UTS UNIVERSITY OF TECHNOLOGY SYDNEY

UTS Business School –
Department of Economics
Dr Chau Chak Wing Building
14 – 28 Ultimo Road, Ultimo NSW 2007

T: +61 2 9514 7713
economics@uts.edu.au

Department of Economics
Working Paper Series

'Planar Beauty Contests'

Mikhail Anufriev [1]
John Duffy [2]
Valentyn Panchenko [3]

[1] University of Technology Sydney, Australia
[2] University of California, United States of America
[3] UNSW Business School, Sydney, Australia

No. 10 / 2019 **June 2019**

Planar Beauty Contests*

Mikhail Anufriev† John Duffy‡ Valentyn Panchenko§

June 12, 2019

Abstract

We introduce a planar beauty contest game where agents must simultaneously guess two, endogenously determined variables, a and b. The system of equations determining the actual values of a and b is a *coupled* system; while the realization of a depends on the average forecast for a, $\overline{a}$, as in a standard beauty contest game, the realization of b depends on both $\overline{a}$ and on the average forecast for b, $\overline{b}$. Our aim is to better understand the conditions under which agents learn the steady state of such systems and whether the eigenvalues of the system matter for the convergence or divergence of this learning process. We find that agents are able to learn the steady state of the system when the eigenvalues are both less than 1 in absolute value (the sink property) or when the steady state is saddlepath stable with the one root outside the unit circle being negative. By contrast, when the steady state exhibits the source property (two unstable roots) or is saddlepath stable with the one root outside the unit circle being positive, subjects are unable to learn the steady state of the system. We show that these results can be explained by either an adaptive learning model or a mixed cognitive levels model, while other approaches, e.g., naïve or homogeneous level-k learning, do not consistently predict whether subjects converge to or diverge away from the steady state.

*We thank Peter Bossaerts, Marcus Giamattei, Frank Heinemann, Cars Hommes, Rosemarie Nagel and Jan Tuinstra for useful discussions. The paper benefited from comments of the participants of the 23rd CEF Conference in Milan, the 2018 ESA meeting, Berlin, the Behavioral and Experimental workshop in Nanyang Technical University, Singapore, BEAM workshop, Amsterdam, and seminars at the European University at St. Petersburg, Monash University and Macquarie University. 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

Keywords: Beauty Contest; Learning; Stability; Simultaneous Equation Systems, Level-k cognitive theory.
JEL Classification: C30, C92, D83, D84.

1 Introduction

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. In the standard version of this game, each player i in a group of N subjects is asked to guess some number, $a_i \in [0, 100]$. Each player i's payoff depends on how close their guess is to some target number, $a^* = p \times \bar{a}$, where $\bar{a} = \sum_{i=1}^{N} a_i/N$ denotes the average guess, and where N and $0 < p < 1$ are known. Since $p < 1$, the unique, dominance solvable Nash equilibrium prediction is for all to guess 0, but this outcome is rarely observed. As Nagel (1995) and Stahl and Wilson (1994, 1995) showed, subjects are heterogeneous with respect to the number of steps of iterated elimination of dominated strategies that they apply to this problem. 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 $p^k \times a_0$. 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 compute the Nash equilibrium and submit a guess of 0.

In this paper we extend the Keynesian beauty contest to a second dimension. That is, we introduce a "planar beauty contest" game where individuals are asked to form guesses about two variables, and where both guesses may matter for the realization of the target values subjects are incentivized to achieve. Thus, the dynamical system describing how target numbers depend on the guesses is *coupled*. In pursuing this exercise we have several aims. First, we seek to understand whether level-k reasoning is a robust characterization of learning behavior when subjects face a two-dimensional coupled dynamical system and may therefore have to expend more thought on the proper solution to that system. In particular, the coupled nature of the problem may cause subjects to think harder about finding the fixed point. Second, such dynamical systems represent a workhorse, reduced-form framework for many models used by economists (we provide two examples). Our planar beauty contest game approach provides a similar reduced-form framework for understanding whether individuals can learn the equilibria of such models that stands in contrast to the "learning-to-forecast" experimental approach (Hommes, 2011). In the latter approach, agents are asked to form forecasts in multivariate systems which are then fed into a dynamic

general equilibrium model to produce realizations of the variables that subjects were forecasting. However, in the learning-to-forecast approach subjects typically have little or no knowledge as to how their forecast matters for the realization of those variables, whereas in our approach, this process is clearly spelled out; the laws of motion of the two variables of interest are perfectly known. Third, we are interested in whether the eigenvalues of the two-dimensional system matter for whether subjects achieve the steady state solution of the system, as well as the speed at which such convergence may obtain. Finally, we are interested in exploring which, among several different learning models found in the literature, can best explain our results. For this purpose, we consider naïve learning, homogeneous level-k learning, adaptive learning and mixed cognitive level learning models.

We consider four linear, planar environments, which comprise the four treatments of our experiment. In all four environments, there is a unique interior Nash equilibrium of the planar system (a^{NE}, b^{NE}) that is held constant; the only changes made are to the equations mapping individual guesses to the target values. Assuming a simple naïve learning process under repeated play of the game, the equations govern the evolution of the a and b variables over time with the interior Nash equilibrium of the game corresponding to the steady state of that dynamical system. In the first environment, both eigenvalues of the planar system are less than 1 in absolute value so that the solution is, dynamically speaking, a "sink" with convergent paths from any initial condition. In the second and third environments, one of the eigenvalues is positive and less than 1 while the other eigenvalue is greater than 1 in absolute value so that the solution displays the so called "saddlepath property." Such dynamical systems are unstable; while there exists a unique converging path (i.e., a one-dimensional set of initial conditions that would converge to the steady state), there exist infinitely many other solution paths leading away from the steady state. In the fourth environment, both of the eigenvalues are greater than 1 in absolute value so that the system is dynamically unstable – a "source" – with no converging paths.

Given these dynamical properties, we expected to observe convergence in the sink environment and non-convergence in the source environment. The saddlepath environment is particularly attractive to macroeconomists, 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. Within the saddlepath case, we consider two different environments. In the "saddlepath negative" case the unstable eigenvalue (i.e., the eigenvalue greater than 1 in absolute value) is negative, and in the "saddlepath positive case" the unstable eigenvalue is positive. As discussed below, there is some evidence that the sign of the eigenvalue in stable *univariate* systems matters for convergence to the steady state of such systems and we wanted to consider whether this property also extends to *multivariate* (planar) systems.

To preview our results, we find that groups of 10 subjects with complete information regarding the planar data generating process are able to learn the steady state of such systems when the interior steady state exhibits the sink property or is saddlepath negative. By contrast, when the steady state exhibits the source property or is saddlepath positive, subjects are unable to coordinate on the interior steady state solution of the system, and instead coordinate on a boundary solution to the dynamical system; absent this boundary, they would simply diverge away form the interior 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, but that such convergence is not a *general* property of saddlepath stable solutions. This finding furthers our understanding of the empirical relevance of saddlepath stable solutions in planar systems.

We also consider what types of learning models could predict the pattern of behavior that we observe across the four treatments of our experiment. 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.[1] Then we 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* to the interior steady state while the human subjects consistently converged to 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

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

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 a *hybrid* adaptive learning model, with some weight assigned to the interior steady state equilibrium, fares the best across all of the models that we consider according to the out-of-sample predictive mean squared error.

The remainder of the paper is organized as follows. Section 2 introduces the planar beauty contest game (PBC) and discusses related literature. We describe our experimental design and implementation of the PBC game in Section 3 and we qualitatively discuss experimental results in Section 4. Section 5 introduces a range of behavioral models of individual play in the game and analyzes their theoretical properties, comparing them to the experimental outcomes. In Section 6 we estimate the models using the experimental data and identify the model that best fits the data across all four treatments. Section 7 concludes. Appendices contain the experimental instructions and additional information.

2 Planar Beauty Contest Game

The basic framework we study is a two-dimensional, coupled, self-referential affine system where expectations of the two endogenous variables, a and b, matter for the realizations of those variables in each period $t = 1, 2, \ldots, T$. Specifically, we have in mind linear (or linearized) economic models that can be put into the following 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 value of the variables a_t and b_t using all information available through time $t-1$ and the values of matrix

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \quad \text{and vector} \quad \boldsymbol{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

are known to all agents. We implement this framework in the lab by asking 10 participants to play the planar beauty contest game for 15 consecutive periods. In

this game, which is presented in Section 3, individuals are motivated to guess the values a_t and b_t determined as in (1) with the averages of the guesses used in place of expectations.

Denote the 2×2 identity matrix by $\mathbf{I}$ and assume that $\mathbf{I} - \mathbf{M}$ is invertible. Under rational expectations, where $a_t^e = a_t$ and $b_t^e = b_t$, system (1) has a unique steady state given by $(\mathbf{I} - \mathbf{M})^{-1}\boldsymbol{d}$. This steady state corresponds to the unique interior Nash equilibrium of the planar beauty contest game.

There are a number of economic models that map into this basic framework. In Appendix A we provide two examples, one from the industrial organization literature and another from the macroeconomics literature on monetary policy. Unlike the univariate case, in this coupled, two-dimensional planar system, expectations about the realization of one variable can matter for realizations of the other variable, and vice versa. A consequence of this *interdependence* is that the eigenvalues of the matrix, $\mathbf{M}$, can matter for the stability of the steady state of the system, and these eigenvalue conditions are a main focus of our analysis.

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 the setup that we study, that is, our basic framework is a static one. However, *learning agents* who do not (or cannot) immediately solve for the steady state of the system, will likely condition their expectations on the past history of realizations, $\mathcal{H}_t = \{a_s, b_s\}_{s=1}^{t-1}$ in forming expectations for the time t values for a_t and b_t. For instance, if agents have naïve expectations, i.e., if $a_t^e = a_{t-1}$ and $b_t^e = b_{t-1}$, then system (1) is effectively a two-dimensional, coupled *first-order dynamical* system of the form:

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} a_{t-1} \\ b_{t-1} \end{pmatrix} + \boldsymbol{d}.$$

Of course, agents may use a variety of different learning rules and we consider the fit of a number of such rules in our analysis of the experimental data. We wish to emphasize that the interdependencies of expectation formation in the two-dimensional system that we study makes for a more challenging test of learning models which have typically been used to examine behavior in univariate or un-coupled systems.[2]

[2]For examples of research where experimental data are used to evaluate and compare various learning models in the univariate framework see Anufriev and Hommes (2012) and Anufriev et al. (2019).

2.1 Related literature

The original beauty contest game, as first explored experimentally by Nagel (1995) and Stahl and Wilson (1995), can be viewed as a single equation system. As noted earlier, the aim of the game is for each player $i = 1, \ldots, N$ 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 $[0, 100]$. In equilibrium, the solution, a^{NE}, must be the same for all players, namely, $a_i = a^{NE} = pa^{NE}$, so that the steady state prediction is that $a_i = 0$ for all i.

Subsequent experimental research by Güth et al. (2002), Sutan and Willinger (2009) examines affine univariate systems where the target value, $a^* = p\overline{a} + d$, and $|p| < 1$. This setup yields a non-corner or "interior" equilibrium solution, with negative or positive feedback depending on the sign of p. The positive feedback version has $0 < p < 1$ and monotonic convergence to the equilibrium steady state, $a^{NE} = d/(1-p)$, while the negative feedback version has $-1 < p < 0$ and oscillatory convergence to the same steady state value. Positive (negative) feedback systems are related to strategic substitutes (complements) as first noted by Haltiwanger and Waldman (1985). Experimentally, 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), Sutan and Willinger (2009), 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 stable 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.[3]

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. (2014), Pfajfar and Žakelj (2016)

[3]More recently, Benhabib et al. (2019) study affine univariate systems where the target value $a^* = p\overline{a} + d_i$, 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.

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 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., 2014, Pfajfar and Žakelj, 2016). Our paper is related to this work in that we consider how the eigenvalues of the system matter for whether agents can learn the steady state. However, we are not interested in considering a specific model (e.g., the New Keynesian model); rather we are interested in how the properties of planar dynamical systems impact on the learning of equilibrium. Toward that end, we do not have an explicitly dynamical framework, though as noted above, if agents are adaptive learners, that adaptation process makes our system dynamical. Unlike 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 beauty contest experimental literature.[4] 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 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 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.

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

3 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, b_i). On the basis of these "guesses", the group averages

$$\bar{a} = \frac{1}{N}\sum_{i=1}^{N} a_i \quad \text{and} \quad \bar{b} = \frac{1}{N}\sum_{i=1}^{N} b_i$$

are computed. Finally, the "target" values a^* and b^* are defined as

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

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

$$\pi_i = \frac{500}{5 + |a_i - a^*| + |b_i - b^*|} \tag{3}$$

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.[5]

Assuming that the matrix $\mathbf{I} - \mathbf{M}$ is invertible (which will be the case in all of our treatments), the one-shot game described above has a unique Nash equilibrium. In this equilibrium every participant submits the guesses given by

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

Indeed, it can be directly checked that with this profile, both targets a^* and b^* coincide with the corresponding guesses, leading to the maximum possible payoff for every participant i.

In this paper we study the behavior of experimental subjects in 4 treatments with

[5]The hyperbolic function in (3) penalizes strongly the smallest deviations from the target, giving participants stronger incentives to guess the targets precisely. This hyperbolic payoff function has been used in Adam (2007) and Assenza et al. (2014) for one dimensional decisions.

different matrices $\mathbf{M}$ and vectors $\boldsymbol{d}$. We construct all treatments so that the interior Nash equilibrium (4) is the same across treatments, with $a^{NE} = 90$, $b^{NE} = 20$. As is typical in beauty contest experiments, we restrict the range of guesses submitted by the participants. That is we inform participants that both a_i and b_i must belong to the interval $[0, 100]$.[6]

We are particularly interested in how changes to the matrix $\mathbf{M}$ affect the outcome of play. Our four treatments differ in the location of the *eigenvalues* of $\mathbf{M}$. In all of our treatments, matrix $\mathbf{M}$ is lower triangular, with $m_{12} = 0$ in (2), and thus the eigenvalues of $\mathbf{M}$ are the diagonal elements. Further, in three of our four treatments, we set $m_{11} = 2/3$; this parameter choice is one that is commonly used in the univariate beauty contest game; indeed the first equation of our coupled system can be thought of as a version of the classic, 2/3 of the average, beauty contest game, (albeit with an interior solution).[7] We now describe our four treatments.

Sink. In this treatment

$$\mathbf{M} = \begin{pmatrix} 2/3 & 0 \\ -1/2 & -1/2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{d} = \begin{pmatrix} 30 \\ 75 \end{pmatrix}.$$

Both eigenvalues of $\mathbf{M}$, 2/3 and $-1/2$, are inside the unit circle. The guessing game with this matrix and restriction of the strategy space to $[0, 100] \times [0, 100]$ has a unique Nash equilibrium of $(90, 20)$.

SaddleNeg. In this "saddle with negative feedback" treatment

$$\mathbf{M} = \begin{pmatrix} 2/3 & 0 \\ -1/2 & -3/2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{d} = \begin{pmatrix} 30 \\ 95 \end{pmatrix}.$$

The matrix $\mathbf{M}$ has eigenvalues 2/3 and $-3/2$. This is the saddle case for dynamical systems, because one of the eigenvalues is inside and another is outside of the unit circle. The unstable eigenvalue is negative and it introduces strategic substitutability

[6] Depending on the treatment, this restriction may lead to other, non-internal, or boundary equilibria of the game, as explained below; see Appendix B for formal proofs. However, in all cases the internal equilibrium $(90, 20)$ will be the payoff-dominating equilibrium. It is also the only rational expectations equilibrium (REE), when we use the game as a representation of model (1).

[7] We chose a lower triangular matrix for $\mathbf{M}$ to make our system coupled, but, at the same time, simple enough that subjects could possibly solve for the steady state.

or "negative feedback" to the b-number strategy: higher guesses imply a lower target. The guessing game with this matrix and the strategy space $[0, 100] \times [0, 100]$ has a unique Nash equilibrium of $(90, 20)$ as in the **Sink** treatment.

SaddlePos. In this "saddle with positive feedback" treatment

$$\mathbf{M} = \begin{pmatrix} 2/3 & 0 \\ -1/2 & 3/2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{d} = \begin{pmatrix} 30 \\ 35 \end{pmatrix}.$$

The matrix $\mathbf{M}$ has eigenvalues $2/3$ and $3/2$. As in the **SaddleNeg** treatment, the first eigenvalue is inside and the second is outside of the unit circle. The eigenvalues in these two saddle treatments have the same absolute values. However, now the unstable eigenvalue is positive leading to strategic complementarity or "positive feedback" to the b-number strategy: higher guesses imply a larger target. The guessing game with this matrix and the strategy space $[0, 100] \times [0, 100]$ has an internal Nash equilibrium of $(90, 20)$ as well as two other boundary Nash equilibria, $(90, 0)$ and $(90, 100)$. We stress that $(90, 20)$ is the only REE, and therefore it payoff-dominates other equilibria.

Source. In this treatment

$$\mathbf{M} = \begin{pmatrix} 3/2 & 0 \\ -1/2 & -3/2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{d} = \begin{pmatrix} -45 \\ 95 \end{pmatrix}.$$

Both eigenvalues of $\mathbf{M}$, $3/2$ and $-3/2$, are outside of the unit circle. Apart from the unique REE $(90, 20)$ the guessing game on the strategy space $[0, 100] \times [0, 100]$ has two other boundary Nash equilibria: $(0, 38)$ and $(100, 18)$.

In all sessions we asked a group of $N = 10$ participants to play the same game for $T = 15$ consecutive periods. In each period, participants were incentivized to independently choose their two numbers to be as close as possible to the target values via the payoff function (3). We will use the subscript t for the time period, and so we denote the submitted guesses of player i in period t by $a_{i,t}$ and $b_{i,t}$ and the realized period target values as a_t^* and b_t^*. Thus in the experiment, the target values

for periods $t = 1, \ldots, 15$ were determined as

$$\begin{pmatrix} a_t^* \\ b_t^* \end{pmatrix} = \mathbf{M} \begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} + \boldsymbol{d}\,, \tag{5}$$

where $\bar{a}_t = \frac{1}{10}\sum_{i=1}^{10} a_{i,t}$, $\bar{b}_t = \frac{1}{10}\sum_{i=1}^{10} b_{i,t}$ and where matrix $\mathbf{M}$ and vector $\boldsymbol{d}$ were held fixed over all 15 periods of the treatment and were known to all 10 participants.

The experimental sessions were conducted in the Experimental Social Science Laboratory at the University of California, Irvine. We conducted four sessions each of the four treatments. Since each session involved 10 inexperienced subjects, 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 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 on the basis of 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 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). Each period ended with a second results screen reminding subjects of their submitted pair of numbers, and informing them of the group average values for the two numbers, $\bar{a}$ and $\bar{b}$, the two target numbers (based on the period averages), a^* and b^*, and the points that the subject earned for the period. Except for the very first period, subjects could see a history of

Treatment	Parameters	REE	Other 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: Characteristics of the Four Experimental Treatments.

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.

Table 1 provides a summary of the four treatment conditions.

4 Results

In this section we have a first look at the data and analyze them qualitatively. We begin in Section 4.1 by discussing individual choices in the first period. Then in Section 4.2 we look at the dynamics of the choices over all 15 periods of the experiment.

4.1 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 decoupled from the second, guessing the a-number is exactly equivalent to playing the standard game with the target given by $m_{11}\bar{a} + d_1$. Moreover three of our treatments

(except for the **Source**) have an identical equation for the a-number, with target given by $(2/3)\bar{a} + 30$. In the **Source** treatment, the target for the a-number is $(3/2)\bar{a} - 45$. Figure 1 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. 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**. Only a few participants submitted the interior Nash Equilibrium quantity for the a-number, 90, 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 $(0 + 100)/2 = 50$.[8] Subjects of level-1 best respond to this level-0 behavior and submit $m_{11}50 + d_1$. 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 1 four vertical lines, corresponding to the levels of 0, 1, 2 and 3 (the legend specifies the corresponding values for these different levels). Note that in the treatments with $m_{11} = 2/3$, the sequence of levels monotonically increases and converges to $a^{NE} = 90$ in an infinite number of steps. 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.[9]

The target for the b-number is affected by the average guesses for *both* numbers, making this guessing task more difficult. The four panels in Figure 2 show histograms of b-number choices in the four different treatments, with a thick vertical line indicating the interior Nash equilibrium choice for b of 20. As in the case of the a-numbers, there is a large heterogeneity in guesses, with spikes around 50 in all treatments, and

[8] An alternative specification for level-0 types has them guessing $d + m_{11} \times 100$, the upper bound of the guessing interval. The level-0 type is more often thought of as an agent's model of others, rather than the guess made by actual person.

[9] 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 $a_i = 0$.

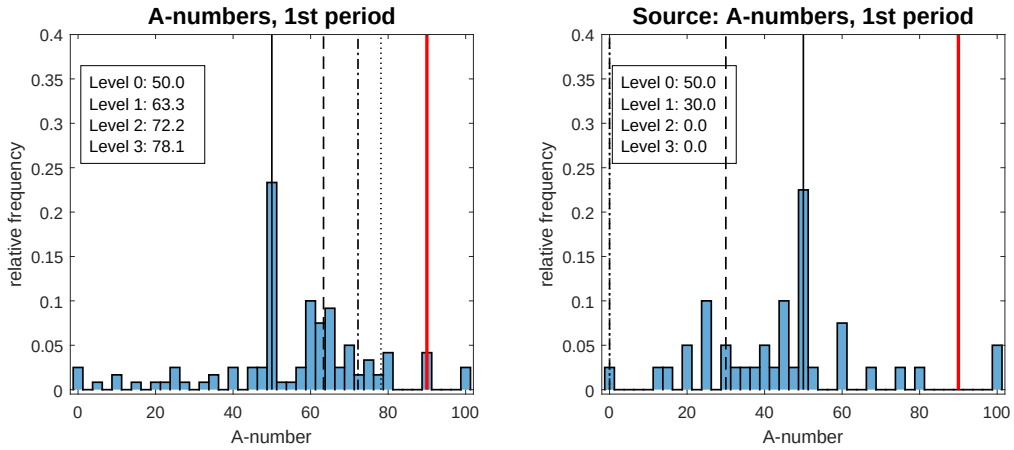


Figure 1: Frequencies of Period 1 a-choices and levels of reasoning (different dashed lines) and internal Nash Equilibrium (red line). *Left:* combined choices from **Sink**, **SaddleNeg** and **SaddlePos** treatments. *Right:* **Source** treatment.

skewness of the remaining choices to the left, except for the **SaddlePos** treatment where there are also many choices are to the right of the mid-point of the guessing interval.[10]

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 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 2 (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

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

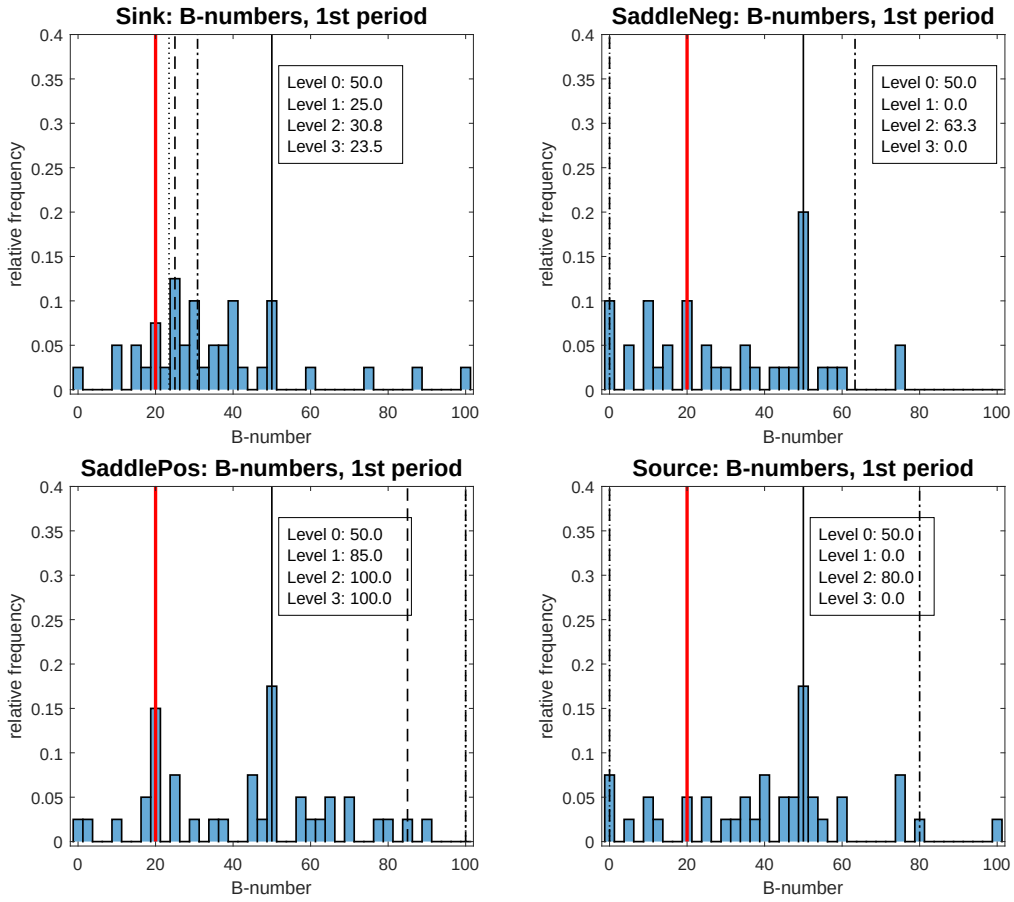


Figure 2: Frequencies of Period 1 b-choices and levels of reasoning (different dashed lines) and internal Nash Equilibrium (red line) for four treatments.

data. The level-2 choice is above 50, and the level-3 choice is 0 again, and so on.[11] 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 Beauty Contest game leads to even further decrease in the level of rationality for the b-number.[12]

The upper part of Table 2 shows the mean, median and standard deviations for choices of variables a and b in the first period. Note that an amount of heterogeneity

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

[12] 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.

Treatment	First period a-number			First period b-number		
	Mean a_1	Median a_1	Std.Dev a_1	Mean b_1	Median b_1	Std.Dev b_1
Sink	57.39	60.00	17.67	34.98	30.00	20.01
SaddleNeg	55.01	56.01	22.28	31.06	29.00	21.46
SaddlePos	56.25	58.80	20.50	42.97	46.75	23.18
Source	44.78	45.75	20.97	40.92	45.00	23.21
Treatment	**Last period a-number**			**Last period b-number**		
	Mean a_{15}	Median a_{15}	Std.Dev a_{15}	Mean b_{15}	Median b_{15}	Std.Dev b_{15}
Sink	84.21	86.85	14.40	25.40	21.48	17.44
SaddleNeg	89.81	89.92	0.55	20.31	20.00	1.49
SaddlePos	87.33	88.00	3.06	92.31	100.00	24.94
Source	7.47	0.00	13.49	35.57	35.00	6.03

Table 2: Aggregate statistics of guesses per treatment for the first and the last period of the experiment in different treatments.

(as measured by the standard deviations) is similar in all treatments. Both means and medians in all cases are substantially different from the middle point 50. Sometimes these statistics are shifted toward the internal Nash equilibrium, sometimes not, but in seven out of eight instances the direction of such shift is consistent with the level-1 choice.[13]

4.2 Dynamics

We next discuss the dynamics of guesses over all periods in the experiment. The lower part of Table 2 reports the mean, median and standard deviations for a and b choices in the last period. Comparing these with the first period choices in the top panel, we see that the heterogeneity in individual choices is reduced over time in all cases, except for the b-number in **SaddlePos**. Most importantly, with experience, the median and mean number choices for both a and b are very close to the internal Nash equilibrium in the **Sink** and **SaddleNeg** treatments and this also holds for the a-number in the **SaddlePos** treatment. In all other cases there is a substantial difference between the final period number choices and the corresponding values in the internal Nash equilibrium.

These results are confirmed when we look at the four panels of Fig. 3 showing

[13]The only exception is the first period b choices in the **SaddlePos** treatment, where the level-1 choice is 85 but most of the guesses are to the left of the middle point.

typical examples of the evolution of the averages, $\bar{a}$ and $\bar{b}$, for all four treatments, **Sink**, **SaddleNeg**, **SaddlePos**, **Source** from one representative session of each treatment (we chose session 3 of all four treatments). The time evolution of $\bar{a}$ (the thick red line) and of $\bar{b}$ (the thin blue line) are graphed in relation to the dashed lines indicating the levels of the internal Nash equilibrium, $a^{NE} = 90$ and $b^{NE} = 20$, in all four cases.[14] We observe that the dynamics of a-number converges to its equilibrium value of 90 in the first three treatments, i.e., in the **Sink** and in both **Saddle** treatments, but does not converge in the **Source** treatment. The fact that the dynamics are rather similar in the first three treatments is not surprising, because the equations governing an evolution of the a-number is the same in these treatments. The average submitted values start around a point close to 55, which results in the targeted value being close to 67. Then both averages and targets start to increase and eventually converge (from below) to the equilibrium level. The dynamics in the **Source** treatment, however, are very different. The average submitted value in the first period is again not far from the middle of the submission range, but that choice now results in the target value being close to 30. Then, both the average and target values decrease so that eventually the target value becomes negative.

The dynamics of the b-number are different in all four treatments. Importantly, both in the **Sink** and in the **SaddleNeg** treatments the dynamics converge to the equilibrium value, $b^{NE} = 20$. In the **SaddleNeg** treatment the convergence is through initial oscillations, whereas in the **Sink** treatment the convergence is rather monotone. In both cases, convergence is very quick so that the variable reaches the equilibrium value by around period 10 (later on, we address criteria for evaluating convergence). On the contrary, in the **SaddlePos** and **Source** treatments, the b-number does not converge to the internal Nash equilibrium value and displays quite peculiar dynamics. In both treatments, the initial b-number average is close to 40. Then, in the **SaddlePos** treatment, this variable quickly increases and stays close to the boundary value of 100 until the end of the experiment. In the **Source** treatment the dynamics is rather stable from the very beginning of the experiment. Moreover the average of the submitted b-numbers in this treatment is already very close to the equilibrium value of 20 in periods $t = 3$ and 5. Afterwards, however, this variable deviates and stays in an increasingly tighter range between 30 and 40 until the end of the experiment.

[14] Overall, the dynamics in the four sessions of each of the four treatments are very similar to those of session 3 as depicted in Fig. 3. See Figs. 7 to 10 in the 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.

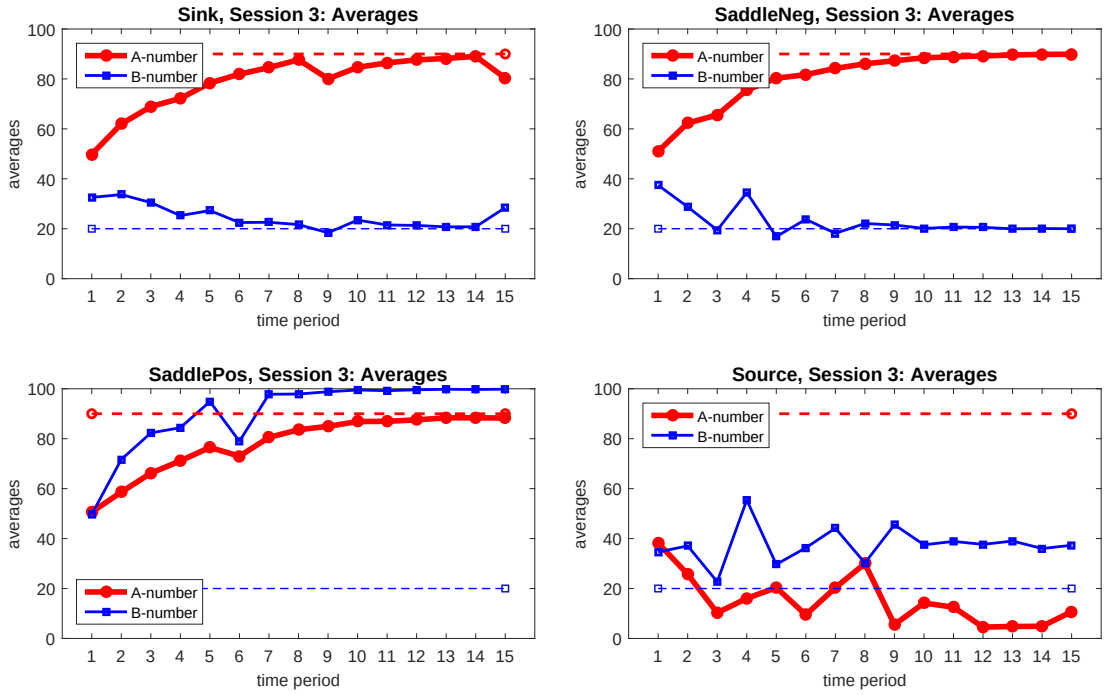


Figure 3: Typical dynamics of the average values, $\bar{a}$ and $\bar{b}$, observed in four different treatments of the experiment. The dashed lines indicate the levels of internal equilibria which are the same across all treatments.

Whereas the dynamics converged in both **Sink** and **SaddleNeg** treatments, 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 when 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 value,

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

Let $t(\varepsilon)$ denote the period when the trajectory for the average values of the a and b numbers belongs 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[15] for five different values of ε. Those cases when the trajectory never reached the neighborhood during the experiment are denoted by the hyphen.

[15]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 $\varepsilon \leq 20$.

	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	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.

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 treatment than for **Sink**. For instance, two sessions in **Sink** did not converged to the 5-neighborhood of equilibrium in 15 periods, whereas in all four sessions of the **SaddleNeg** treatment such convergence occurred (in 8 periods in average). The only session of the **Sink** treatment where dynamics converged to 0.5-neighborhood was still exhibited slower convergence than session 4 of the **SaddleNeg** treatment, where in three sessions convergence to this neighborhood occurred.[16]

We summarize the results of the experiment as follows:

1. The dynamics converge to the internal Nash equilibrium in the **Sink** and **SaddleNeg** treatments and do not converge to that equilibrium in the **SaddlePos** and **Source** treatments. Convergence is quickest in the **SaddleNeg** treatment.

2. The a-number converges to its internal Nash equilibrium value almost monotonically in the **Sink**, **SaddleNeg** and **SaddlePos** treatments. Convergence is quickest in **SaddleNeg** treatment and the slowest in **SaddlePos**.[17]

3. The b-number converges to its interior Nash equilibrium value almost monotonically in the **Sink** treatment and via an oscillatory path in the **SaddleNeg**

[16] As sometimes trajectory would exit the ε-neighborhood, we also compared the *latest* periods when the trajectory *did not* belong to the ε-neighborhood of the equilibrium. The conclusions were similar: the trajectories in the **Sink** treatment would stay around equilibrium for the shorter time than in the **SaddleNeg** treatment. See Appendix D.

[17] This conclusion follows from applying the "first hit time" convergence criterion to the trajectory of a-number only, see the details in Appendix D. Note that convergence is slowest in **SaddlePos** despite the fact that in this treatment the first period value of the a-number was the closest to the interior equilibrium value for a. As the a-variable is subject to the same one-dimensional Beauty Contest Game in all three treatments, we can conclude that the dynamics of the irrelevant b-variable affect the speed of convergence of a-number to its equilibrium value in those three treatments.

treatment.

4. The b-number does not converge to its internal equilibrium value in **SaddlePos** and **Source** treatments. In both cases, however, the dynamics are rather stable, near some constant level.[18]

The most remarkable finding of the experiment is the contrast in dynamics between two saddle treatments. Participants *are* able to learn a steady state with the saddlepath property under negative feedback, as in our **SaddleNeg** treatment, quickly converging to the internal Nash equilibrium. Participants *are not* able to learn a steady state with the saddlepath property under positive feedback, as in our **SaddlePos** treatment, converging instead to another, payoff-dominated, boundary Nash equilibrium. The outcome in these two saddle treatments are different, despite the fact that these treatments differ only in the sign of the target for b-number as dependent on the average of the b-guesses.

The notable differences in behavior between treatments translate into differences in subjects' payoffs. Table 4 shows the average payoffs per experimental session (in points).[19] We show the averages and the standard deviations of the total payoff earned over all 15 periods of experiments and compare payoffs earned in the first and last periods. As can be expected, 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 quicker convergence in the **SaddleNeg** treatment allowed subjects to earn the largest total payoff across all four treatments.

[18] 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 boundardy equilibrium $(0, 38)$. Recall that these equilibria are payoff-dominated and neither of them is the rational expectation equilibrium.

[19] 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.

Treatment	Session	Periods 1–15		Period 1		Period 15	
		Ave	Std.Dev	Ave	Std.Dev	Ave	Std.Dev
Sink	Sess. 1	923.43	95.64	23.38	12.84	92.68	4.28
	Sess. 2	549.32	171.95	16.97	12.79	35.70	13.39
	Sess. 3	734.63	111.98	23.33	14.46	41.55	13.79
	Sess. 4	758.11	173.62	24.38	17.07	62.05	16.81
	Average	741.37	138.30	22.02	14.29	58.00	12.07
SaddleNeg	Sess. 1	581.20	72.67	14.18	8.54	61.39	12.74
	Sess. 2	888.13	89.05	11.69	6.98	92.59	3.45
	Sess. 3	791.62	77.39	17.02	21.24	95.06	3.51
	Sess. 4	892.81	94.30	13.40	3.92	93.32	4.80
	Average	788.44	83.35	14.07	10.17	85.59	6.12
SaddlePos	Sess. 1	189.63	58.73	10.85	5.05	13.19	3.51
	Sess. 2	210.00	49.44	14.40	12.33	21.16	9.88
	Sess. 3	165.70	31.43	15.98	17.48	10.70	0.27
	Sess. 4	168.94	20.26	10.86	2.77	10.80	0.28
	Average	183.57	39.96	13.02	9.41	13.96	3.48
Source	Sess. 1	160.09	35.70	15.65	22.24	9.04	1.08
	Sess. 2	162.27	35.86	6.11	2.11	11.88	3.58
	Sess. 3	150.24	38.00	9.56	2.05	11.21	3.25
	Sess. 4	162.51	34.17	11.06	3.90	9.39	1.38
	Average	158.78	35.93	10.59	7.58	10.38	2.32

Table 4: Payoff statistics for each session of the corresponding treatment: average per subject and standard deviation between subjects

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 correspoding b-number guesses are shown in the bottom panels. The steady state values are indicated by the vertical red line. Notice that the variance of individual choices decreases over time, with the greatest reduction occuring 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.[20]

[20]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.

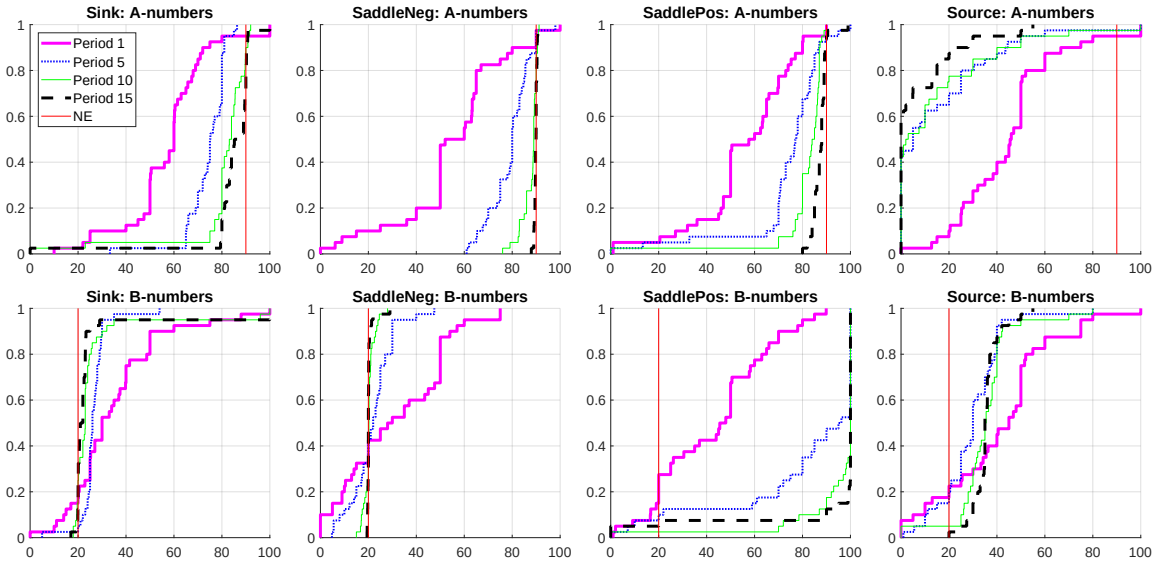


Figure 4: Cumulative frequencies of individual choices in periods 1, 5, 10 and 15 for **Sink**, **SaddleNeg**, **SaddlePos**, and **Source** treatments (from left to right) for a-number (upper panels) and b-number (lower panels). Vertical line indicates internal NE.

5 Behavioral Models

As we have seen, the ability of participants to converge to the internal Nash equilibrium and achieve the highest possible payoff depends on the treatment. In this section and the next, we address two important questions: *Can a single model of behavior explain the observed differences across our four treatments? If so, which model does provide the best fit to our data?*

In this section we will present several dynamic models and discuss their qualitative dynamics to see whether they can match the dynamics of average guesses in the experimental data. Specifically, we focus on the *simulated path* of 15 periods for each model. This path may depend on initial conditions and parameters, but it does not use any information from the experimental data.[21] 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.

[21] Later, in Section 6 we fit the models to the experimental data using quantitative measures. For that exercise 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.

5.1 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:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} \equiv \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} . \tag{6}$$

Using (5), we can write the naïve model as two-dimensional dynamical system in terms of variables $\{\bar{a}, \bar{b}\}$:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + \boldsymbol{d} . \tag{7}$$

This system has a unique steady state given by the internal Nash equilibrium of the game. The naïve model is linear and, therefore, its dynamic properties can be understood from the eigensystem of matrix $\mathbf{M}$, see Table 5. The steady state stability depends on the eigenvalues of matrix $\mathbf{M}$. 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 5 to provide more information about the dynamics. Note that in all treatments, the eigenvector $\mathbf{v}_2$ is the same, and is given by the vertical line.[22]

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, intialized at any feasible point, will only converge (is globally stable) when both eigenvalues are less than 1 in absolute value. Generally, the speed of

[22] 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).

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 5: 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 (8).

convergence is determined by the slowest dimension, i.e., by the largest (in absolute value) eigenvalue of the system.

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 (7) 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.[23] To illustrate how qualitatively 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 5 we observe that in the **Sink** treatment, both eigenvalues are inside the unit circle and therefore the dynamics is 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. Convergence of the model is slower in the direction of a-number: this is because the

[23] In 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.

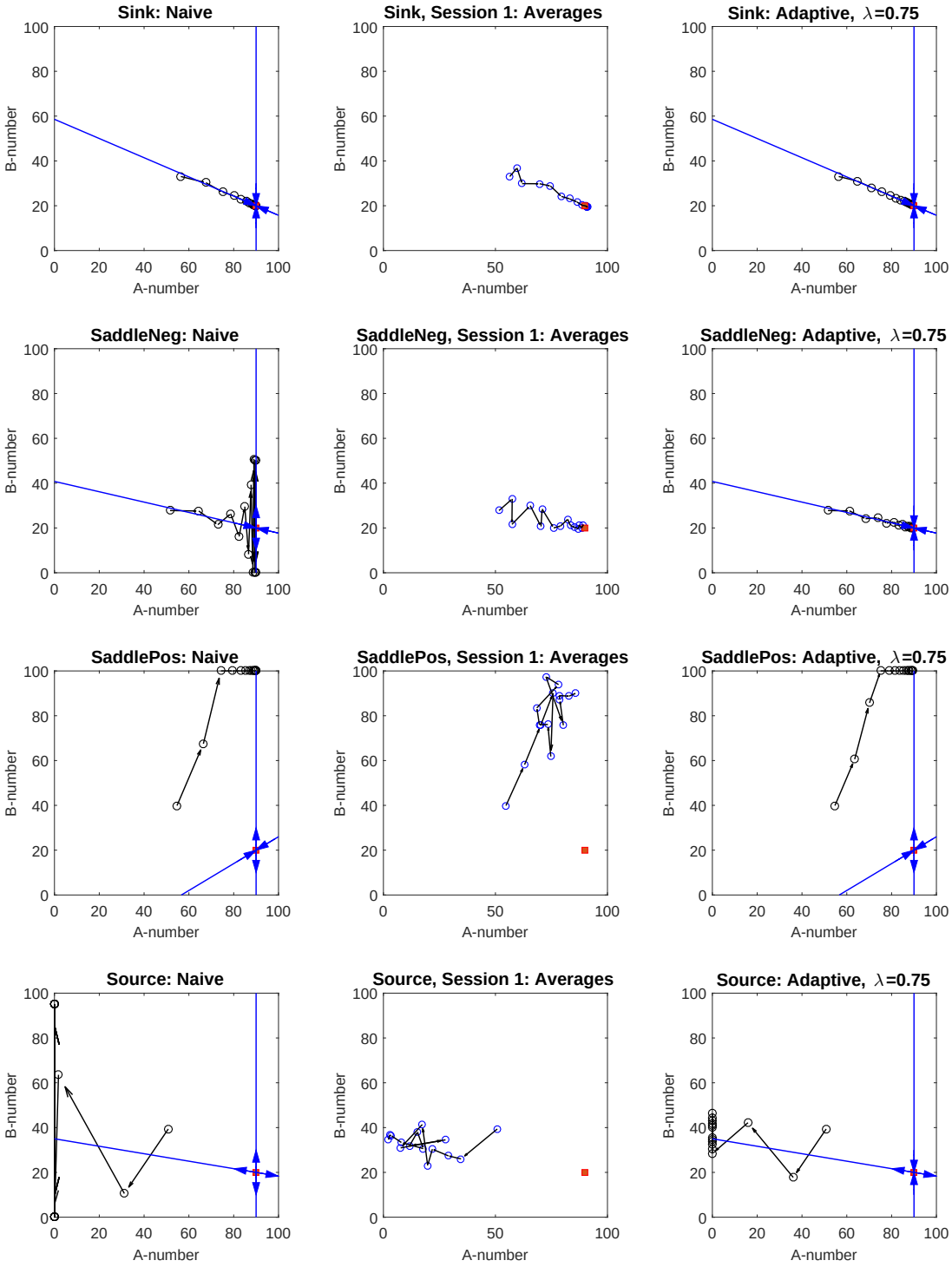


Figure 5: Dynamics of averages in an experimental group (*middle panels*) as compared with the dynamics of naïve model (7) (*left panels*) and adaptive model (12) with $\lambda = 0.75$ (*right panels*). For both learning models the first 15 periods are simulated with the same initial point as in the experimental group. The two blue lines are the eigenvectors with the arrows indicating whether the direction is stable or not.

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 5. 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 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 (6) and simulate, instead, the following dynamics

$$\bar{a}_t = \max\left\{0, \min\left\{100, a^*_{t-1}\right\}\right\} \quad \text{and} \quad \bar{b}_t = \max\left\{0, \min\left\{100, b^*_{t-1}\right\}\right\}. \tag{8}$$

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.[24] 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 5. 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.[25]

Finally, in the **Source** treatment both eigenvalues are outside of the unit circle

[24] 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.

[25] 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 (8) assumes.

and the dynamics diverge away from the interior equilibrium except in the special case where the system starts exactly at that equilibrium, (a^{NE}, b^{NE}). Again, we illustrate the dynamics in the left panel of Fig. 5, in the bottom row. While in both the experimental data and the naïve model simulations, the dynamics diverge from the interior equilibrium, 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 is not consistent with the experimental evidence. Above all, it does not predict convergence to the interior Nash equilibrium in the* ***SaddleNeg*** *treatment. There are discrepancies in other treatments as well, including very different dynamics in the* ***Source*** *treatment.*

5.2 Homogeneous level-k learning model

In Section 4.1 we analyzed first period choices in light of the level-k model of thinking. 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,[26] i.e.,

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} \tag{9}$$

Agents who follows the level-0 choice in each time period can also be labeled as "stubborn" agents, since their guesses do not change over 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 agents would continue to guess 50 for both numbers in all subsequent periods of the experiment.

[26] One may argue that our level-0 is too simplistic. We have chosen this definition over possible alternatives on the basis of 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.

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 when all agents are of level 1, i.e., the homogeneous level-1 model, thus coincides precisely with the naïve model (6) 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.,

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \mathbf{M} \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} + \boldsymbol{d}\,. \tag{10}$$

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 of 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., naïve) model,[27] see the level-2 model in Table 7 and compare the convergence predictions for that model with those for the naïve model in Table 5.

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. On the basis of this we conclude

Result 2. *The homogeneous level-k learning model for any finite k is qualitatively inconsistent with the experimental behavior.*

[27] 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 $|\mu^k| < 1$.

Note that our results hold only if we restrict ourselves to the model with a *homogeneous* level of iterative thinking. Later, in section 5.4, 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 result can then be reversed.

5.3 Adaptive Learning Model

The adaptive learning model that we consider next, generalizes the naïve model by assuming that subjects change their guesses in response to past errors.[28] According to this model, players' choices are adapted towards the targets in a constant proportion to the guessing error. This model implies the following evolution of average choices:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + \lambda \left[\begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} - \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} \right] = \lambda \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} + (1-\lambda) \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}, \quad (11)$$

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

The second part of expression (11) suggests the following interpretation of the adaptive model. Assume that the population consists of individuals with different levels of rationality. If fraction $1 - \lambda$ of the population uses the level-0 model, and 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 (11). This interpretation provides a useful connection between the homogeneous level-k models that we introduced in Section 5.2 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* (9)*, 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

[28] 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 modelling, see Section 3.3 in Evans and Honkapohja (2001).

Treatment	Eigen System μ_1, $\mathbf{v}_1$	μ_2, $\mathbf{v}_2$	Converges to a^{NE} and b^{NE}? for a	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 6: 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 predictions of the model in terms of convergence for a generic initial point.

discussed further in the next section. Here we will continue with properties of the adaptive model. Substituting the target values from (5) into equation (11) we obtain a linear dynamical system for the averages:

$$\begin{pmatrix}\bar{a}_t\\ \bar{b}_t\end{pmatrix} = (\lambda\mathbf{M} + (1-\lambda)\mathbf{I})\begin{pmatrix}\bar{a}_{t-1}\\ \bar{b}_{t-1}\end{pmatrix} + \lambda\boldsymbol{d}\,. \tag{12}$$

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 internal Nash equilibrium $(a^{NE}, b^{NE}) = (\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}$. However, the linear dynamics are now governed by the matrix $\lambda\mathbf{M} + (1-\lambda)\mathbf{I}$. Given that in all our treatments, the matrix $\mathbf{M}$ is lower triangular, the eigenvalues of this matrix are easy to determine. They are:

$$\mu_1 = 1-\lambda+\lambda m_{11} \quad \text{and} \quad \mu_2 = 1-\lambda+\lambda m_{22}\,. \tag{13}$$

The eigenvalues of the adaptive model are thus the weighted averages of 1 (with 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 $\mathbf{M}$.[29] This allows us to characterize the convergence properties of the adaptive model in different treatments for different values of λ, see Table 6.

We find that in order to obtain covergence to the interior equlibrium of the **SaddleNeg** treatment under the adaptive learning dyanmics, the weight, λ, that is as-

[29] 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}.

signed 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" behaviour 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 diverging b-number in the **SaddlePos** treatment (100) and for 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 qualitative properties of the adaptive model, with weight $\lambda \in (\frac{2}{3}, \frac{4}{5})$ are consistent with the experimental evidence in all four treatments.*

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.[30] 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 (12), 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 (8) 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

[30] Online Appendix G provides more discussion on the qualitative properties of the adaptive model and illustrates the dynamics of that model using other values for λ.

thus does not account for discrepancies between the model's prediction and the more noisy experimental dynamics at every time step.

5.4 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, represented by 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.[31]

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 from the Nash Equilibrium) as follows

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

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 $f_2 = 1 - f_1$.[32] The matrix of the corresponding

[31] 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 response 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 behaviour and in Result 2 earlier on.

[32] If our level-0 seems too simplistic, this model can be more reasonable as a mixing model of low

Treatment	Level-2 EVs μ_1	Level-2 EVs μ_2	Level-2 Converges? for a	Level-2 Converges? for b	Mixed model of level-1 (naïve) and level-2 EVs μ_1	Mixed model of level-1 (naïve) and level-2 EVs μ_2	Converges to a^{NE} and b^{NE}? for a	Converges to a^{NE} and b^{NE}? 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 7: Properties of the level-2 and mixed models.

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 7 and observe that for $1/3 < f_1 < 3/5$ the dynamics are qualitatively similar to the experimental data.

Result 5. *Both the mixed cognitive model with a population consisting of level-*0 *and level-*1 *agents and the mixed cognitive model with a population consisting of level-*1 *and level-*2 *agents match, qualitatively, the converging/diverging properties observed in the experiment.*

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 6.

5.4.1 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.[33] 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 discussed so far, as they play a best response to the rest of 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

levels, see footnote 26. However, when our data are used to compare models quantitatively, this mixing model will not be better than the adaptive learning model, as shown in the next section.

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

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

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

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

$$\begin{pmatrix}\bar{a}_t - a^{NE}\\ \bar{b}_t - b^{NE}\end{pmatrix} = \prod_{i=1}^{K}(\lambda_i\mathbf{M} + (1-\lambda_i)\mathbf{I})\begin{pmatrix}\bar{a}_{t-1} - a^{NE}\\ \bar{b}_{t-1} - a^{NE}\end{pmatrix}, \tag{14}$$

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 5.1 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 and 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 (14), 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 (14) converge.[34]

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 by $e^{-\tau}\frac{\tau^k}{k!}$. The only parameter of this distribution, $\tau > 0$, characterizes the average level of rationality in the population.[35] In the special case when the Poisson CH model is truncated to two levels, 0 and 1, the proportion of the level-1 types is given by $\tau/(1+\tau)$. Applying Proposition 5.1 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 5.1. *The adaptive model* (11) *corresponds to the Poisson CH model from Camerer et al. (2004) with two levels of rationality,* 0 *and* 1*, and* $\tau = \lambda/(1-\lambda)$.

We conclude this section by presenting another mixed cognitive levels model. Recall that some subjects in our experiment submitted the internal Nash Equilibrium values as their guesses starting from the very first period. For this reason it is useful to add some weight to this type of behavior[36] and consider the mixed model with levels 0, 1 and the Nash equilibrium:[37]

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = f_0 \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix} + f_1 \begin{pmatrix} a^*_{t-1} \\ b^*_{t-1} \end{pmatrix} + f_{NE} \begin{pmatrix} a^{NE} \\ b^{NE} \end{pmatrix}, \tag{15}$$

where $f_0 + f_1 + f_{NE} = 1;\ f_0, f_1, f_{NE} \in [0,1]$.

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

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

[36] In the **Sink** treatment, where the process of iterative best responses with increasing k converges to the rational expectation equilibrium, such behavior can be thought as level ∞ of rationality. However, in the other three treatments there is no such convergence. Nevertheless, in all four treatments subjects could simply solve the coupled system (which they knew) for the fixed point to find (a^{NE}, b^{NE}), and some subjects seem to have done this. Our model assumes that such "fixed-point-solver" subjects are present in proportion f_{NE}.

[37] We can add the Nash prediction to the other mixed models as well. In Section 6 we estimate the mixed model consisting of levels 0, 1, 2, and Nash.

In deviations from the Nash Equilibrium, the dynamics of this model are given by

$$\begin{pmatrix} \bar{a}_t - a^{NE} \\ \bar{b}_t - b^{NE} \end{pmatrix} = f_0 \begin{pmatrix} \bar{a}_{t-1} - a^{NE} \\ \bar{b}_{t-1} - b^{NE} \end{pmatrix} + f_1 \begin{pmatrix} a^*_{t-1} - a^{NE} \\ b^*_{t-1} - b^{NE} \end{pmatrix} = \\ = \Big(f_0 \mathbf{I} + f_1 \mathbf{M} \Big) \begin{pmatrix} \bar{a}_{t-1} - a^{NE} \\ \bar{b}_{t-1} - b^{NE} \end{pmatrix} .$$

The convergence properties of this model depend on the eigenvalues, which are

$$\mu_1 = f_0 + f_1 m_{11} \quad \text{and} \quad \mu_2 = f_0 + f_1 m_{22} \, . \tag{16}$$

Thus, the eigenvalues determining the model's properties are convex combinations of the eigenvalues of matrix $\mathbf{M}$, 1, and 0. This is similar to the adaptive model (11), except that the presence of Nash/fixed point guessers, with $f_{NE} > 0$ stabilizes the model even more than when adaptive learning stabilized the naïve model (since f_0 and f_1 are now smaller). The model thus is able to generate qualitatively the same dynamics as the adaptive model, whereas the extra parameter f_{NE} gives more flexibility, particularly in adjusting the speed of convergence.

5.5 Average Learning Models

One possible reason why the dynamics of the adaptive and other mixed models is consistent with subjects' behavior is that those models effectively take into account all past target values (and those values were available to the participants in our experiment). For instance in the **SaddleNeg** treatment, initially the dynamics for the b-number are oscillating around the Nash equilibrium. If agents use past averages for the b-number, these oscillations can be dampened and convergence can be obtained.[38]

There are two standard ways of modelling learning using past averaging. First, by rewriting the adaptive model (11) recursively, we obtain for any $t \geq 2$:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \lambda \begin{pmatrix} a^*_{t-1} + (1-\lambda) a^*_{t-2} + \cdots + (1-\lambda)^{t-2} a^*_1 \\ b^*_{t-1} + (1-\lambda) b^*_{t-2} + \cdots + (1-\lambda)^{t-2} b^*_1 \end{pmatrix} + (1-\lambda)^{t-1} \begin{pmatrix} \bar{a}_1 \\ \bar{b}_1 \end{pmatrix} . \tag{17}$$

The adaptive model written in this way involves a constant gain, λ, attached to

[38] 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.

the (exponentially declining) weighted average of all past target values (and some diminishing weight on the initial average). For this reason the model (17) is also called the exponentially weighted moving average (EWMA) model.

A second, alternative model averages the previous L target values with equal weights. We refer to this averaging model as the moving average model, MAve(L).[39] When $t > L$, the average guesses given by this model are:

$$\begin{pmatrix} \bar{a}_t \\ \bar{b}_t \end{pmatrix} = \frac{1}{L} \begin{pmatrix} a^*_{t-1} + \cdots + a^*_{t-L} \\ b^*_{t-1} + \cdots + b^*_{t-L} \end{pmatrix} . \tag{18}$$

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 the corresponding variable.[40]

Table 8 summarizes properties of MAve(L) model (properties of the EWMA averaging model are the same as the adaptive model). When $L = 1$, we have the naïve model, and so the left columns of Table 8 are repeated from Table 5. 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.[41] 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, see Online Appendix H. The last columns of Table 8 summarize the results of this analysis (where the conditions for convergence are that the reported eigenvalues of the T map are negative). As Table 8 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

[39] 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 (18), and so this model is known as the model with *decreasing* gain, which approximates recursive least squares learning.

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

[41] 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) EVs		$L=1$ (naïve) Converges?		$L=2$ Converges?		$L=3$ Converges?		$L=\infty$ (LS learning) EVs of T-map		$L=\infty$ (LS learning) 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 8: 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.

was observed in the experiment and divergence otherwise (**SaddlePos** and **Source**).

To simulate the moving average model we initialize the targets for the first period only. Then, equation (18) is used for $t > L$, whereas when $t \leq L$, all available observations are equally weighted.[42] Figs. 19 and 20 in Online Appendix H 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 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 in Online Appendix H shows simulations of the model when at each step *all available* targets are averaged. This is, in effect, the Ave15 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 from the period 1 target 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

[42] In Eq. (18) the factor on the right-hand side becomes $1/(t-1)$ and the last term in the sums are from period 1.

very slow, and this is because in this average model, 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. Because of this property of weighting new observations with increasingly smaller weights, the average model (18) is often referred to as the *decreasing* gain model.

Result 6. *The convergence properties of MAve(L) model, with $L \geq 2$, are consistent with experimental dynamics across treatments. However, there are some discrepancies in speed and patterns of convergence/divergence.*

6 Estimation

In this section we compare the models discussed in Section 5 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 throug period $t-1$, including previous targets and averages of both numbers) to the average guesses for period t:

$$\begin{pmatrix} \bar{a}_t^m \\ \bar{b}_t^m \end{pmatrix} = \mathcal{L}\left[\begin{pmatrix} a_{t-1}^* \\ b_{t-1}^* \end{pmatrix}, \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}, \begin{pmatrix} a_{t-2}^* \\ b_{t-2}^* \end{pmatrix}, \begin{pmatrix} \bar{a}_{t-1} \\ \bar{b}_{t-1} \end{pmatrix}, \ldots; \boldsymbol{\theta}\right] . \tag{19}$$

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 do not add a session-specific index to the values in the right-hand side of (19).[43]

We estimate and compare the following models:

- the internal Nash equilibrium (NE) prediction;
- the naïve model (6);
- the average models, including the moving average models (18) with 2 to 5 lags;[44]

[43]The presence of the experimental data in the right-hand side of (19) distinguishes this *one-period ahead* prediction from the simulated paths analyzed in the previous section, which does not make use of the experimental data.

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

and the exponential weighted moving average (EWMA) model (17).[45]

- homogeneous level-k models with $k = 0$ (stubborn), $k = 1$ (naïve) and $k = 2$;
- two versions of adaptive model: the standard model as in (11), and its generalization with different parameters, λ_a and λ_b, for a and b numbers (see Appendix G);
- mixed cognitive models with various combinations of level-k models and NE-model;
- the cognitive hierarchy (CH) model with a Poisson distribution over levels up to $k = 5$.

We estimate the parameters of the models (containing parameters) by minimizing the sum of squared errors (SSE) between model predictions and the actual data from the experiment. Since estimating the adaptive model requires one lag, for 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.[46] Parameter estimates resulting from this estimation for different models are shown in column 2 of Table 9. 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 boot-

[45] We stress that although the EWMA model (17) is equivalent analytically to the adaptive model (11), 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 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 (17). The initial $\begin{pmatrix} \bar{a}_1 \\ \bar{b}_1 \end{pmatrix}$ on the right-hand-side of this equation is set to $\begin{pmatrix} a_1^* \\ b_1^* \end{pmatrix}$.

[46] We also allowed for variations in parameters between different treatments, but this did not result in superior model fit (in terms of the leave-one-out MSEs as explained further). As a robustness check, in addition to the full sample estimation, we restimated all models using the data from the converging **Sink** and **SaddleNeg** treatments only. There were quantitative changes in some parameters, but the overall performance rankings of the models remained similar (see Appendix I).

Models		Parameter Estimates	Out-of-sample MSEs Sink	SadNeg	SadPos	Source	Overall
NE		-	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	.42 (.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	$\lambda_a = \lambda_b$	.44 (.03)	32.33	13.42	84.68	114.58	61.25
	$\lambda_a \neq \lambda_b$	.38 (.03) .47 (.04)	32.48	13.06	87.86	112.34	61.44
Mixed	0 1 (Adaptive)	.56(.03) .44(.03)	32.33	13.42	84.68	114.58	61.25
	0 1 2	.56(.03) .44(.03) .00(.00)	32.33	13.42	84.68	114.58	61.25
	1 2	.84(.03) .16(.03)	48.73	15.68	232.68	279.69	144.19
	0 NE	1.00(.00) .00(.00)	48.82	59.01	138.87	211.58	114.57
	1 NE	.86(.01) .14(.01)	58.94	27.29	113.92	129.42	82.39
	0 1 NE	.41(.02) .52(.02) .07(.01)	34.05	9.05	74.06	87.14	51.07
	0 1 2 NE	.40(.03) .50(.02) .03(.02) .07(.01)	34.98	9.82	74.61	86.70	51.53
	1 2 NE	.71(.02) .16(.02) .14(.01)	65.89	23.06	127.17	93.00	77.28
CH-Poisson $\max k = 5$		.56 (.05)	32.54	16.10	83.20	117.03	62.22

Table 9: Estimation and performance of various learning models in terms of the out-of-sample, one-step-ahead MSE using the leave-one-out procedure. MSEs are computed for periods 2 to 15.

strap 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.

Next, we compare the performances of the 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 9. 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.[47] We perform this procedure 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 internal NE 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 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 averaging model. This finding is predominantly due to the better performance of the stubborn, level-0 model in the case of the non-convergent treatments. Next we observe that the adaptive models perform substantially better than all prior models considered in Table 9. 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, 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. Interestingly, the estimated parameter of the single parameter adaptive model, $\widehat{\lambda} = 0.44$ is lower than the theoretically derived parameter range $\left(\frac{2}{3}, \frac{4}{5}\right)$ which can explain the fluctuations observed in the experimental data (see Result 4). One of the explanations is that the econometric model allows the exogenous noise (fluctuations), while in Section 5 we considered only the deterministic path. Moreover, when we re-estimate the parameter λ considering only the converging **Sink** and **SaddleNeg** treatments, we get $\widehat{\lambda} = 0.61$, which is closer to the theoretical lower bound. Finally, the mixed cognitive model with levels 0, 1 and the internal NE outperforms all other considered models (in terms of average MSE). Note that this model is equivalent to

[47] It is called "out-of-sample" because the data of the session for which the MSE is computed were not used to estimate the model parameter(s). This procedure has an implicit penalty for the number of parameters (over-parametrized models perform better "in-sample", but show worse "out-of-sample" performance because of over-fitting). AIC or BIC are popular model selection criteria which have explicit penalty for the number of parameters. In fact, leave-one-out cross-validation is asymptotically equivalent to AIC (Stone, 1977).

Level	0	1	2	NE
Periods 1-15	**0.52**	**0.40**	0.03	**0.06**
	(0.02)	(0.02)	(0.02)	(0.01)
Period 1	**0.68**	**0.18**	0.00	**0.14**
	(0.03)	(0.04)	(0.01)	(0.03)
Periods 2-15	**0.40**	**0.50**	0.02	**0.07**
	(0.03)	(0.02)	(0.02)	(0.01)

Table 10: Mixed model with levels 0, 1 and 2 and NE (internal Nash Equilibrium) for different periods. All parameter estimates in bold significant at 5% level.

the one-parameter adaptive model with the added internal NE. The mixed model with levels 0, 1, 2 and the internal NE is very close, but in second place. The CH-Poisson model shows reasonably good performance, but it performed slightly worse than a simpler adaptive model and the best mixed model. The parameter estimate of the CH-Poisson model $\widehat{\tau} = 0.56$ indicates that the average level of rationality is somewhere between level-0 and level-1.

Next, we estimate the weights of the general mixed model with levels 0, 1, 2 and NE, on all periods 1–15 and then investigate whether they change between periods. In particular, we focus on the change between period 1 and periods 2–15. Table 10 shows parameter estimates and their standard errors (in parenthesis) for the mixed model using the whole sample and these sub-periods. The standard errors are estimated by the bootstrap method previously described. The parameter estimates that are significant at the 5% level are denoted by bold font. Moreover, for these estimates the differences between periods 1 and 2-15 are also significant. We observe that in period 1, there is more weight on level-0, while in the remaining periods the weight of level-1 is significantly higher. The weight on the internal NE in periods 2-15 falls in comparison to period 1. This may be viewed as some evidence that more participants learned to use level-1 and those who computed the internal NE in the first period learned not to abandon that strategy. The parameter estimates on level-2 are not significantly different from zero in all sub-periods. Robustness checks using data from just the two converging treatment, **Sink** and **SaddleNeg** yield similar estimates.

7 Conclusion

We conduct the first experiment generalizing the Beauty Contest game to two dimensions. The model we study is connected to the IO, trade and macroeconomic literatures, where two dimensional coupled systems in which expectations matter often arise. For simplicity, we have focused on the simplest possible structure, such that variable a is decoupled, but variable b depends on the average predictions of both a and b.

We conduct four treatments with various eigenvalues for the planar system. Convergence is achieved in the Sink treatment and in the Saddle treatment but only when there is negative feedback (the sign of the unstable eigenvalue is negative). Rather remarkably, participants are able to learn and follow the saddle path in this case without initially being placed on the saddle path itself. Saddle path 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 the Saddle with positive feedback (the sign of the unstable eigenvalue is positive) and the Source dynamics reliably result in divergence away from the interior steady state. 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 both convergence dynamics and fit of the estimated model. Indeed, a further 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 we consider, we find that a mixed, cognitive levels model (with levels 0, 1 and NE), which is analogous to the single-parameter adaptive learning model with some weight on the internal NE, 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.

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. forthcoming.

Assenza, Tiziana, Peter Heemeijer, Cars Hommes, and Domenico Massaro, "Managing Self-organization of Expectations through Monetary Policy: a Macro Experiment," CeNDEF Working paper 14-07, University of Amsterdam 2014.

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.

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, 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.

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.

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.

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.

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.

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.

APPENDIX

A Motivating Examples

In this appendix, we provide two motivating examples of economic models that map into the basic framework described in Section 2.

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 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 average price in market A.[48] 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 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

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

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

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

Thus price setting game of N firms is exactly equivalent to our planar beauty contest game as introduced in Section 3.

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:

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

[48] 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).

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.[49]

The system can be rewritten in the following matrix form:

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

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 multivarite, 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).

[49]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.

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 $a_i = a^{NE}$ and $b_i = b^{NE}$ for every i, the averages are $\bar{a} = a^{NE}$, $\bar{b} = b^{NE}$, 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 $(a^{NE}, b^{NE}) = (90, 20)$ 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) 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 $a^{NE} = 90$ 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 3 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:

A* = 30 + (2/3) × Average of all "A-numbers"
B* = 75 − (1/2) × Average of all "A-numbers" − (1/2) × Average of all "B-numbers".

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

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

50.50 100 20.25

and "B-numbers"

10.79 30 0 .

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

$$A^* = 30 + (2/3) \times 56.92 = 67.95; \text{ and}$$
$$B^* = 75 - (1/2) \times 56.92 - (1/2) \times 13.6 = 39.74.$$

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

About your task: The experiment lasts for 15 periods. Each period your only task is to choose two numbers, an "A-number" and a "B-number". Your goal is to choose these numbers to be as close as possible (in absolute value) to the target values A*

and B* which will be determined in each period, using the procedure explained above after all participants in your group have made their choices. Both numbers you choose should be between 0 and 100, inclusive. You may enter a real number with up to 2 decimals.

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

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

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

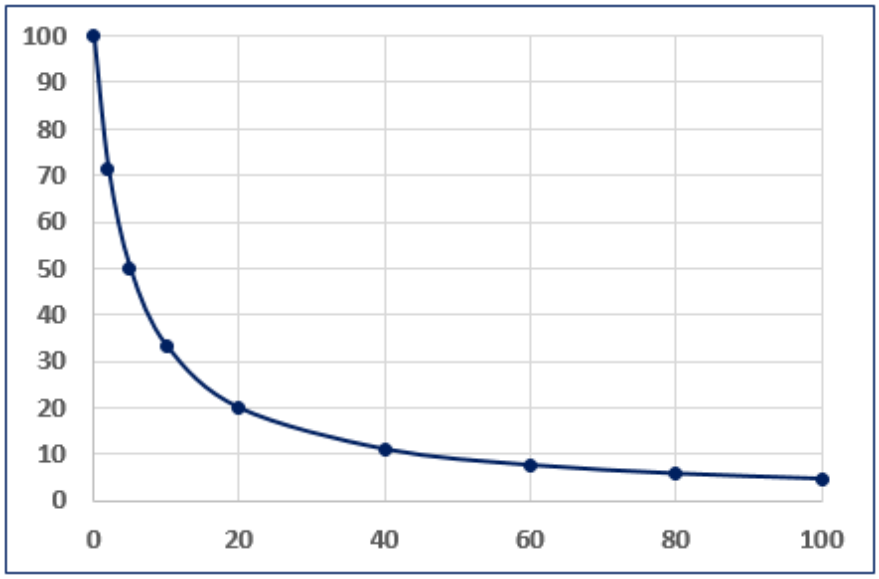
Points Table The table gives the number of points for a given discrepancy of "A-number" from the target value A* (the first column) and a given discrepancy of "B-number" from the target value B* (the first row).

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

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

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

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



Additional Information

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

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	-	6	15	9	5	5	5
5	8	-	-	-	11	6	8	7
1	10	-	-	-	-	12	12	9
0.5	15	-	-	-	-	14	13	14

Table 11: 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.

- 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 11 confirm the conclusion made in the main text for comparison between the **SaddleNeg** and **Sink** treatments. Not only convergence 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 $U_\varepsilon = (a^{NE} - \varepsilon, a^{NE} + \varepsilon)$. Table 12 shows the first hit time (and the period of irreversible entry in parentheses) for all sessions of

	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 12: 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.

the **Sink**, **SaddleNeg** and **SaddlePos** treatmentsfor five different values of ε. (In the **Source** treatment, the trajectories never entered the U_ε neighborhood for $\varepsilon \leq 20$.

Table 12 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 5.1

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

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

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

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

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 4.1. The thick red lines correspond to the rational expectation choice (a^{NE}, b^{NE}). 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 internal Nash equilibrium. 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 internal equilibria, $a^{NE} = 90$ and $b^{NE} = 20$. The middle and right panels show the same evolution as the phase diagram (the right panels 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 levels of internal equilibria, $a^{NE} = 90$ and $b^{NE} = 20$. In the only equilibrium where all participants guess the 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**.

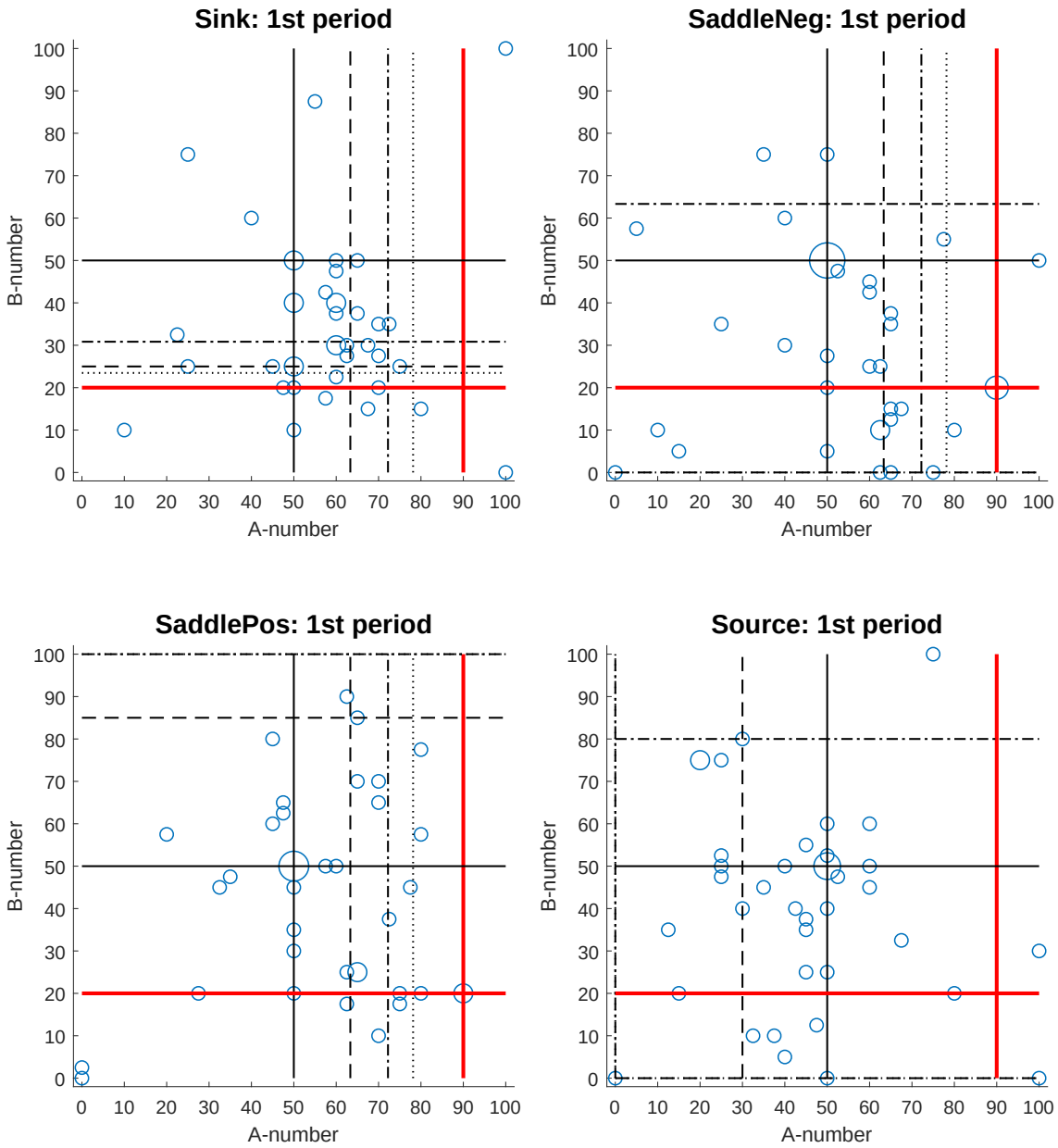


Figure 6: Frequencies of individual pair of guesses in period 1 and levels of reasoning (different dashed lines) and internal Nash equilibrium (red thick line) for four experimental treatments.

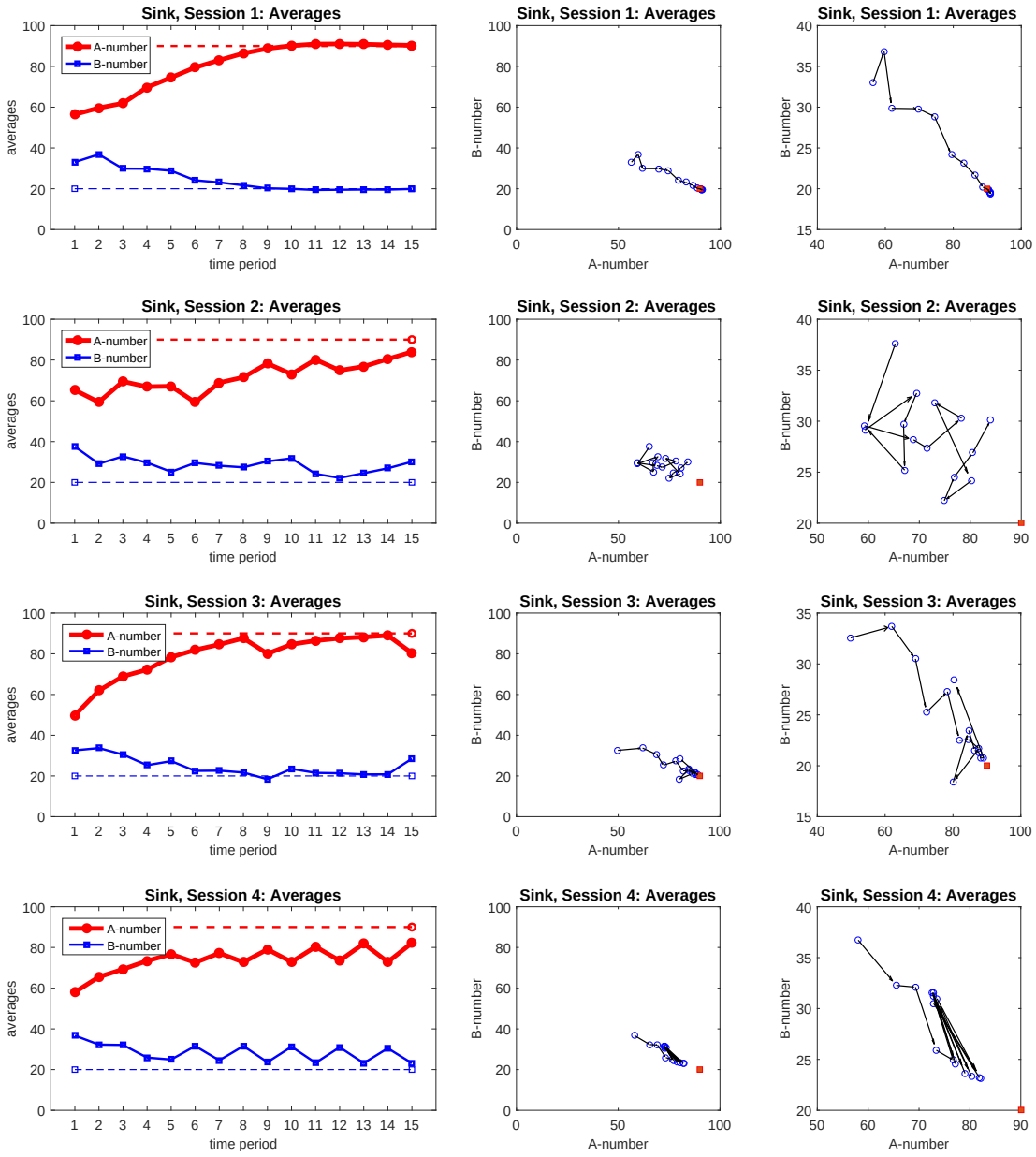


Figure 7: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **Sink** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

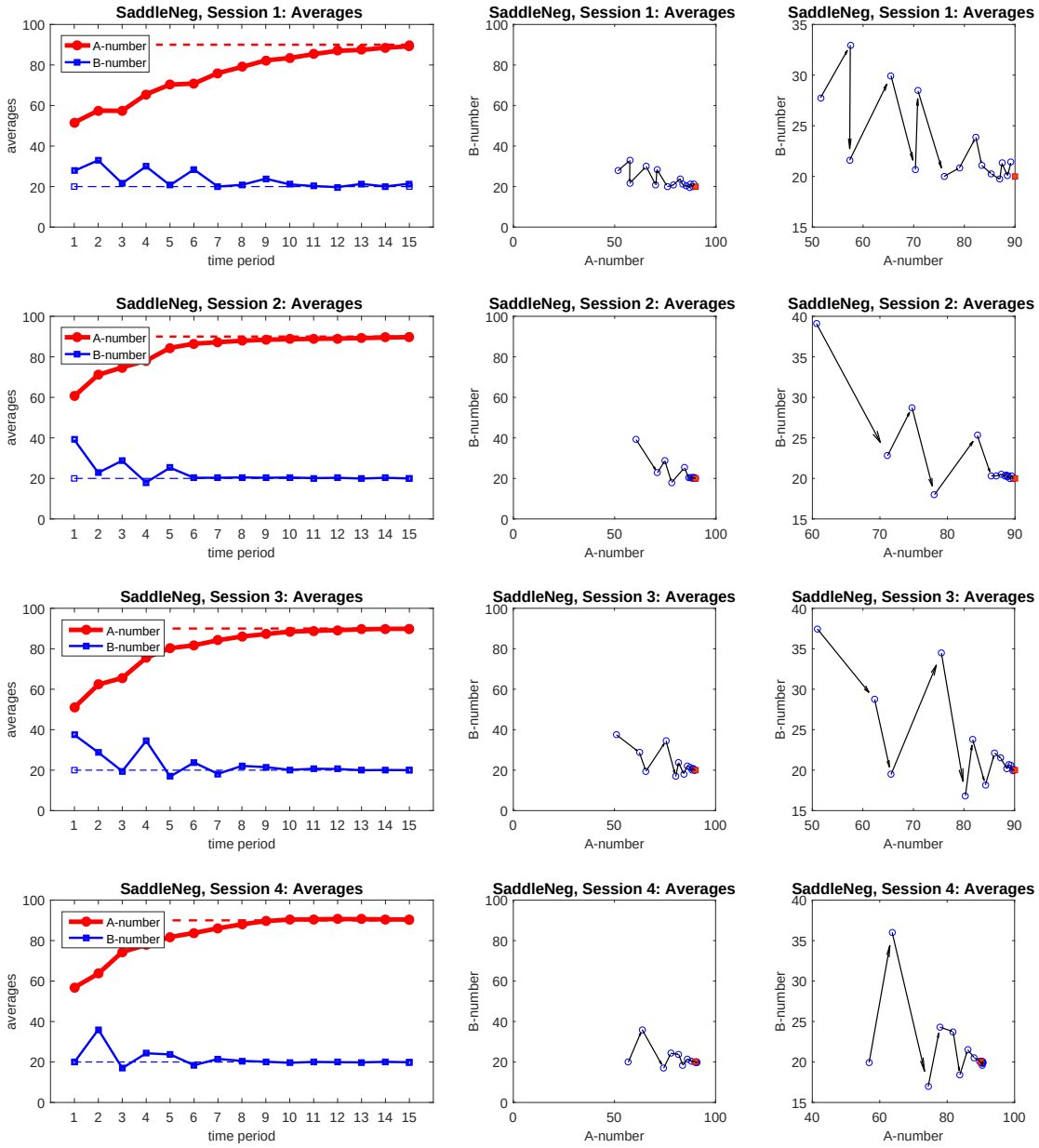


Figure 8: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **SaddleNeg** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

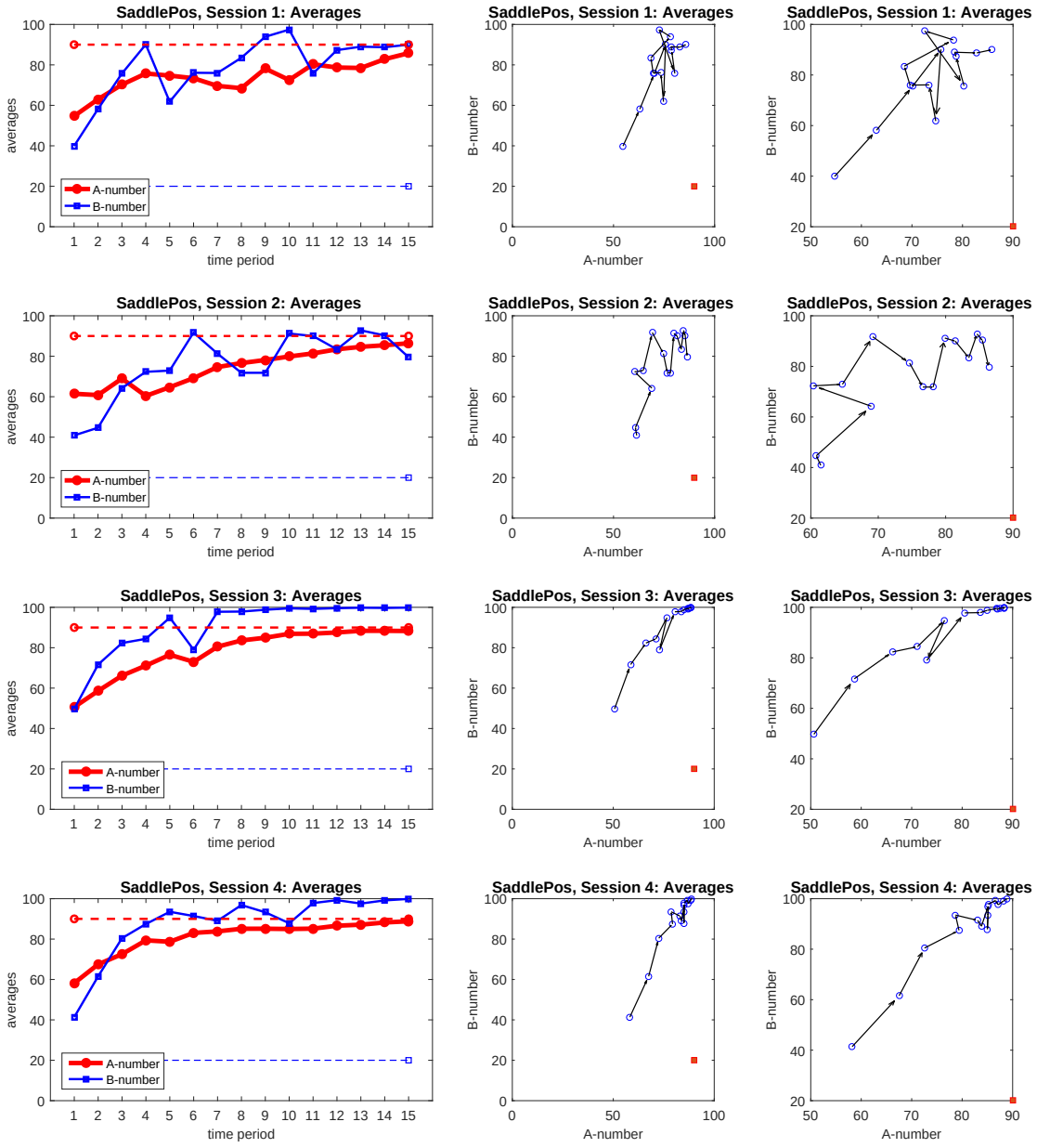


Figure 9: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **SaddlePos** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

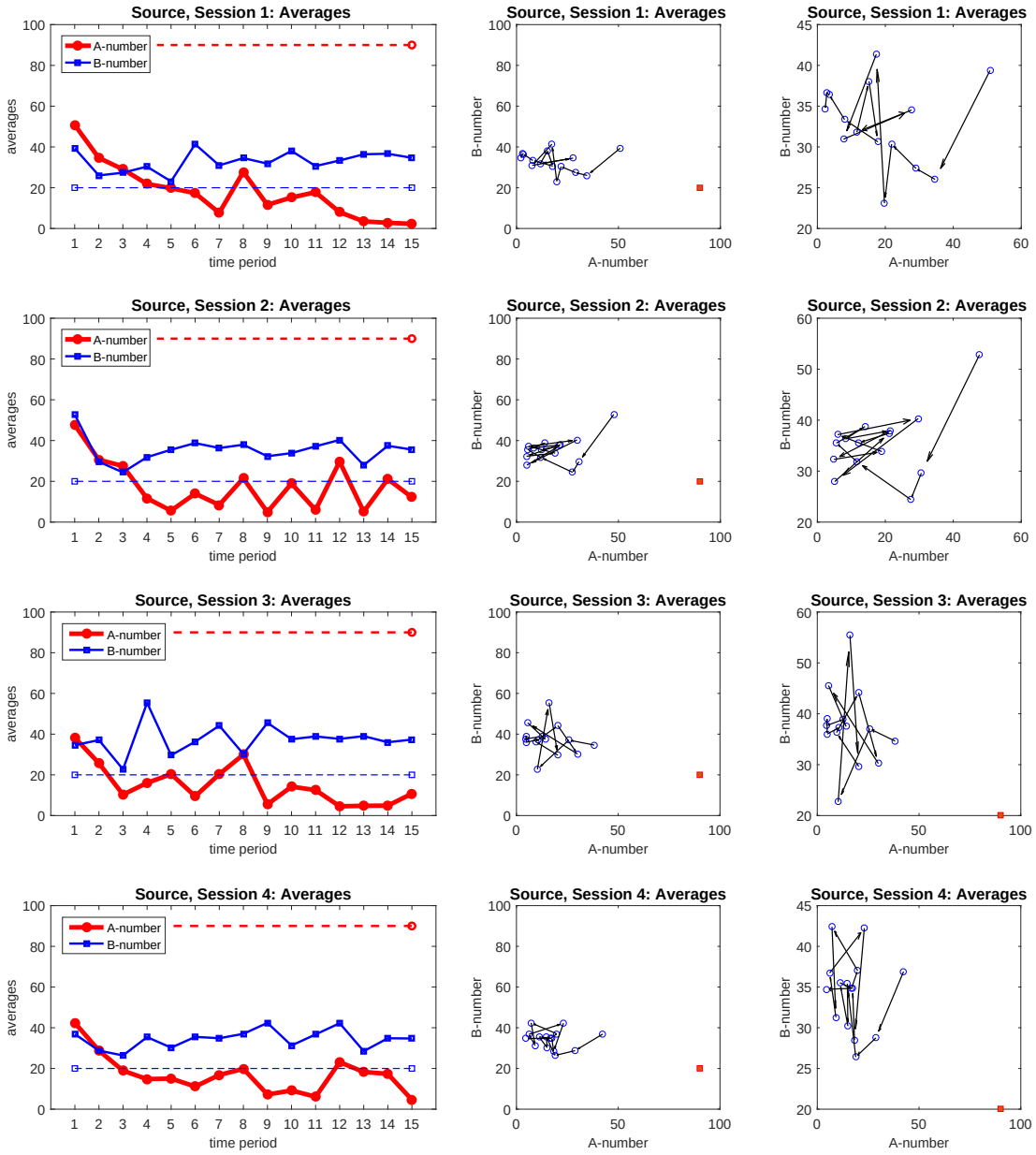


Figure 10: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **Source** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

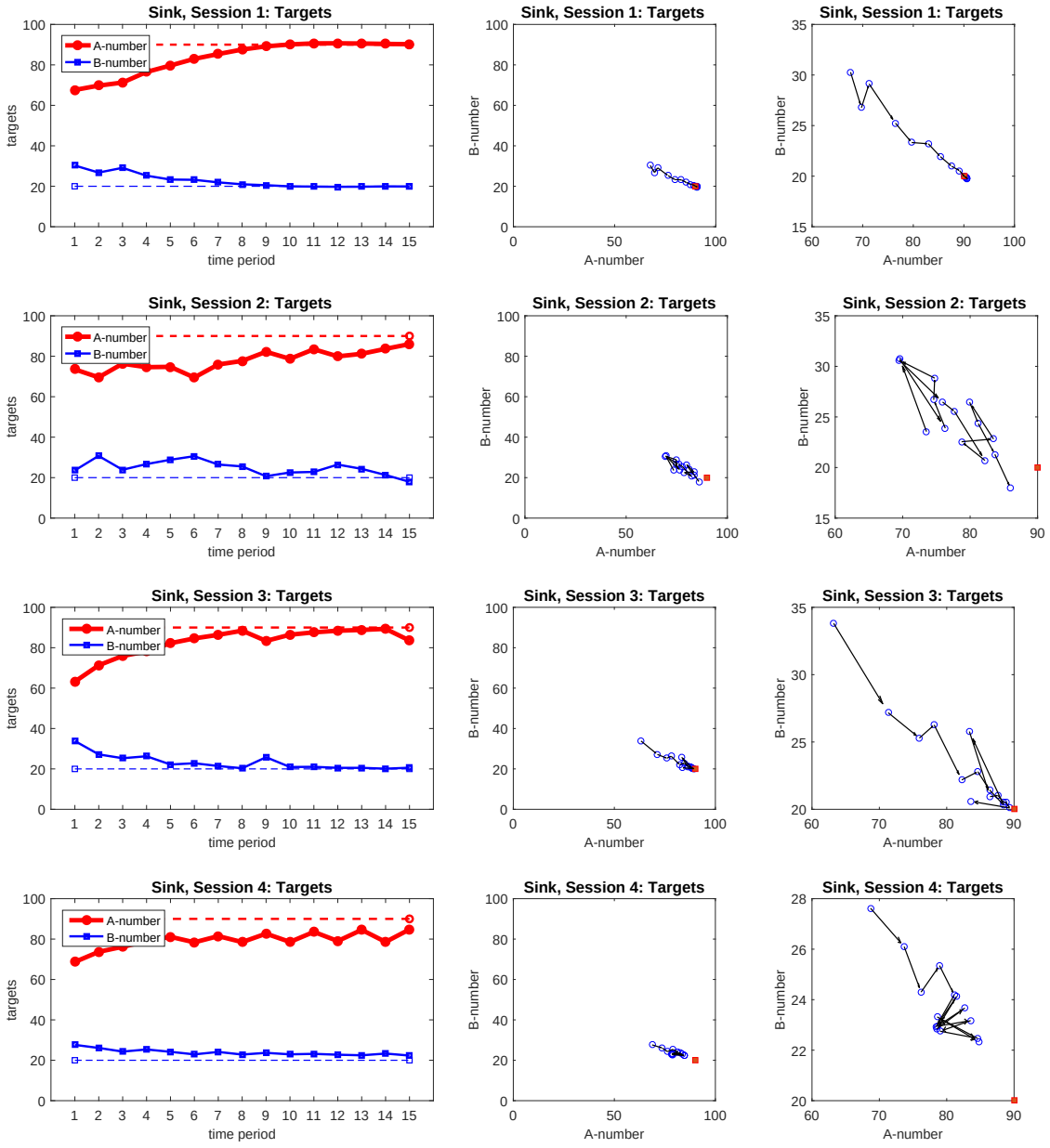


Figure 11: Dynamics of the target values, a^* and b^* in the **Sink** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

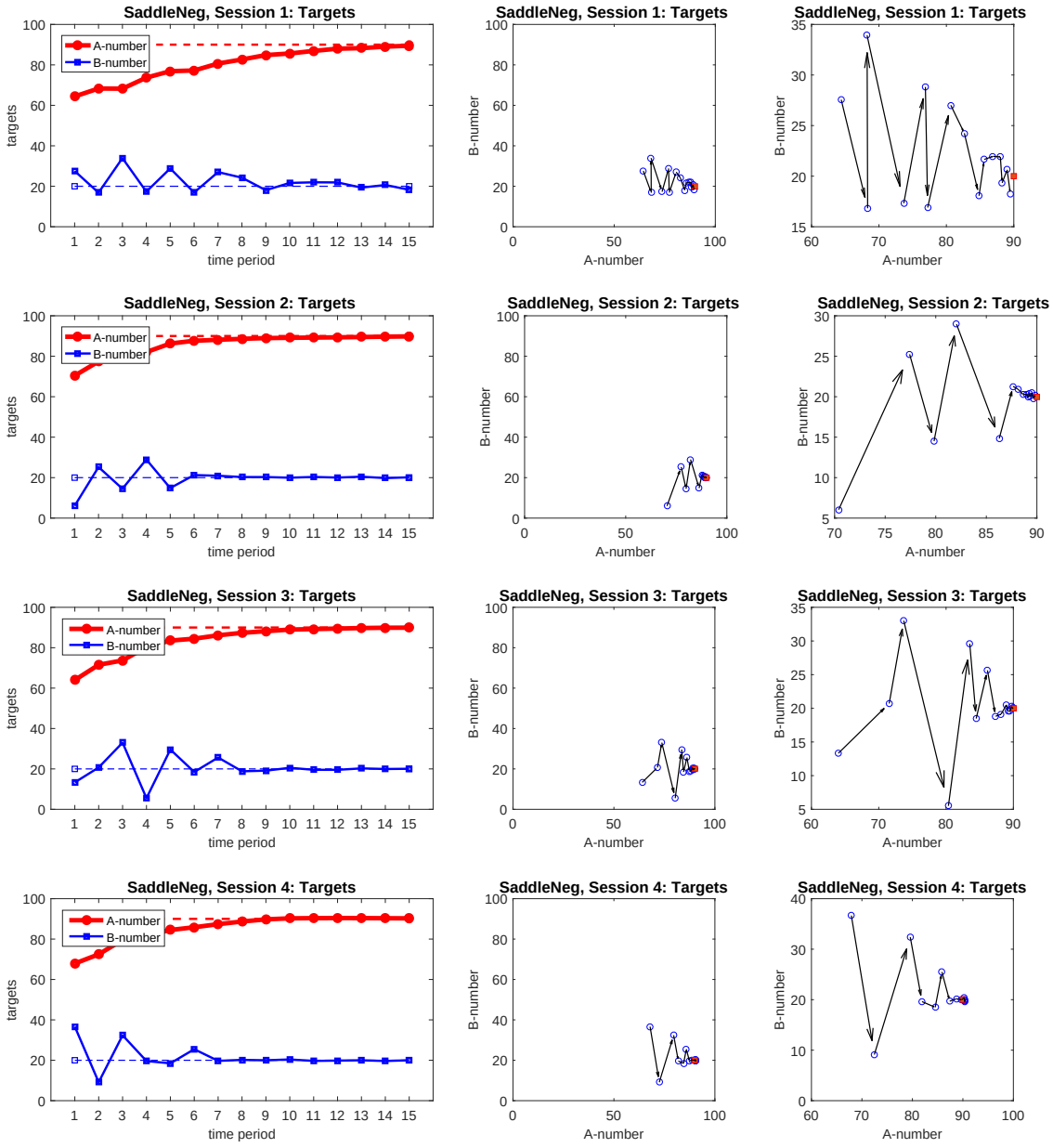


Figure 12: Dynamics of the target values, a^* and b^* in the **SaddleNeg** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

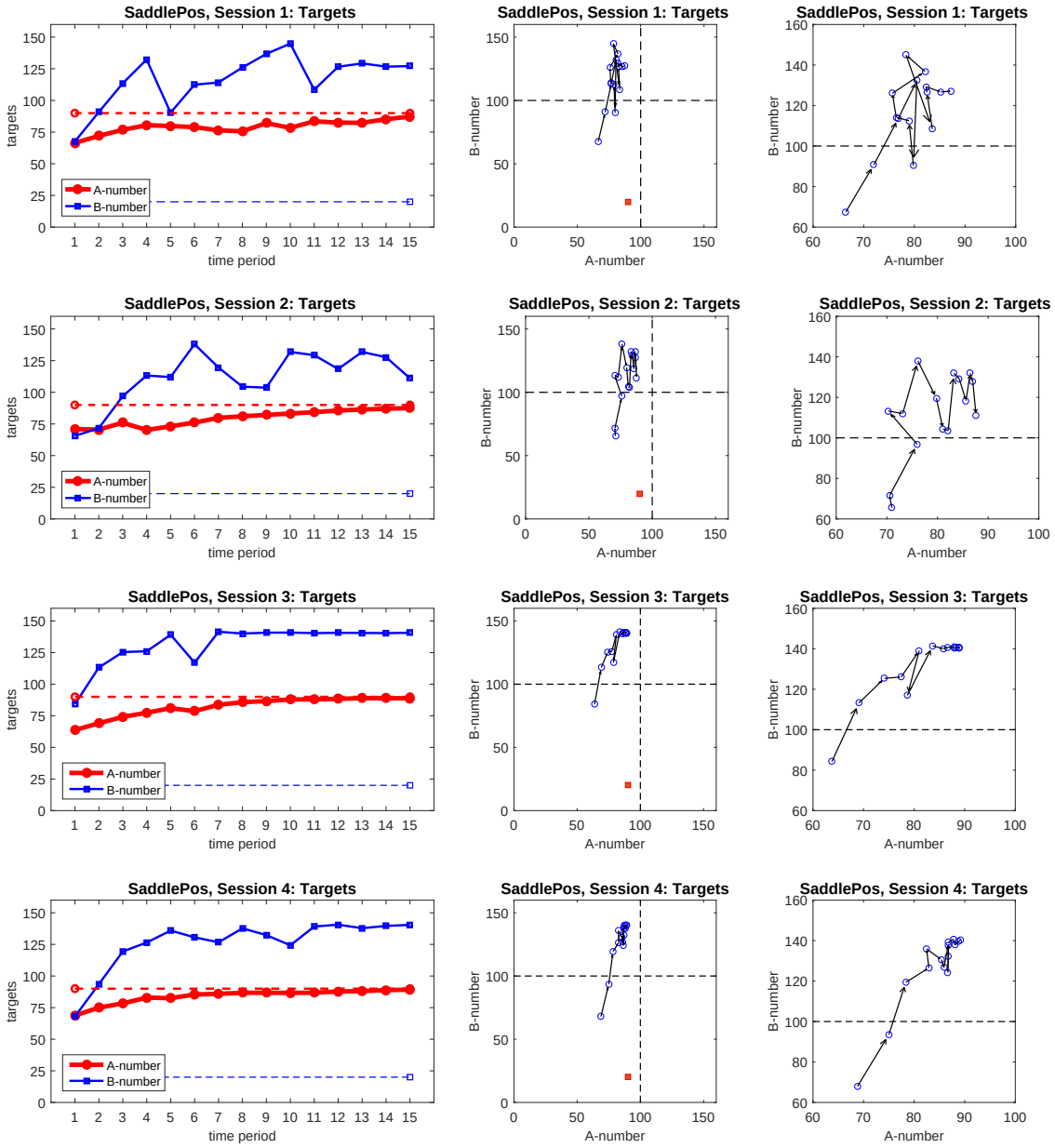


Figure 13: Dynamics of the target values, a^* and b^* in the **SaddlePos** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

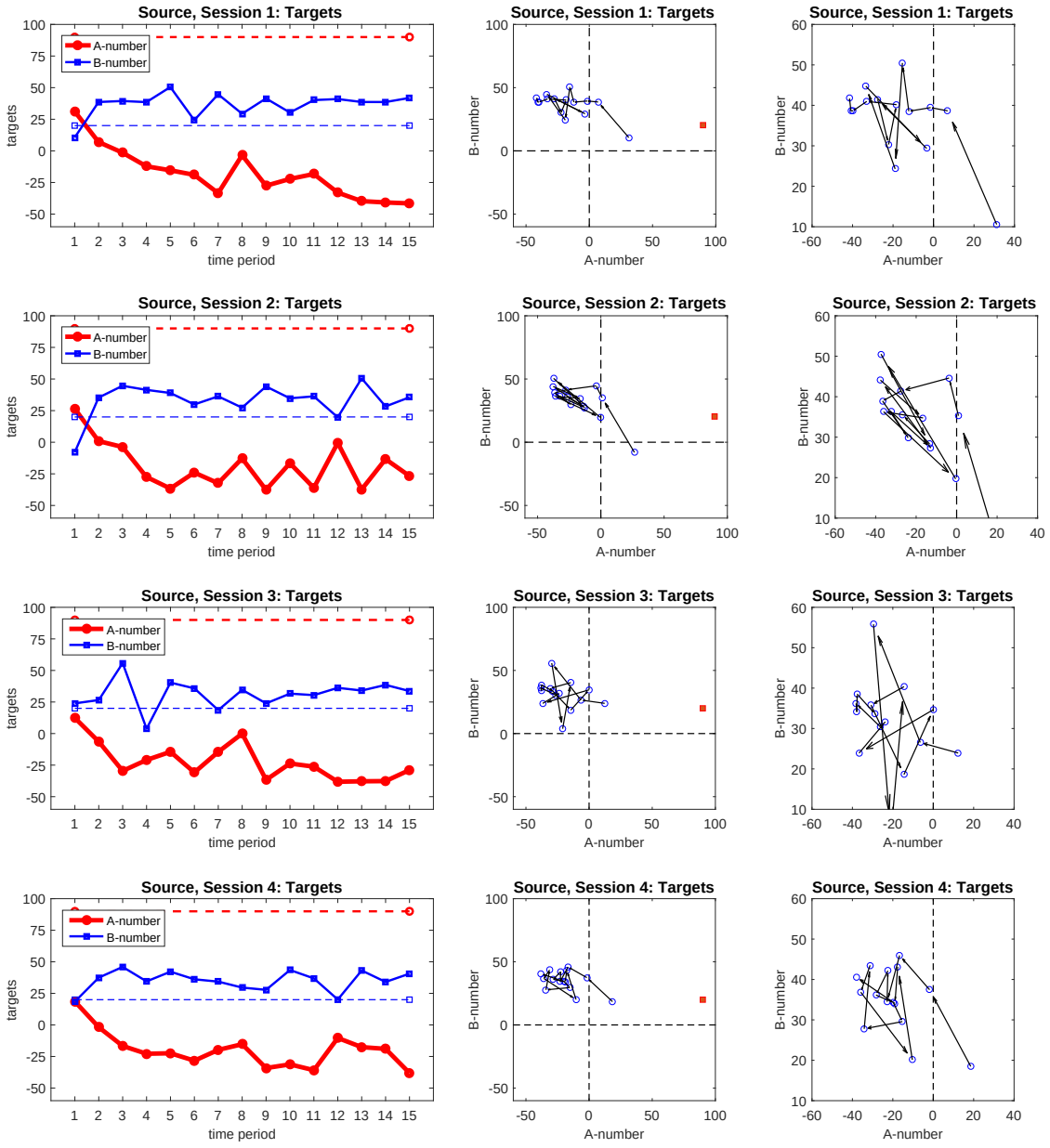


Figure 14: Dynamics of the target values, a^* and b^* in the **Source** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

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

In Section 5.3 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 5.3 for four various values of parameter λ and discuss the qualitative properties of this model for our four treatments; see Table 6 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 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 such model in Section 6.

Substituting into (G.1) equation (5) that defines the target value, we obtain

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

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 $\begin{pmatrix} \bar{a}^{NE} \\ \bar{b}^{NE} \end{pmatrix}$. Then at the steady state we have

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

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

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

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

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

with matrix $\widetilde{\mathbf{M}}=\mathbf{I}-\mathbf{\Lambda}+\mathbf{\Lambda}\mathbf{M}$. 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 5. First, the naïve model, as introduced in Section 5.1, 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 5.

Next, the adaptive model, as described in Section 5.3, is a special case of model (G.1) with common $\lambda = \lambda_a = \lambda_b$. This justifies our focus on eigenvalues as in (13) and results in Table 6. 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 qualitative dynamics of the adaptive model presented in Table 6 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 (13)) 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.[50] When $\lambda > 0.8$ the dynamics of b variable diverges. Given that in our simulations we impose bounds as in (8), 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. 8, 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 5).

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 diverges to 0. The rate of divergence is $\mu_1 = 1 + \lambda/2$ and thus the smaller λ is, the

[50] 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.

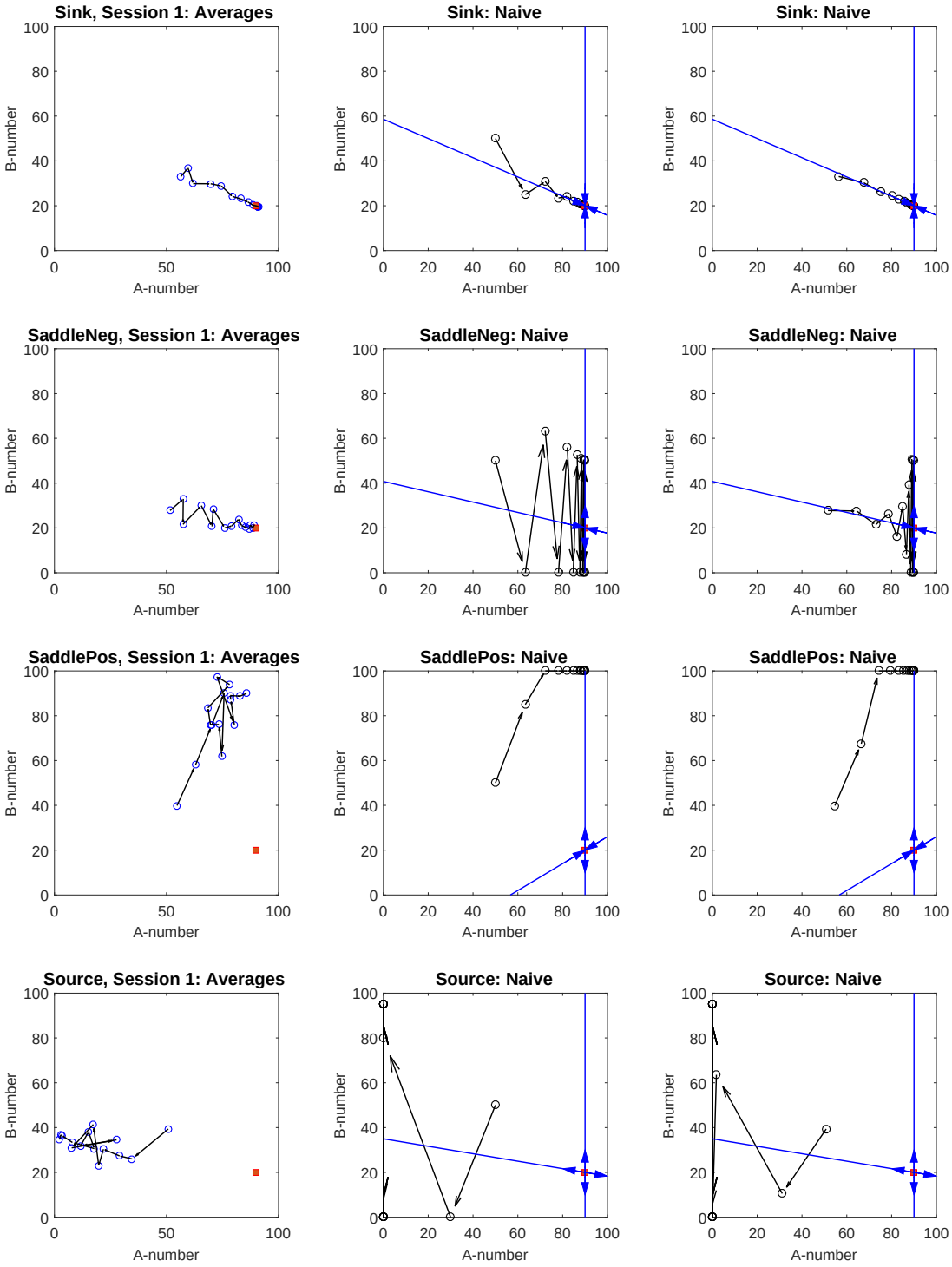


Figure 15: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the naïve model. First 15 periods are simulated using the naïve model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group (*right panels*). The two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

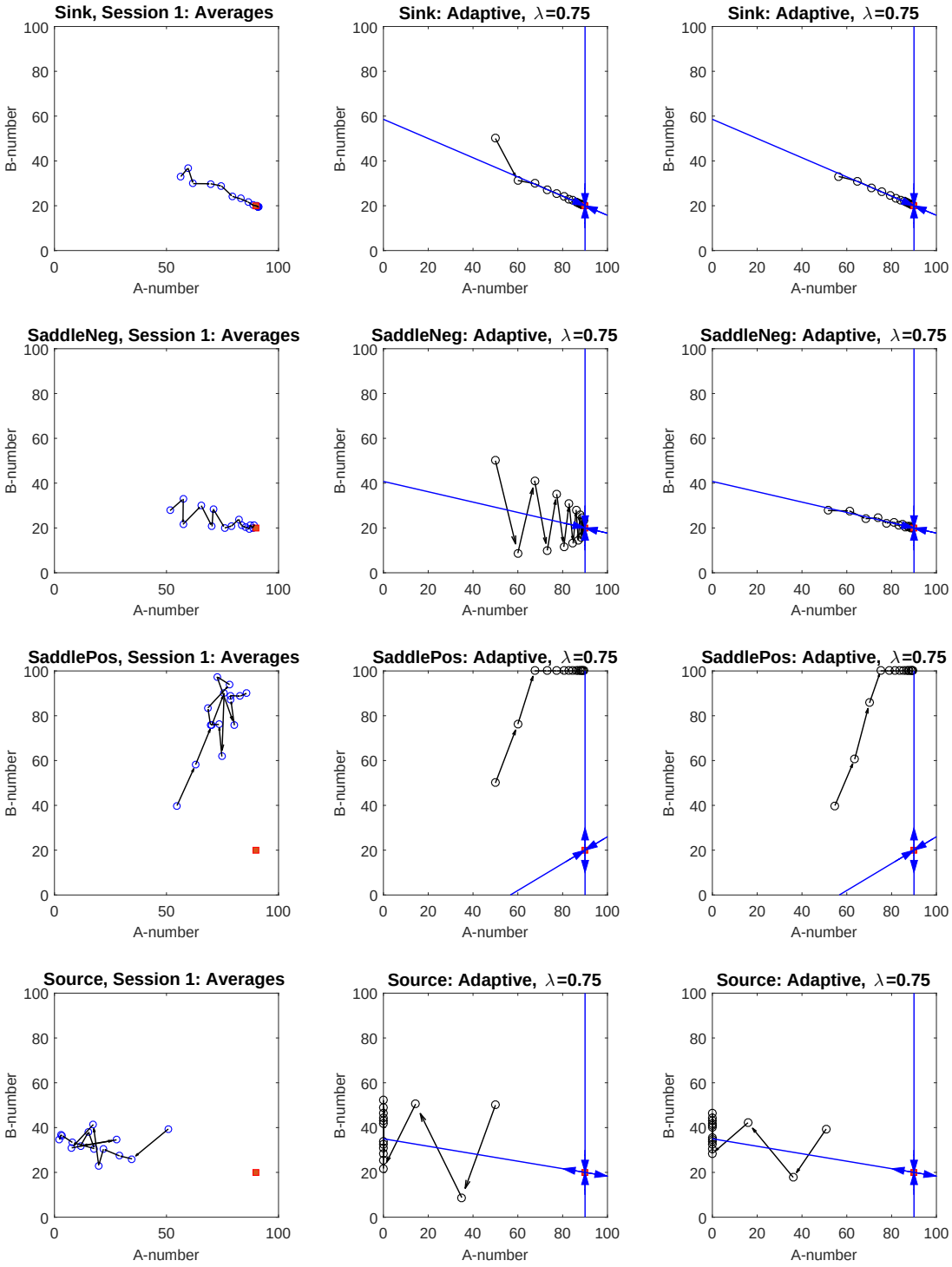


Figure 16: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the adaptive model with $\lambda = 0.75$. First 15 periods are simulated using the adaptive model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group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

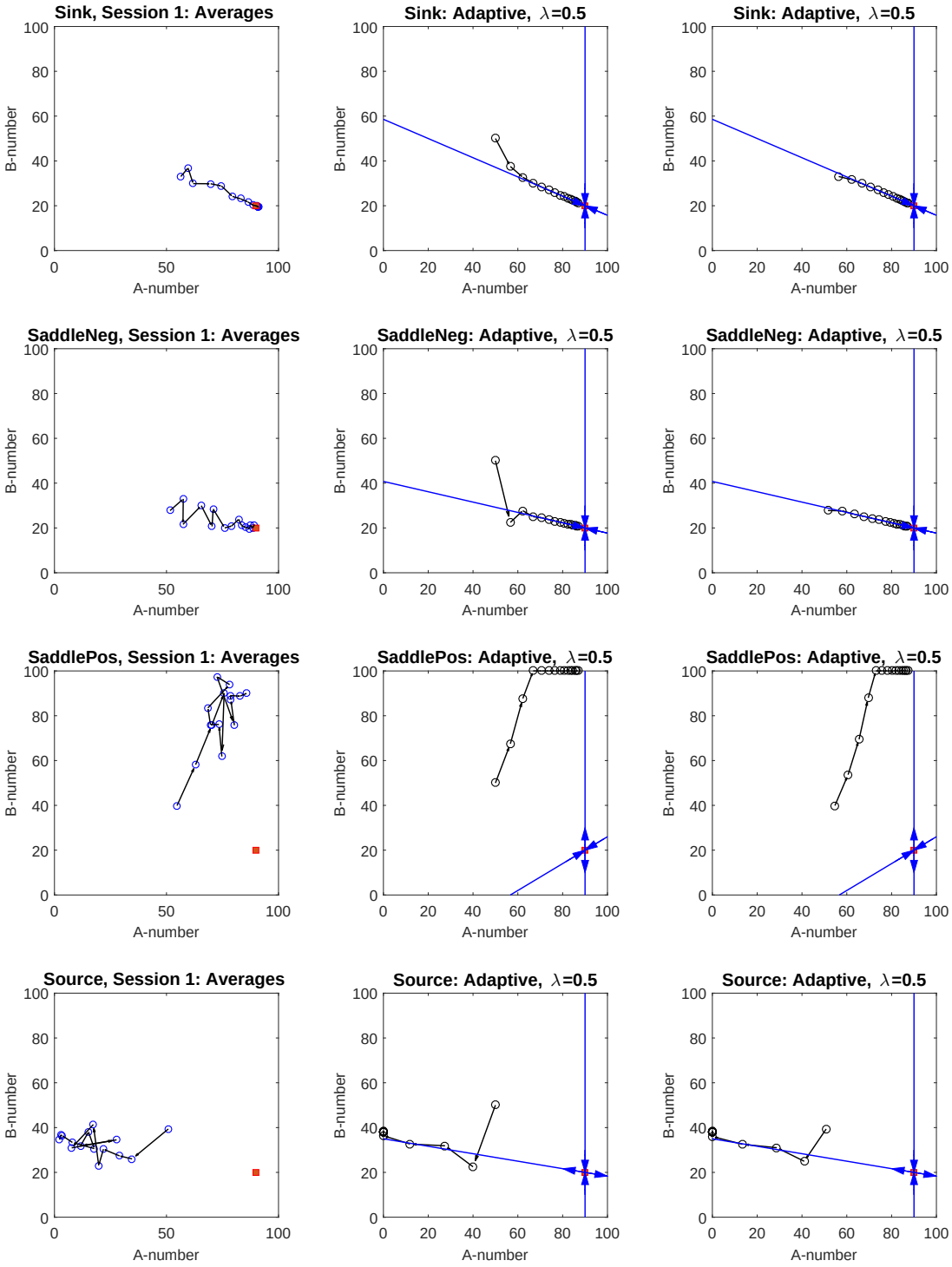


Figure 17: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the adaptive model with $\lambda = 0.5$. First 15 periods are simulated using the adaptive model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group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

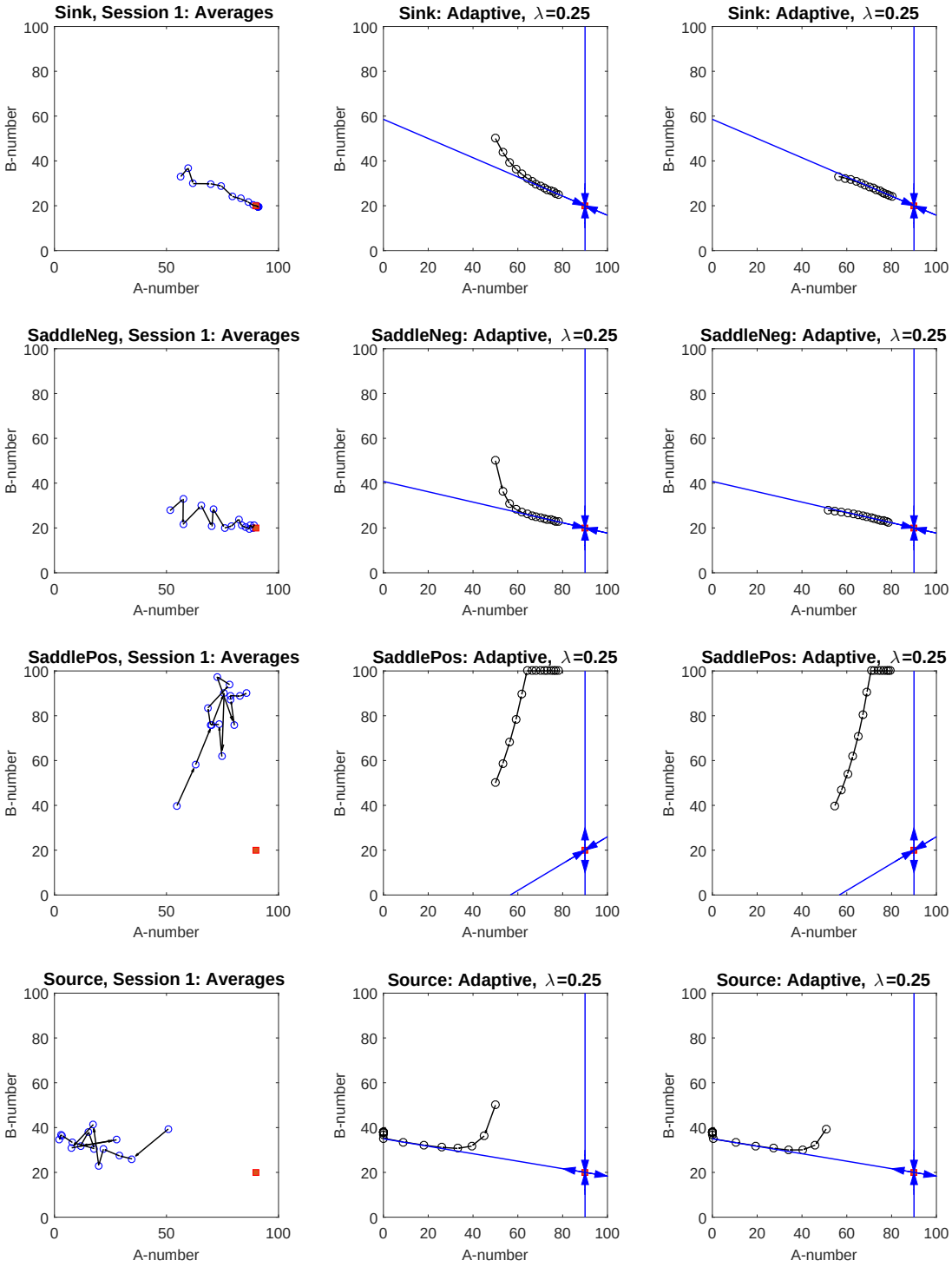


Figure 18: Dynamics of averages in an experimental group (*left panels*) as compared with the dynamics of the adaptive model with $\lambda = 0.25$. First 15 periods are simulated using the adaptive model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experimental group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

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.[51] 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 $y = 18$.

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

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

In Section 5.5 we introduced the moving average model with L lags. Figs. 19 to 21 compare the dynamics in the experiment with dynamics of the moving average model with $L = 2$, $L = 3$ and $L = 15$ lags, respectively.

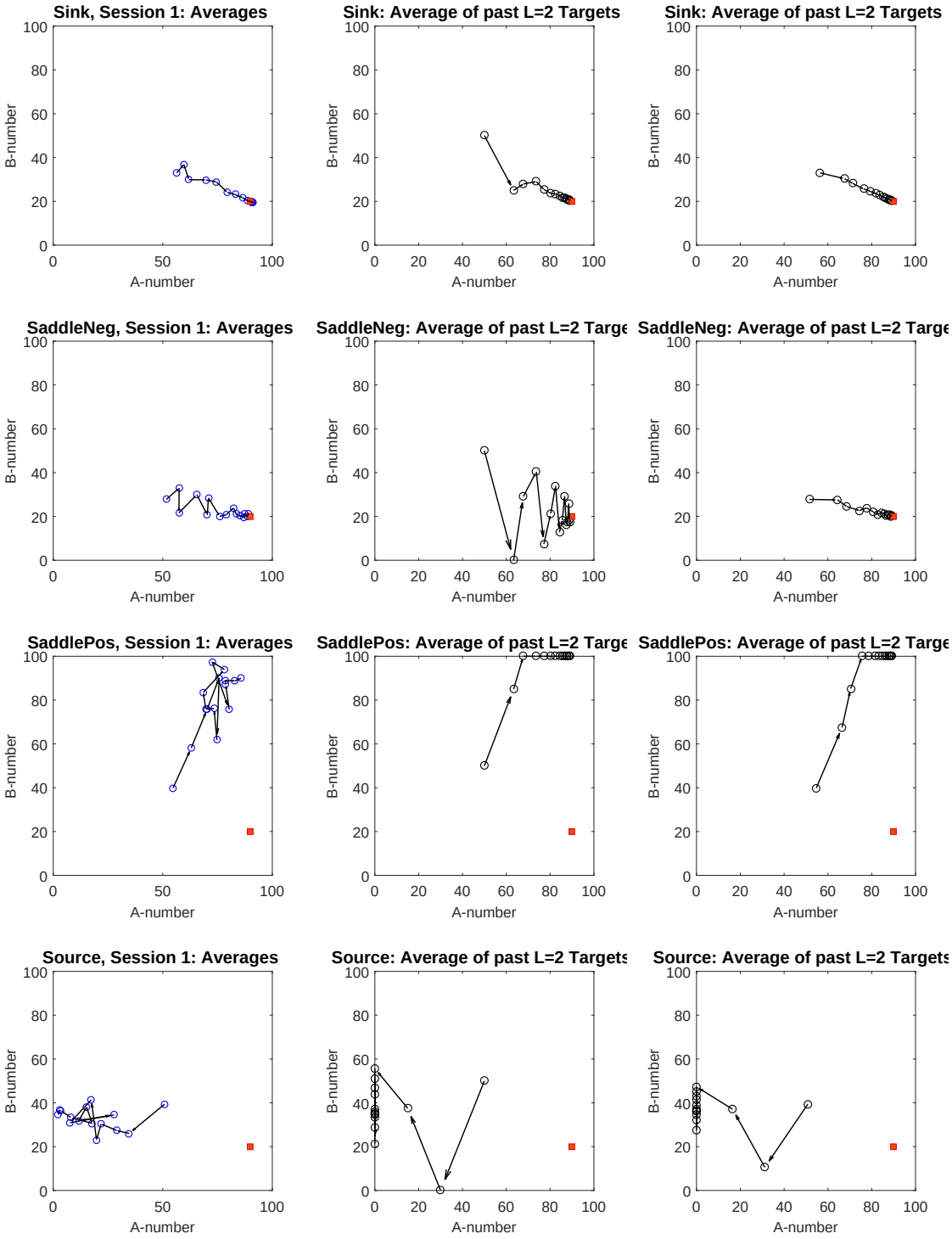


Figure 19: Dynamics of averages in the experimental group (*left panels*) as compared with the dynamics of the MAve(2) model. First 15 periods are simulated using the moving average model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experiment (*right panels*).

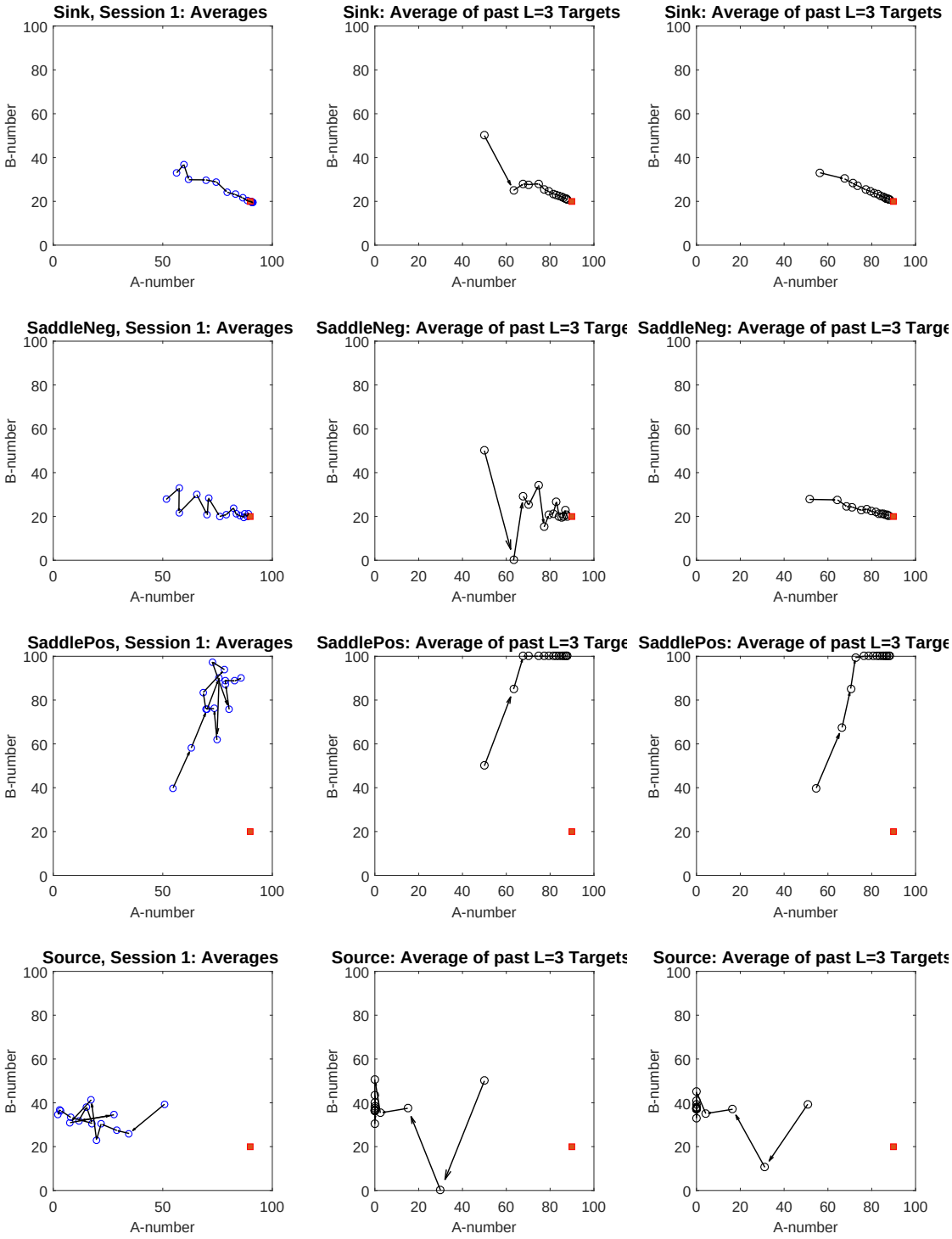


Figure 20: Dynamics of averages in the experiment (*left panels*) as compared with the dynamics of the MAve(3) model. First 15 periods are simulated using the averaged model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experiment (*right panels*).

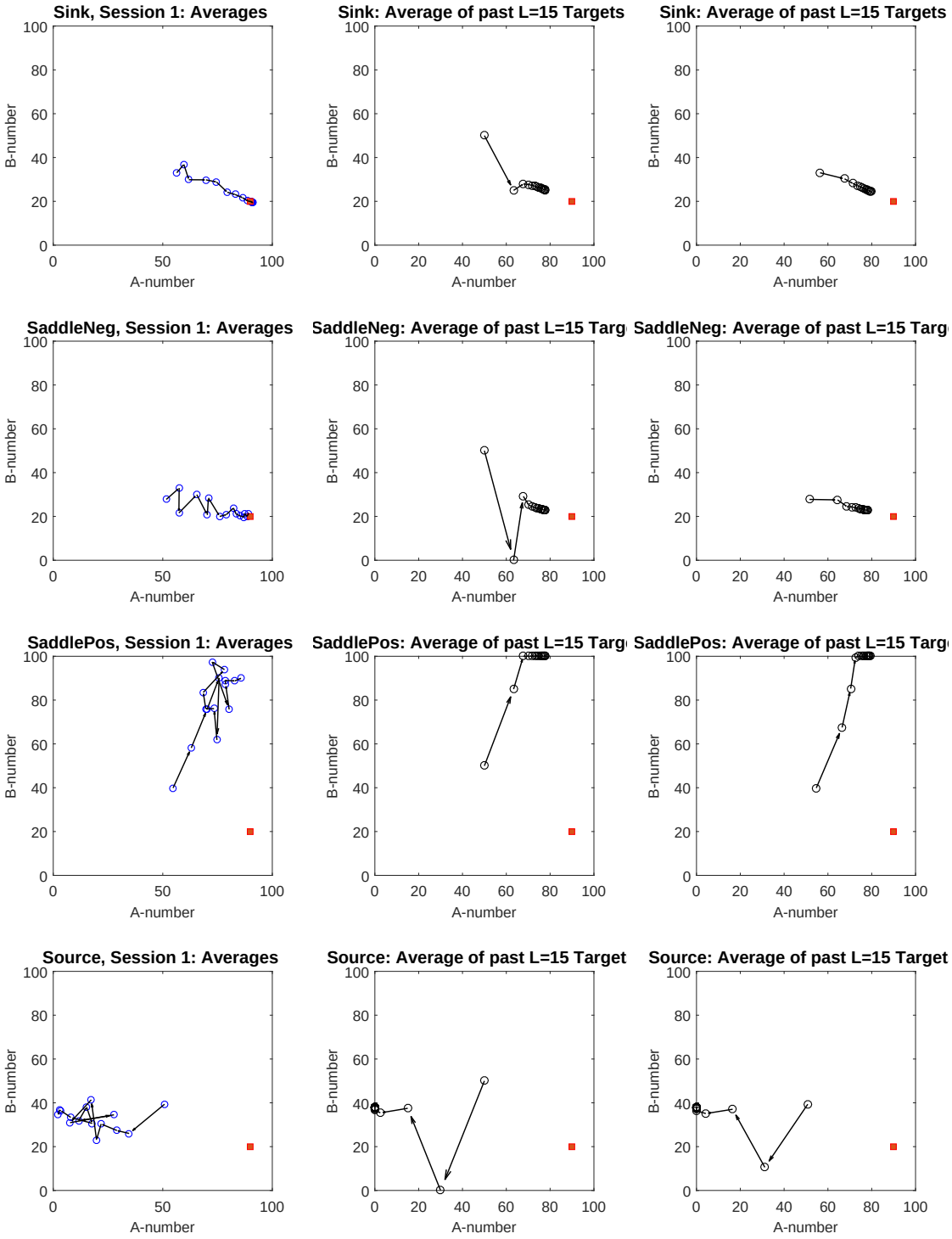


Figure 21: Dynamics of averages in the experiment (*left panels*) as compared with the dynamics of the **AveragedAll** model with $L = 15$. First 15 periods are simulated using the averaged model with initial point at $(50, 50)$ (*middle panels*) and the point observed in the experiment (*right panels*).

Average model as recursive least squares

Generally, sample mean of a variable a_t with $t \geq 1$ can be recursively expressed as

$$\mu_t = \frac{t-1}{t}\mu_{t-1} + \frac{1}{t}a_t = \mu_{t-1} + \frac{1}{t}(a_t - \mu_{t-1})\,.$$

In the average model, the mean of all past target values, $\boldsymbol{\mu}_{t-1} = \sum_{s=1}^{t-1} \boldsymbol{z}_s^*/(t-1)$, is used as the submitted number. Thus the model dynamics can be written as

$$\bar{\boldsymbol{z}}_t = \boldsymbol{\mu}_{t-1} = \boldsymbol{\mu}_{t-2} + \frac{1}{t-1}(\boldsymbol{z}_{t-1}^* - \boldsymbol{\mu}_{t-2}) = \bar{\boldsymbol{z}}_{t-1} + \frac{1}{t-1}\big(\boldsymbol{d} + (\mathbf{M} - \mathbf{I})\,\bar{\boldsymbol{z}}_{t-1}\big)\,. \quad \text{(H.1)}$$

Asymptotically, for large t, system (H.1) 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):

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

The local stability condition of this system can be analyzed using the Jacobian matrix of $\boldsymbol{F}$ at the fixed point (which coincide with the Nash equilibrium of our game). This matrix is, of course, simply $\mathbf{M} - \mathbf{I}$ and the eigenvalues (for lower triangular matrix $\mathbf{M}$ are $m_{11} - 1$ and $m_{22} - 1$. The solution of ODE system is asymptotically stable, if both these values are negative. This is the condition of asymptotic convergence of the average model.

I Online Appendix (Not for publication). Parameter estimation and performance of learning model: robustness check

Models		Parameter	Out-of-sample MSEs Sink	SadNeg	Overall
NE		-	221.16143	97.53310	159.34726
Naïve		-	44.03442	25.00151	34.51797
Average	MAve(2)	-	35.75808	23.14639	29.45224
	MAve(3)	-	36.57404	24.94528	30.75966
	MAve(4)	-	36.29821	29.37378	32.83599
	MAve(5)	-	39.01747	34.03913	36.52830
	EWMA	.66(.06)	37.98032	19.33101	28.65567
Level-k	0 (Stubborn)	-	49.01650	59.20243	54.10947
	1 (Naïve)	-	44.03442	25.00151	34.51797
	2	-	89.36464	153.33727	121.35095
Adaptive	$\lambda_a = \lambda_b$	.61(.04)	32.29442	8.41438	20.35440
	$\lambda_a \neq \lambda_b$	.59(.07) .62(0.03)	32.48933	8.81117	20.65025
Mixed	0 1 (Adaptive)	.39(.04) .61(.04)	32.29442	8.41438	20.35440
	0 1 2	.39(.04) .61(.03) .00(.01)	32.29443	8.41440	20.35441
	1 2	.86(.04) .14(.04)	48.51259	17.87918	33.19589
	0 NE	.77(.03) .23(.03)	38.93502	36.67750	37.80626
	1 NE	1.00(.00) .00(.00)	44.03451	25.00151	34.51801
	0 1 NE	.39(.03) .61(.03) .00(.01)	32.29440	8.41439	20.35439
	0 1 2 NE	.39(.04) .61(.04) .00(.01) .00(.00)	32.29444	8.41439	20.35442
	1 2 NE	.86(.04) .14(.04) .00(.00)	48.51270	17.87917	33.19594
CH-Poisson $\max k = 5$		1.01 (.16)	32.48609	18.95240	25.71925

Table 13: Estimation and performance of various learning models in terms of the out-of-sample one-step-ahead MSE using leave-one-out procedure for **Sink** and **SaddleNeg** data only. MSEs are computed for periods 2 to 15.

Level	**0**	**1**	**2**	**NE**
Periods 1-15	**0.48**	**0.50**	0.03	0.00
	(0.03)	(0.04)	(0.03)	(0.00)
Periods 1	**0.49**	**0.42**	0.09	0.00
	(0.04)	(0.06)	(0.06)	(0.02)
Periods 2-15	**0.39**	**0.61**	0.00	0.00
	(0.04)	(0.03)	(0.01)	(0.00)

Table 14: Mixed model with levels 0, 1 and 2 and NE (internal Nash Equilibrium) for different periods on the data from **Sink** and **SaddleNeg** treatments only. All parameter estimates in bold significant at 5% level.