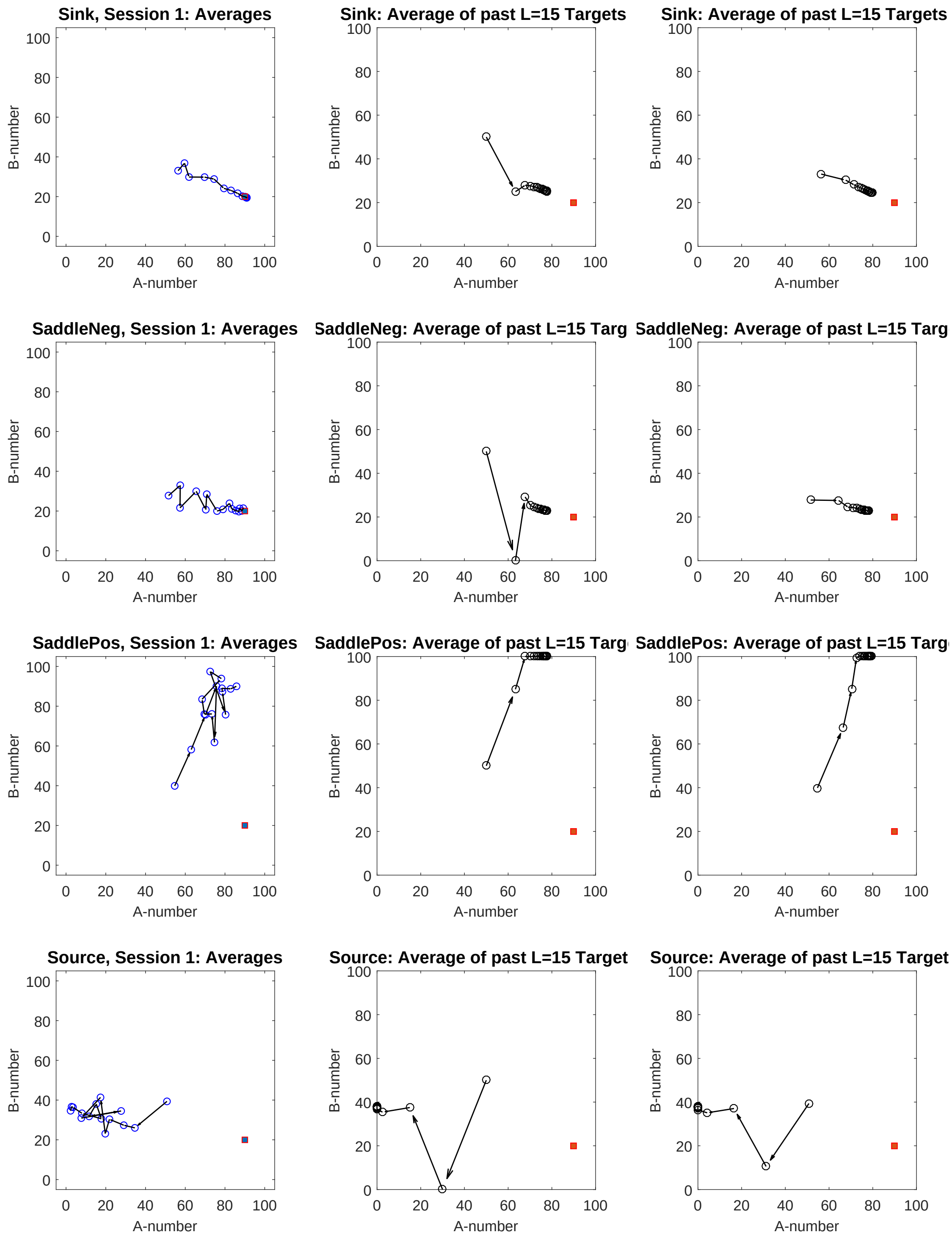


Figure 21: Dynamics of averages in the experiment (*left panels*) as compared with the dynamics of the **AveragedAll** model with $L = 15$. First 15 periods are simulated using the averaged model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experiment (*right panels*).

I Online Appendix (Not for publication). Robustness checks

We provide the results of the robustness check for the econometric estimations in Section 5. This check addresses an issue of possible heterogeneity of model parameters across treatments, which may lead to biased estimates as explained in Wilcox (2006). Table 12 shows the results of the estimation of the models for each treatment separately. The last column computes the average over four treatments. The results of the convergent **Sink** and **SaddleNeg** treatments confirm the ranking of learning models, while the overall ranking slightly favors a similar mixed model with the addition of level 2.

Models		MSEs				
		Sink	**SaddleNeg**	**SaddlePos**	**Source**	**Overall**
REE		221.16	97.53	4559.81	5685.28	2640.94
Naïve		44.03	25.00	212.42	318.09	149.88
Average	MAve(2)	35.76	23.15	199.29	269.72	131.98
	MAve(3)	36.57	24.95	187.91	263.17	128.15
	MAve(4)	36.30	29.37	183.30	252.84	125.45
	MAve(5)	39.02	34.04	182.05	255.46	127.64
	EWMA	39.38	18.70	214.12	261.44	133.41
Level-k	0 (Stubborn)	49.02	59.20	138.00	210.57	114.20
	1 (Naïve)	44.03	25.00	212.42	318.09	149.88
	2	89.36	153.34	397.90	683.08	330.92
Adaptive	$\lambda_a = \lambda_b$	**33.25**	**8.35**	87.29	115.07	60.99
	$\lambda_a \neq \lambda_b$	33.47	9.68	86.58	**112.61**	60.59
Mixed	0 1 (Ada)	**33.25**	**8.35**	87.29	115.07	60.99
	0 1 2	33.25	9.30	**82.46**	115.07	**60.02**
	1 2	44.03	15.92	212.42	270.16	135.63
CH-Poisson	$\max k = 1$ (Ada)	**33.25**	**8.35**	87.29	115.07	60.99
	$\max k = 2$	33.37	11.70	85.85	117.52	62.11
	$\max k = 3$	33.39	12.30	85.71	117.76	62.29

Table 12: Estimation and performance of learning models in terms of the out-of-sample one-step-ahead MSE (for periods 2 to 15) using leave-one-out procedure when models' parameters are estimated separately for each treatment.

References

Adam, Klaus, "Experimental Evidence on the Persistence of Output and Inflation," *The Economic Journal*, 2007, *117* (520), 603–636.

Anufriev, Mikhail and Cars Hommes, "Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments," *American Economic Journal: Microeconomics*, 2012, *4* (4), 35 – 64.

_ , _ , and Tomasz Makarewicz, "Simple Forecasting Heuristics that Make us Smart: Evidence from Different Market Experiments," *Journal of European Economic Association*, 2019, *17*, 1538–1584.

Assenza, Tiziana, Peter Heemeijer, Cars Hommes, and Domenico Massaro, "Managing Self-organization of Expectations through Monetary Policy: a Macro Experiment," *Journal of Monetary Economics*, 2020. Forthcoming.

Bao, Te and John Duffy, "Adaptive versus eductive learning: Theory and evidence," *European Economic Review*, 2016, *83*, 64–89.

Benhabib, Jess, John Duffy, and Rosemarie Nagel, "(De)-anchoring beliefs in beauty contest games," Working paper 2019.

Blanchard, Olivier Jean and Charles M Kahn, "The solution of linear difference models under rational expectations," *Econometrica*, 1980, pp. 1305–1311.

Bulow, Jeremy I, John D Geanakoplos, and Paul D Klemperer, "Multimarket oligopoly: Strategic substitutes and complements," *Journal of Political economy*, 1985, *93* (3), 488–511.

Camerer, Colin F, *Behavioral game theory: Experiments in strategic interaction*, Russell Sage Foundation, 2003.

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

Chong, Juin-Kuan, Teck-Hua Ho, and Colin F Camerer, "A generalized cognitive hierarchy model of games," *Games and Economic Behavior*, 2016, *99*, 257–274.

Costa-Gomes, Miguel and Vincent Crawford, "Cognition and behavior in two-person guessing games: An experimental study," *American Economic Review*, 2006, *96* (10), 1737–1768.

_ , _ , and Bruno Broseta, "Cognition and behavior in normal-form games: An experimental study," *Econometrica*, 2001, *69* (5), 1193–1235.

Crawford, Vincent, Miguel Costa-Gomes, and Nagore Iriberri, "Boundedly rational versus optimization-based models of strategic thinking and learning in games," *Journal of Economic Literature*, 2013, *51*, 512–527.

Duffy, John and Rosemarie Nagel, "On the robustness of behaviour in experimental beauty contest games," *The Economic Journal*, 1997, *107* (445), 1684–1700.

Ellison, Martin and Joe Pearlman, "Saddlepath learning," *Journal of Economic Theory*, 2011, *146* (4), 1500–1519.

Evans, George W and Seppo Honkapohja, *Learning and Expectations in Macroeconomics*, Princeton University Press, 2001.

Fehr, Ernst and Jean-Robert Tyran, "Limited rationality and strategic interaction: the impact of the strategic environment on nominal inertia," *Econometrica*, 2008, *76* (2), 353–394.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental economics*, 2007, *10* (2), 171–178.

Fudenberg, Drew and David K Levine, *The Theory of Learning in Games*, MIT press, 1998.

Galí, Jordi, *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press, 2015.

Güth, Werner, Martin Kocher, and Matthias Sutter, "Experimental 'beauty contests' with homogeneous and heterogeneous players and with interior and boundary equilibria," *Economics Letters*, 2002, *74* (2), 219–228.

Haltiwanger, John and Michael Waldman, "Rational expectations and the limits of rationality: An analysis of heterogeneity," *American Economic Review*, 1985, *75* (3), 326–340.

Hansen, Bruce, *Econometrics*, University of Wisconsin, 2019.

Heemeijer, Peter, Cars Hommes, Joep Sonnemans, and Jan Tuinstra, "Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation," *Journal of Economic Dynamics and Control*, 2009, *33* (5), 1052–1072.

Ho, Teck-Hua, Colin Camerer, and Keith Weigelt, "Iterated dominance and iterated best response in experimental 'p-beauty contests'," *The American Economic Review*, 1998, *88* (4), 947–969.

Hommes, Cars, "Dynamics of the cobweb model with adaptive expectations and nonlinear supply and demand," *Journal of Economic Behavior & Organization*, 1994, *24* (3), 315–335.

_ , "The heterogeneous expectations hypothesis: Some evidence from the lab," *Journal of Economic Dynamics and Control*, 2011, *35* (1), 1–24.

_ , *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*, Cambridge University Press, 2013.

Lambsdorff, Johann Graf, Manuel Schubert, and Marcus Giamattei, "On the role of heuristics – Experimental evidence on inflation dynamics," *Journal of Economic Dynamics and Control*, 2013, *37* (6), 1213–1229.

Ljung, Lennart, "Analysis of recursive stochastic algorithms," *IEEE transactions on automatic control*, 1977, *22* (4), 551–575.

Mauersberger, Felix and Rosemarie Nagel, "Levels of reasoning in Keynesian beauty contests: A generative framework," in Cars Hommes and Blake LeBaron, eds., *Handbook of Computational Economics*, Vol. 4, North-Holland, 2018, pp. 541–634.

Nagel, Rosemarie, "Unraveling in guessing games: An experimental study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Nerlove, Marc, "Adaptive expectations and cobweb phenomena," *The Quarterly Journal of Economics*, 1958, *72* (2), 227–240.

Pfajfar, Damjan and Blaž Žakelj, "Inflation expectations and monetary policy design: Evidence from the laboratory," *Macroeconomic Dynamics*, 2016, pp. 1–41.

Sargent, Thomas J, *Bounded Rationality in Macroeconomics*, Oxford University Press, 1993.

Sonnemans, Joep and Jan Tuinstra, "Positive expectations feedback experiments and number guessing games as models of financial markets," *Journal of Economic Psychology*, 2010, *31* (6), 964–984.

Stahl, Dale and Paul Wilson, "Experimental evidence on players' models of other players," *Journal of Economic Behavior & Organization*, 1994, *25* (3), 309–327.

_ **and** _ , "On players' models of other players: Theory and experimental evidence," *Games and Economic Behavior*, 1995, *10* (1), 218–254.

Stone, Mervyn, "Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society: Series B (Methodological)*, 1974, *36* (2), 111–133.

_ , "An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion," *Journal of the Royal Statistical Society: Series B (Methodological)*, 1977, *39* (1), 44–47.

Sutan, Angela and Marc Willinger, "Guessing with negative feedback: An experiment," *Journal of Economic Dynamics and Control*, 2009, *33* (5), 1123–1133.

Wilcox, Nathaniel T, "Theories of learning in games and heterogeneity bias," *Econometrica*, 2006, *74* (5), 1271–1292.