Supplementary Online Appendix

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

Mikhail Anufriev* John Duffy† Valentyn Panchenko‡

November 22, 2021

This Online Appendix to the paper is divided in 4 sections.

Section F1 collects additional information about our experimental design to those presented in Section 2 of the paper.

Section F2 provides more information about data that we obtained in all 16 experimental sessions and thus supplements Section 3 of the paper. In particular, Section F2.2 discusses the first period data at length.

Section F3 collects additional results of the analysis of the behavioral models and supplements Section 4 of the paper. The average models are analyzed in Section F3.3 where we formulate Result 5.

Section F4 presents specifications of all additional models (i.e., QRE, NI, EWA, and Gill and Prowse model) whose estimation results are given in Section 5. Section F4.5 shows the results of estimation of non-learning models on the first period experimental data.

*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

F1 Additional Information on Experiment

F1.1 Screenshots of the Experiment

Fig. F.1 reproduces information from the instructions that was projected on a large screen in the lab for the duration of each experimental session (with appropriate changes for the different treatments). Figs. F.2 to F.5 show computer screenshots for different stages of the experiment. All examples are from the **SaddlePos** treatment.

Order of Moves in a Period

1. Choose an A-number and a B-number each in [0,100]
2. The Average of all A-numbers and of all B-numbers is computed.
3. Target A* and B* numbers are computed as follows:

A* = 30 + (2/3) x Average of all "A-numbers"
B* = 35 – (1/2) x Average of all "A-numbers" + (3/2) x Average of all "B-numbers".

4. Your points are determined by

$$\text{Points} = \frac{500}{5 + |\text{your "A–number"} - A^*| + |\text{your "B–number"} - B^*|}$$

1 point = 1 cent

Figure F.1: The slide projected on the large screen in the lab during the experiment.

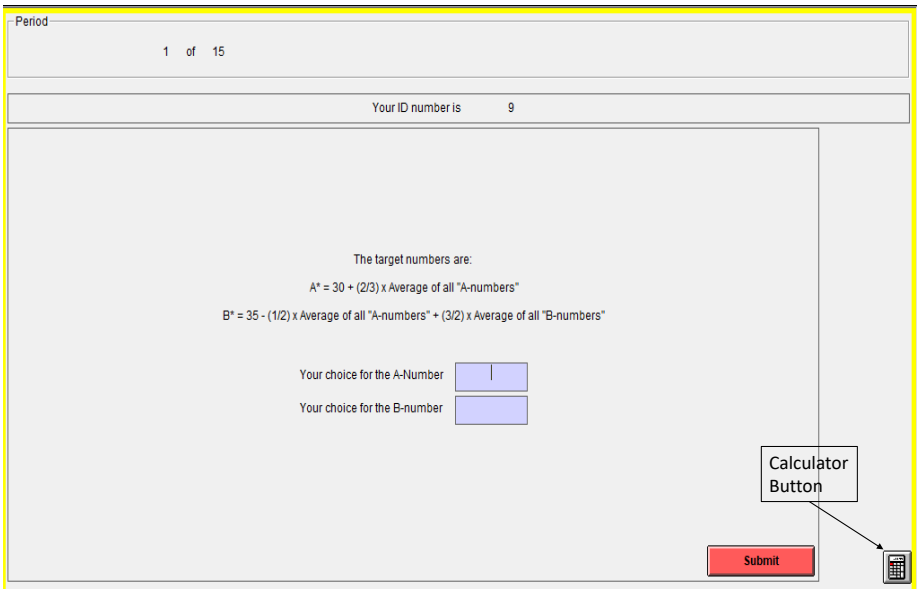

Figure F.2: Decision screen in period 1.

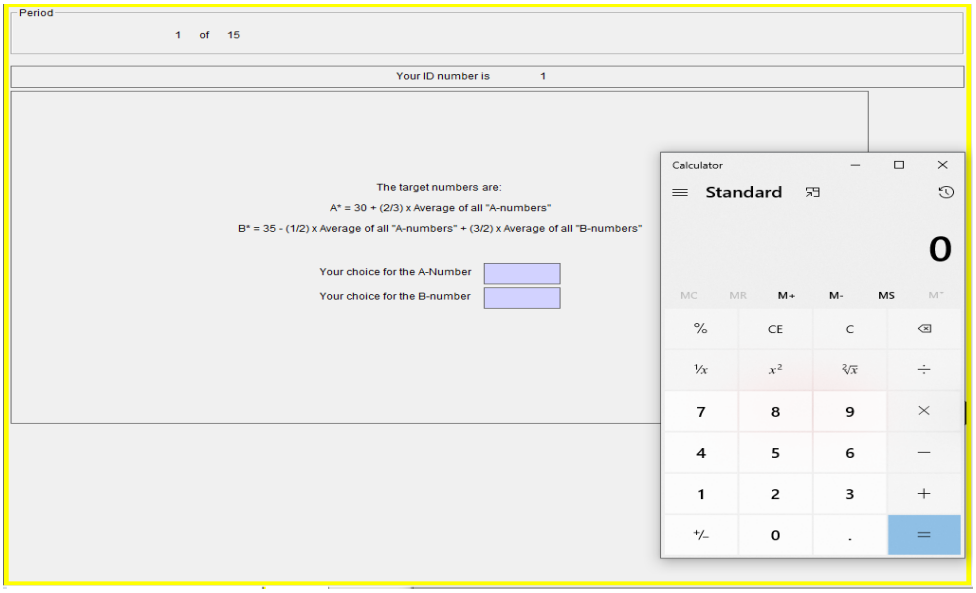

Figure F.3: Calculator display illustration.

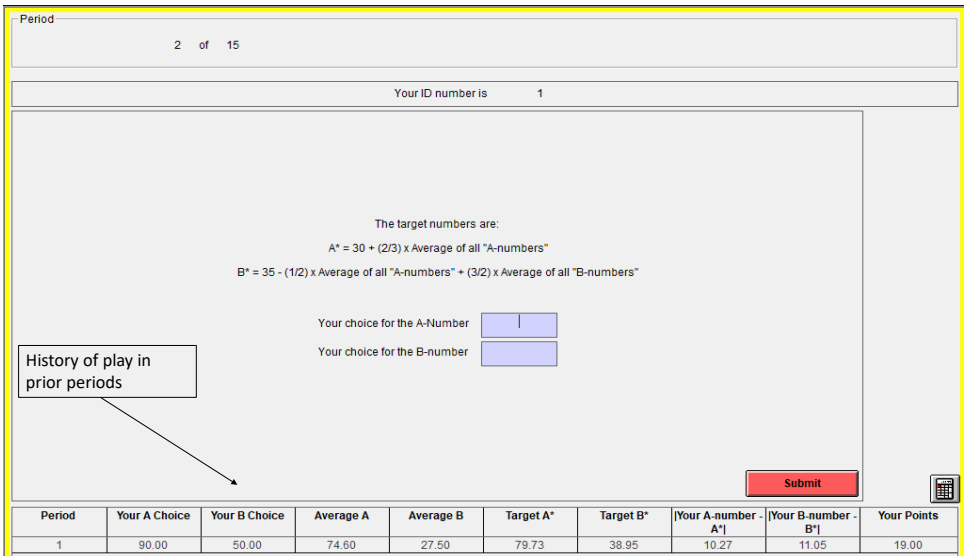

Figure F.4: Decision screen in period 2, showing the history of play in the table.

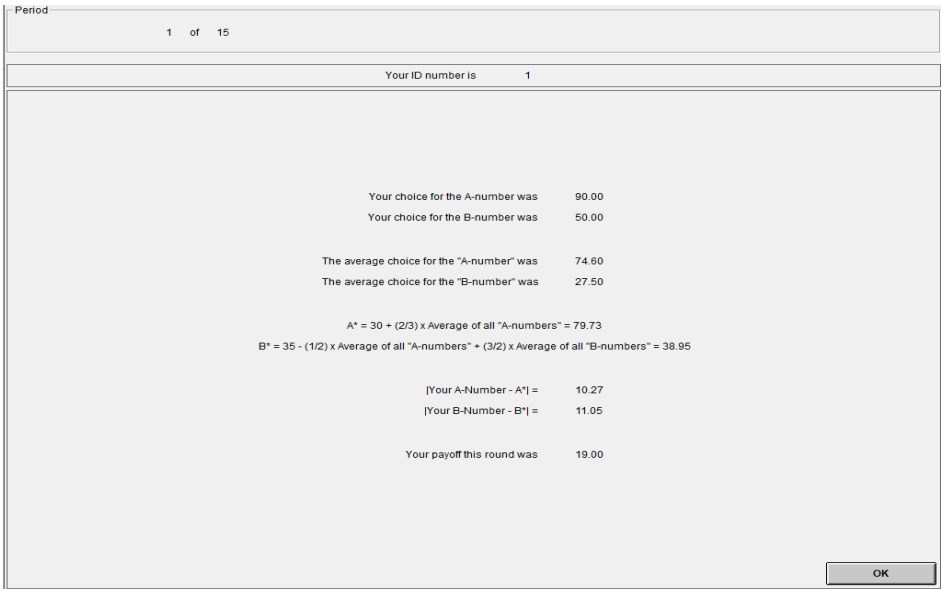

Figure F.5: Results screen displayed between the rounds.

F2 Additional Results of the Experiment

F2.1 Dynamics in all experimental sessions

Figs. F.6 to F.9 show the dynamics of average guesses for each session in treatments **Sink**, **SaddleNeg**, **SaddlePos**, **Source**, respectively. The left panels show the evolution over time of $\bar{a}$ (thick red line) and $\bar{b}$ (thin blue line). The dashed lines in these figures indicate the levels of the PONE, $a^E = 90$ and $b^E = 20$. The middle and right panels show the same evolution as in the phase diagram (i.e., X-Y plots of the a number versus b number over time), with the right panel showing a zoomed-in version of the middle panel.

F2.2 Choices in the first period

In the first period of our experiment, all 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 $m_{11}\bar{a} + d_1$. For this reason, it is interesting to look at subjects' choices in the first period alone.

Figure F.10 presents histograms of individual guesses for the a-number in period 1. The left panel combines all groups from 3 treatments **Sink**, **SaddleNeg** and

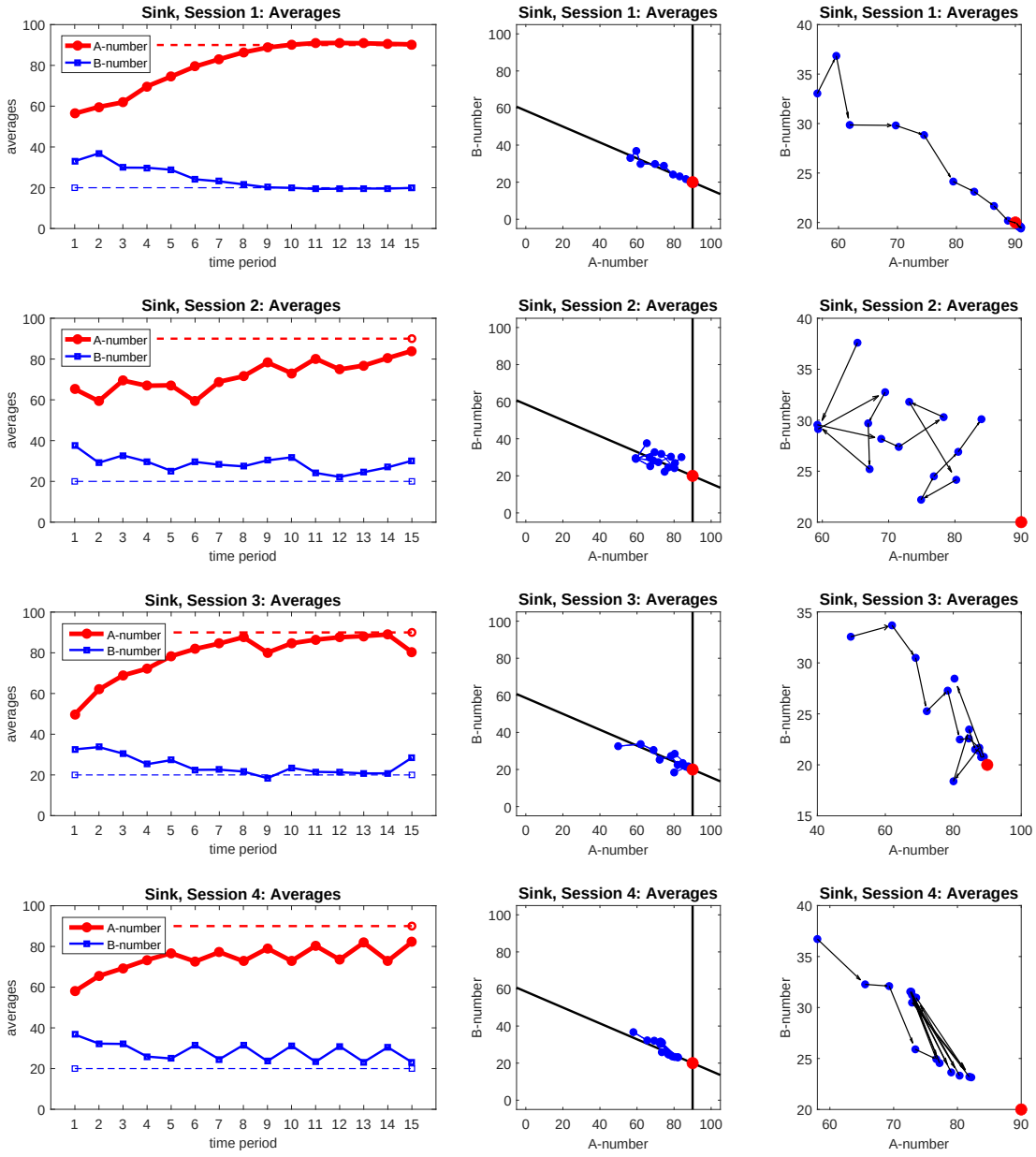

Figure F.6: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **Sink** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

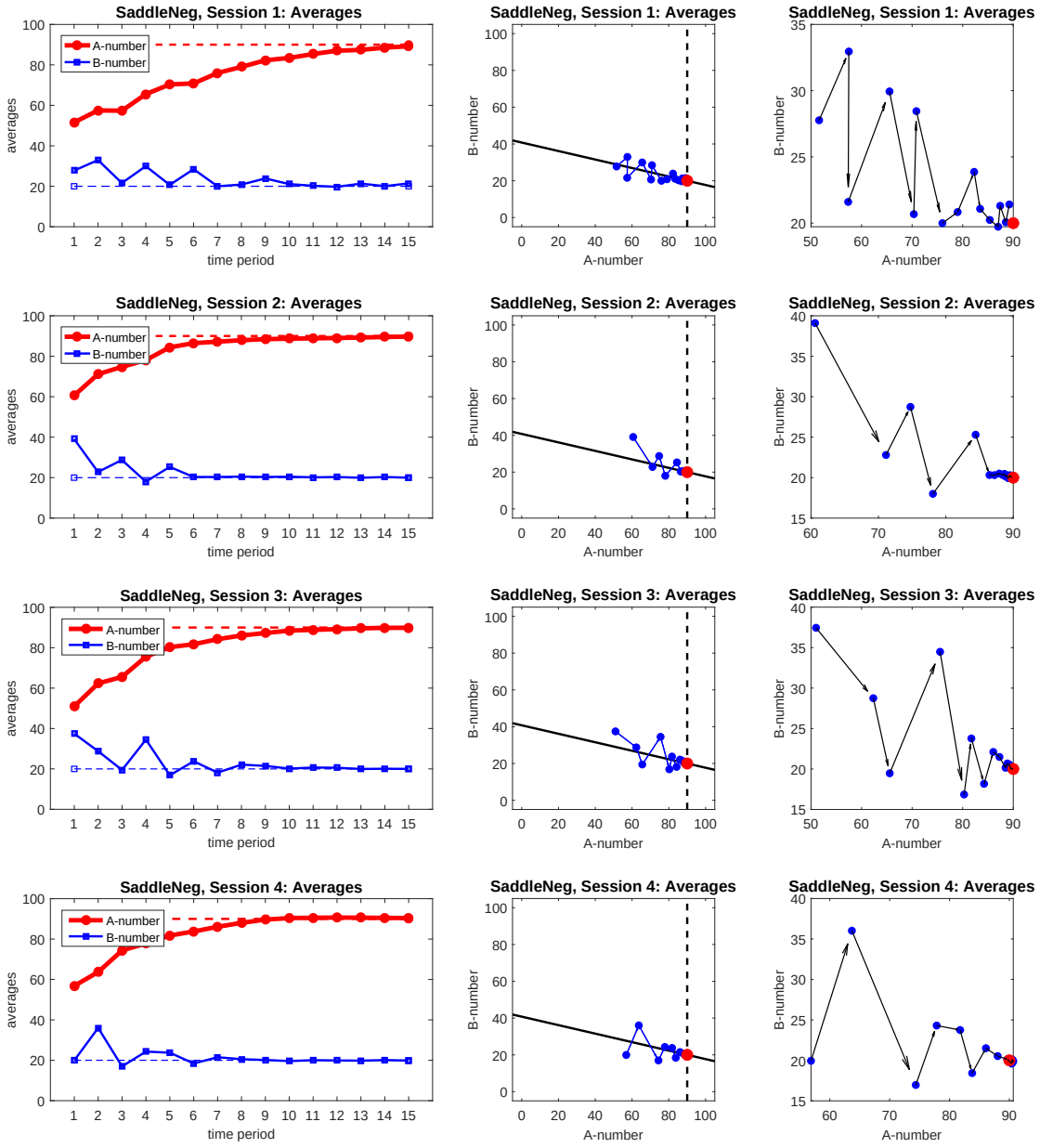

Figure F.7: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **SaddleNeg** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

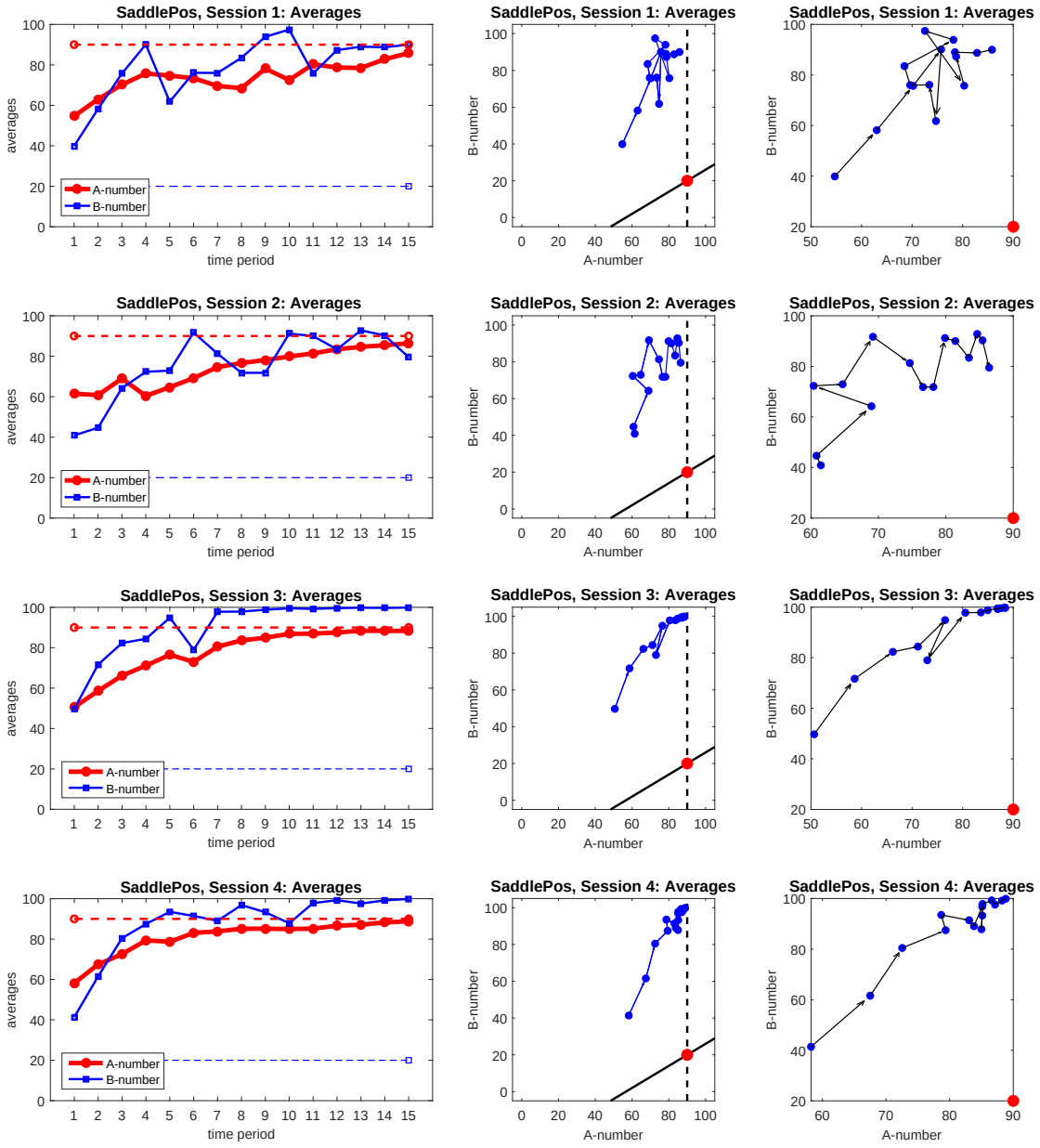

Figure F.8: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **SaddlePos** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

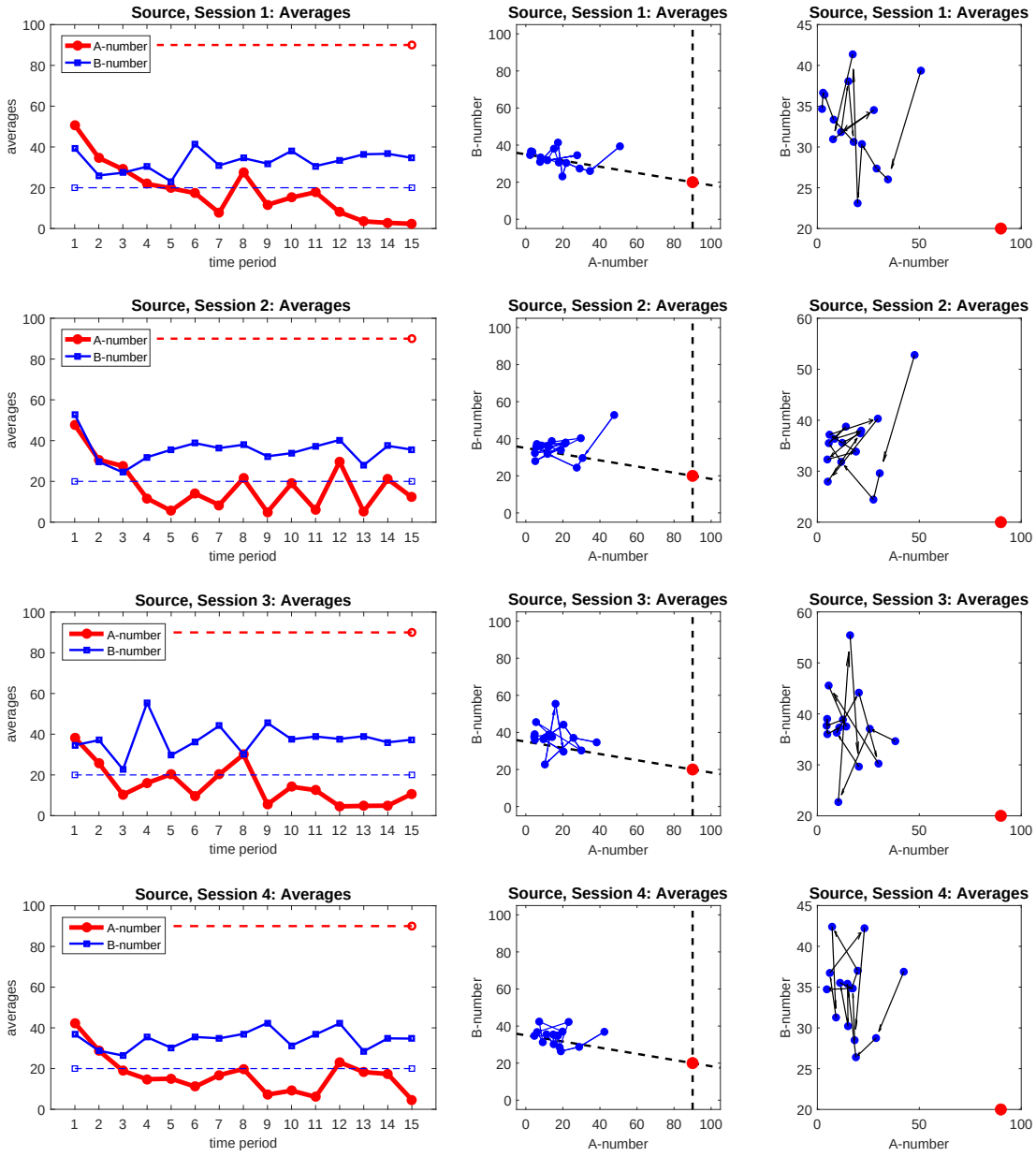

Figure F.9: Dynamics of the average values, $\bar{a}$ and $\bar{b}$, in the **Source** treatment of the experiment. *Left:* Time series. *Middle:* Phase diagrams. *Right:* Phase diagrams (zoomed version).

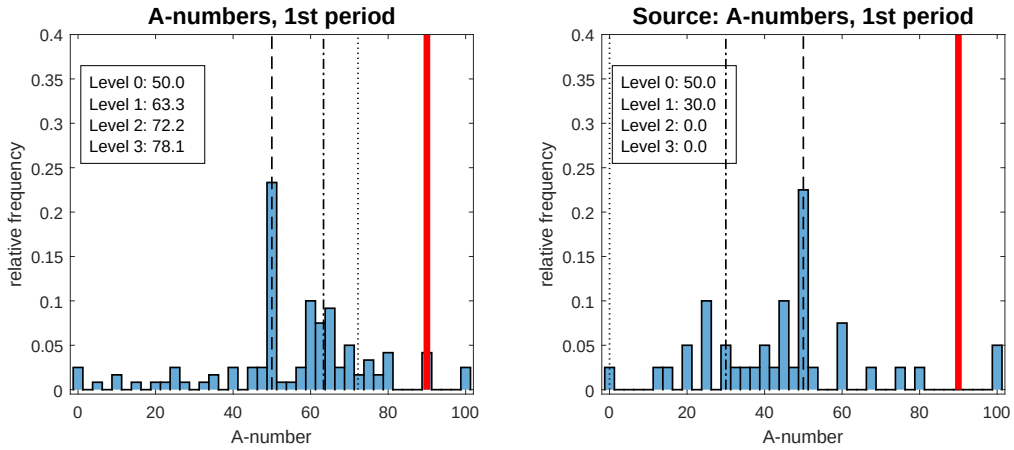

Figure F.10: Frequencies of Period 1 a-choices and levels of reasoning (different dashed lines) and the PONE (solid red line). *Left:* combined choices from **Sink**, **SaddleNeg** and **SaddlePos** treatments. *Right:* **Source** treatment.

SaddlePos, and the right panel shows the histogram for the **Source** treatment.[1] We observe that first period guesses are heterogeneous with a large spike around the middle of the $[0, 100]$ guessing interval. The remaining choices are concentrated in the right half of the interval for the first three treatments, and in the left half for the **Source** treatment. A few participants in the first three treatments submitted the PONE quantity, $a^E = 90$, as represented by the thick vertical line. Figure F.11 shows histograms for b-number choices for the four different treatments. As with the a-numbers, guesses are heterogeneous, there are spikes at 50 in all four treatments, and some participants again submit the PONE quantity, $b^E = 20$. The choices are skewed to the left in all treatments, except for the **SaddlePos** treatment where there are also many choices to the right of the mid-point of the guessing interval.

A standard approach to systematize the choices employs level-k reasoning. According to this classification (see, e.g., Nagel, 1995), level-0 subjects submit guesses distributed uniformly over the guessing interval, having a mean of $(0 + 100)/2 = 50$, whereas other types iteratively play the best response.[2] For the a-number, it means that subjects of level-1 submit $50m_{11} + d_1$, subjects of level-2 best respond to this, and so on. For the b-number, to extend this approach, we make the "lock-step" assumption, that subjects at level-k (with $k > 0$) play a best response to others at *the*

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

[2] When the best responses fall outside of the interval of strategies $[0, 100]$, they are truncated to the closest boundary. An alternative specification for level-0 types has them guessing $m_{11} \times 100 + d$, as 100 is the upper bound of the guessing interval.

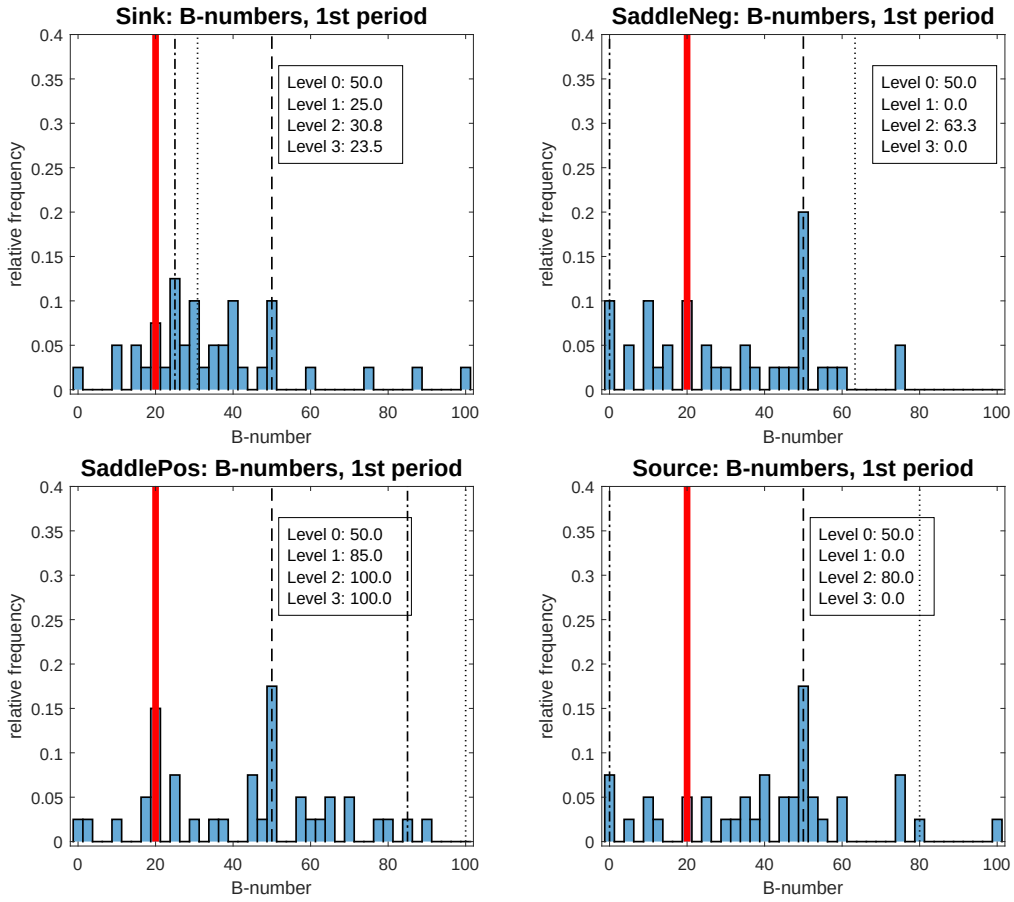

Figure F.11: Frequencies of Period 1 b-choices and levels of reasoning (different dashed lines) and the PONE (red line) for four treatments.

same level $k-1$ for both the a and b numbers.[3] 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.

To illustrate this approach, we superimpose on the histograms in Figs. F.10 and F.11 three vertical lines, corresponding to the levels of 0, 1, and 2 (the legend specifies the corresponding values for these levels and level 3 as well). Note that for the a-number, the sequence of levels is monotonic. It increases and converges to $a^E = 90$ in the three treatments with $m_{11} = 2/3$, and decreases to 0 (staying there for any $k \geq 2$) in the **Source** treatment.[4] For the b-number, the levels converge to $b^E = 20$ only in the **Sink** treatment. The convergence is oscillatory, making it

[3]The lockstep assumption is equivalent to a recursive definition of levels given by Eq. (9) in the main paper.

[4]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 $a_i = 0$.

difficult to identify the actual levels played by subjects, but we do observe spikes around level-1 and level-2 predictions. In the **SaddleNeg** and **Source** treatments, the level-1 choice is 0 (after truncation), where we also see a spike in our data. The level-2 choice is above 50, and the level-3 choice is 0 again, and so on.[5]

Comparing this level-k model with the data for b-choices, we conclude that the presence of a coupled variable in the 2DBC game leads to an even further decrease in the level of rationality for the b-number. Looking back at statistics for the first period choices in Table 2 of the paper, note that in 14 out of 16 sessions, the average a-guess is shifted from the mid-point of 50 towards the first level of rationality, whereas for b-guesses this happens in only 11 out of 16 sessions.[6]

Fig. F.12 shows a scatter plot of individual choices in period 1. The four panels correspond to our four treatments. Every point corresponds to one or more individuals submitting an a-guess as shown on the horizontal axis and a b-guess as shown on the vertical axis. 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 lines indicating various levels of rationality. The thick red lines correspond to the PONE, (a^E, b^E). The other thinner lines indicate various levels of rationality: level-0, is indicated by the dashed lines intersecting at $(50, 50)$, level-1 by the dashed-dotted lines, and level-2 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. F.12 shows that even if the choices are quite dispersed, there are a few clear cases of consistency where participants apply levels 0 or 1 to both of their choices or are able to derive the PONE. For other levels of rationality, we do not observe such consistency.

F2.3 Speed of convergence

In Section 3 of the main paper, we observe that the dynamics converge in the **Sink** and **SaddleNeg** treatments and that the a-guesses converge in the **SaddlePos** treatment. There are some further dynamic features of our experimental data.

[5]This process converges to a two-cycle between 0 and 50 in the **SaddleNeg** and to a two-cycle between 0 and 95 in the **Source** treatment. In the **SaddlePos** treatment, the levels increase monotonically to 100.

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

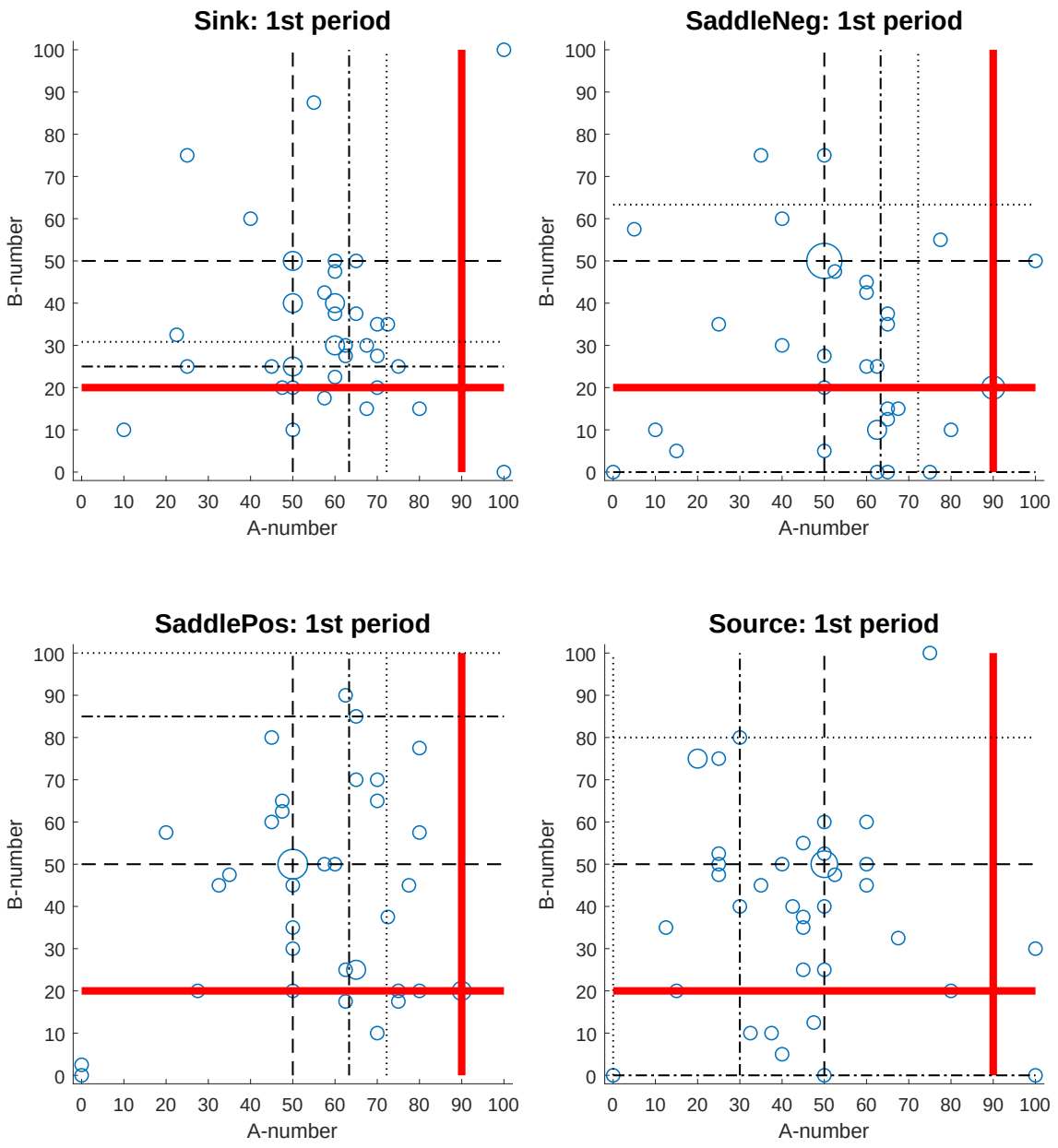

Figure F.12: Frequencies of individual pairs of guesses in period 1 and levels of reasoning (different dashed lines) and the PONE (red thick line) for the four experimental treatments.

First, there is a clear difference in the type of convergence observed. The a-number converges almost monotonically from below in all three treatments. The b-number, however, converges almost monotonically in the **Sink** treatment and through oscillations in the **SaddleNeg** treatments.[7]

[7]There is one exception, Session 4, of the **Sink** treatment, see Fig. F.6. The oscillations in both numbers of that session are due to a single subject.

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

Table F.1: The first period when the trajectory enters the ε-neighborhood of the PONE. Cases where the trajectory never reached the neighborhood are indicated by the '-' symbol.

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 F.2: The latest period when a trajectory enters the ε-neighborhood of the PONE irreversibly, i.e., to stay there until the end of the experiment. Cases where the trajectory was outside of the ε-neighborhood in the last period of the experiment are indicated by the '-' symbol.

Second, there is a difference in the *speed* of convergence. We illustrate this difference in Table F.1 by comparing the "first hit time", i.e., the first instance in which the trajectories for the average a and b numbers enter some ε-neighborhood of the PONE, in each of the four sessions of the **Sink** and **SaddleNeg** treatments. Let us fix $\varepsilon > 0$ and define the neighborhood as an open square around the equilibrium,

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

Let $t(\varepsilon)$ denote the period when the trajectory for the average values of the a and b numbers belong to the ε-neighborhood of the equilibrium for the first time. Formally, $t(\varepsilon)$ is such that $(\bar{a}_{t(\varepsilon)}, \bar{b}_{t(\varepsilon)}) \in U_\varepsilon$ and $(\bar{a}_t, \bar{b}_t) \notin U_\varepsilon$ for any $t < t(\varepsilon)$. Table F.1 shows the first periods defined in this way for all sessions of the **Sink** and **SaddleNeg** treatments[8] for five different values of ε. Cases where the trajectory never reached the neighborhood during the experiment are indicated by the '-' symbol.

In addition to the "first hit time", we made an across treatment comparison of the

[8]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 $\varepsilon \leq 20$.

latest experimental period when the trajectory entered the ε-neighborhood of PONE to stay there until the end of the experiment (so to say, the period of "irreversible entry"). These statistics are only relevant for those sessions where the trajectory entered the neighborhood at least once. 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 $t \geq \tau(\varepsilon)$.

Tables F.1 and F.2 suggest 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 the other convergent **Sink** treatment.[9]

F3 Additional Analysis of Behavioral Models

F3.1 The mixed levels 0-1 model

Section 4.2 presented results for the mixed levels 0-1 model. Here we illustrate and discuss the dynamics of this model. The results are presented in Table 4 of the main text. The dynamics of the mixed levels 0-1 model is written in (11) as

$$\bar{\boldsymbol{z}}_t = (\lambda \mathbf{M} + (1-\lambda)\mathbf{I})\bar{\boldsymbol{z}}_{t-1} + \lambda \boldsymbol{d}\,. \tag{F.1}$$

Proposition F3.1. *Let us assume that matrix* $\mathbf{I}-\mathbf{M}$ *is not invertible and that* $\lambda \neq 0$. *Then there exists a unique steady state for the dynamics* (F.1) *given by* $(\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}$.

Proof. Denote the steady state of the system as $\boldsymbol{z}^*$. Then at the steady state we have

$$\boldsymbol{z}^* = (\lambda \mathbf{M} + (1-\lambda)\mathbf{I})\boldsymbol{z}^* + \lambda \boldsymbol{d} \quad \Leftrightarrow \quad \mathbf{0} = -\lambda(\mathbf{I}-\mathbf{M})\boldsymbol{z}^* + \lambda \boldsymbol{d}\,.$$

As $\lambda \neq 0$, we can simplify the last equality to

$$(\mathbf{I}-\mathbf{M})\boldsymbol{z}^* = \boldsymbol{d} \quad \Leftrightarrow \quad \boldsymbol{z}^* = (\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}\,.$$

This proves the statement. □

Thus, the system (F.1) has a unique steady state coinciding with the PONE. The

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

stability properties of the model depend on the eigensystem of matrix $\widetilde{\mathbf{M}} = \lambda\mathbf{M} + (1-\lambda)\mathbf{I}$. For the case of a triangular $\mathbf{M}$ we have the following result.

Proposition F3.2. *Consider the dynamics of the mixed levels 0-1 model as given by* (F.1) *and assume that matrix* $\mathbf{M}$ *is lower triangular. The steady state of this model,* $\boldsymbol{z}^E = (\mathbf{I}-\mathbf{M})^{-1}\boldsymbol{d}$*, is globally stable if and only if the following conditions are satisfied*

$$-2 < \lambda(m_{11}-1) < 0 \quad \textit{and} \quad -2 < \lambda(m_{22}-1) < 0\,, \tag{F.2}$$

where m_{11} *and* m_{22} *are the diagonal elements of matrix* $\mathbf{M}$*. Matrix* $\widetilde{\mathbf{M}}$ *has the same eigenvectors as matrix* $\mathbf{M}$ *for any* λ*.*

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

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

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

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

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 global stability of the steady state. This proves conditions (F.2).

Let us assume that $\mathbf{v}_1$ is the eigenvector of $\lambda\mathbf{M} + (1-\lambda)\mathbf{I}$, associated with the eigenvalue μ_1. Then $(\lambda\mathbf{M} + (1-\lambda)\mathbf{I})\mathbf{v}_1 = \mu_1\mathbf{v}_1 = (1-\lambda+\lambda m_{11})\mathbf{v}_1$. Simplifying, $\lambda\mathbf{M}\mathbf{v}_1 = \lambda m_{11}\mathbf{v}_1$, leading to $\mathbf{M}\mathbf{v}_1 = m_{11}\mathbf{v}_1$. Thus, $\mathbf{v}_1$ is the eigenvector of $\mathbf{M}$ associated with the eigenvalue m_{11}. A similar statement can be proven for the vector $\mathbf{v}_2$. □

These statements lead to the results reported in Section 4.1 of the paper for the homogeneous level-1 model, see Table 3. Indeed, this model is a special case of the model above where $\lambda = 1$ so that $\widetilde{\mathbf{M}} = \mathbf{M}$. We also obtain the results in Section 4.2 of the paper for the mixed levels 0-1 model and, in particular, Table 4.

Fig. F.13 compares the dynamics for the mixed levels 0-1 model in all four treatments, when $\lambda = 0.1$, $\lambda = 0.5$, and $\lambda = 0.9$, in the left, middle, and right panels, respectively. For each simulation, we take the average values in period $t = 1$ of all groups in the corresponding treatment (these average values are reported in Table 2 of the main text) as initial conditions.

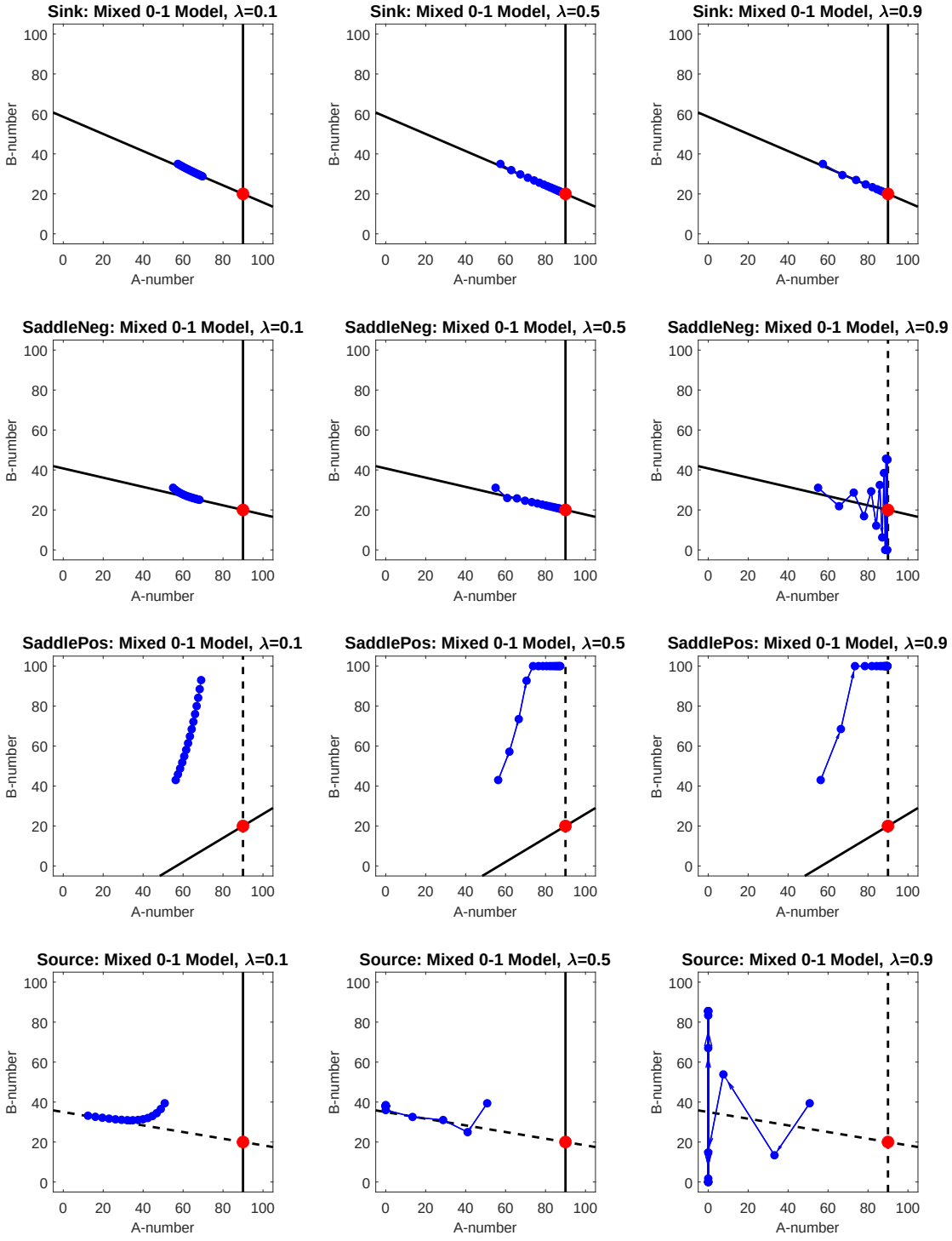

Figure F.13: Phase diagrams of $\boldsymbol{z}_t$-dynamics in the mixed levels 0-1 model for different λ's: $\lambda = 0.1$ (*left panels*), $\lambda = 0.5$ (*middle panels*), and $\lambda = 0.9$ (*right panels*). For all three models, 15 periods are simulated with the initial point given by the a and b-averages of all first period guesses in the corresponding treatment. The black lines are the eigenvectors, stable (solid) and unstable (dashed).

Treatment	Eigenvalues		Attractor and dynamics type		Data	
	μ_1	μ_2	for a	for b	for a	for b
Sink	$f_1\frac{2}{3}+f_2\frac{4}{9}$	$-f_1\frac{1}{2}+f_2\frac{1}{4}$	a^E monotone	b^E oscillates for $f_1>1/3$	✔	✔
SaddleNeg	$f_1\frac{2}{3}+f_2\frac{4}{9}$	$-f_1\frac{3}{2}+f_2\frac{9}{4}$	a^E monotone	b^E for $1/3<f_1<13/15$ oscillates for $f_1>3/5$	✔	✔ $f_1\in(\frac{3}{5},\frac{13}{15})$ ✘ otherwise
SaddlePos	$f_1\frac{2}{3}+f_2\frac{4}{9}$	$f_1\frac{3}{2}+f_2\frac{9}{4}$	a^E monotone	0 or 100† monotone	✔	✔
Source	$f_1\frac{3}{2}+f_2\frac{9}{4}$	$-f_1\frac{3}{2}+f_2\frac{9}{4}$	0 or 100† monotone	38‡ for $1/3<f_1<13/15$, oscillates for $f_1>3/5$	✔	✔ $f_1\in(\frac{1}{3},\frac{13}{15})$ ✘ otherwise

† depending on the initial conditions

‡ assuming that the a-number dynamics converges to 0

Table F.3: Properties of the mixed levels 1-2 models for four experimental treatments. We report two eigenvalues, the attractor of the model when guesses are truncated as in Eq. (10) of the paper, and the type of dynamics (monotone or oscillatory). The last two columns verify whether the dynamics match the experimental data (✔) or not (✘).

We observe that, as λ increases, three changes in the dynamics of the model occur. First, the dynamics in the **SaddleNeg** treatment change from converging (for $\lambda = 0.1$ and $\lambda = 0.5$) to diverging. Relatedly, the dynamics in the **Source** treatment become diverging along the vertical dimension corresponding to the b-number. Thus, for a high value of λ (such as $\lambda = 0.9$), the dynamics in these treatments are inconsistent with the experimental data. Second, the speed of convergence in the **Sink** treatment and the speed of divergence in the **SaddlePos** and **Source** treatments) increases with increases in λ. In fact, for small $\lambda = 0.1$ the model is clearly too slow when compared with the experimental data. Third, the dynamics in the **Sink**, **SaddleNeg** and **Source** treatments change from monotone to oscillating as λ increases. In the data, these oscillations are clearly seen in the **SaddleNeg** treatment.

Taken together, these three observations suggest that the mixed level 0-1 model is consistent with our data *only* for intermediate values of λ, as Result 2 of the paper establishes.

F3.2 Mixed level-k models

In the main text, we introduced the mixed level-k model that generalizes the mixed levels 0-1 model. The eigenvalues, μ_1 and μ_2, of the mixed level-k model are the convex combinations of the k-th powers of the eigenvalues of matrix $\mathbf{M}$, weighted by the corresponding fractions, f_k, as follows from Eq. (12) of the paper.

To illustrate Result 3 of the paper, Table F.3 presents the properties of the mixed

Treatment	Eigenvalues μ_1	Eigenvalues μ_2	Attractor and dynamics type for a	Attractor and dynamics type for b	Data for a	Data for b
Sink	$f_0 + \frac{2f_1}{3}$	$f_0 - \frac{f_1}{2}$	a^E monotone	b^E osc. for $\lambda > 2/3$	✔	✔
SaddleNeg	$f_0 + \frac{2f_1}{3}$	$f_0 - \frac{3f_1}{2}$	a^E monotone	b^E for $\lambda < \frac{2}{5} \cdot \frac{2-f_E}{1-f_E}$ osc. for $\lambda > 2/5$	✔	✔ $\lambda \in \left(\frac{2}{5}, \frac{2}{5} \cdot \frac{2-f_E}{1-f_E}\right)$ ✘ otherwise
SaddlePos	$f_0 + \frac{2f_1}{3}$	$f_0 + \frac{3f_1}{2}$	a^E monotone	0 or 100† for $\lambda > 2f_E/(1-f_E)$; monotone	✔	✔ $\lambda > \frac{2f_E}{1-f_E}$ ✘ otherwise
Source	$f_0 + \frac{3f_1}{2}$	$f_0 - \frac{3f_1}{2}$	0 or 100† for $\lambda > 2f_E/(1-f_E)$; monotone	38‡ for $\lambda < \frac{2}{5} \cdot \frac{2-f_E}{1-f_E}$ osc. for $\lambda > 2/5$	✔ $\lambda > \frac{2f_E}{1-f_E}$ ✘ otherwise	✔ $\lambda < \frac{2}{5} \cdot \frac{2-f_E}{1-f_E}$ ✘ otherwise

† depending on the initial conditions
‡ assuming that the a-number dynamics converges to 0

Table F.4: Properties of the mixed levels 0-1-E model with fractions f_0, f_1 and f_E. The conditions are written using the re-parametrization, $f_1 = \lambda(1-f_E)$ and $f_0 = (1-\lambda)(1-f_E)$. See the caption of Table F.3 for details.

levels 1-2 model, with agents of levels 1 and 2 present in proportions f_1 and $1 - f_1$, respectively. In the **SaddleNeg** treatment, μ_2 is the convex combination of $-3/2$ and $(-3/2)^2 = 9/4$. For $3/5 < f_1 < 13/15$, we have $-1 < \mu_2 < 0$ and the dynamics of the model converges to the PONE via oscillations.

Table F.4 presents the properties of the mixed levels 0-1-E model, where the equilibrium type is added to the mixed levels 0-1 model. We include it here as that model, when estimated against the actual experimental data, turned out to be the winner of the models contest, see Section 5 of the paper.

To facilitate a comparison between the 0-1 and 0-1-E models, we re-introduce parameter λ, the fraction of level-1 types among all non-equilibrium agents, and we reparametrize the 0-1-E model by $f_0 = (1-\lambda)(1-f_E)$ and $f_1 = \lambda(1-f_E)$. We find that all features of the experimental data are reproduced when

$$\max\left\{\frac{2}{5}, \frac{2f_E}{1-f_E}\right\} < \lambda < \frac{2}{5} \cdot \frac{2-f_E}{1-f_E}. \tag{F.3}$$

This is a generalization of Result 2 of the paper, as inequality (F.3) reduces, when $f_E = 0$, to $\lambda \in (\frac{2}{5}, \frac{4}{5})$. When the fraction of the equilibrium type increases, the interval in (F.3) shifts to the right and the condition to match the data becomes more stringent. The intuition is that the equilibrium type pushes the system to the PONE faster. Thus, to reproduce both the oscillatory and diverging dynamics, it is necessary to counteract the equilibrium types with a relatively high frequency of level-1 types, i.e., with a higher value for λ.

F3.3 Average Models

In this section, we discuss models that *average* past information that was available to the participants in the experiment. The mixed levels 0-1 model (adaptive expectations) effectively takes into account all past target values. By rewriting Eq. (11) of the paper recursively, we obtain for any $t \geq 2$:

$$\bar{\boldsymbol{z}}_t = \lambda \left(\boldsymbol{z}^*_{t-1} + (1-\lambda)\boldsymbol{z}^*_{t-2} + \ldots + (1-\lambda)^{t-2}\boldsymbol{z}^*_1 \right) + (1-\lambda)^{t-1}\bar{\boldsymbol{z}}_1 \,. \tag{F.4}$$

Model (F.4) is called the exponentially weighted moving average (EWMA) model, because the past target values have exponentially declining weights. Result 2 of the main text establishes the λ's for which the EWMA is consistent with our experimental data.

An alternative model, the moving average model MAve(L), averages the previous L targets with *equal* weights.[10] When $t > L$, the average guesses given by this model are[11]

$$\bar{\boldsymbol{z}}_t = \frac{1}{L}\left(\boldsymbol{z}^*_{t-1} + \ldots + \boldsymbol{z}^*_{t-L} \right) \,. \tag{F.5}$$

When $L = 1$, this model coincides with the homogeneous level-1 model. Result 1 of the paper states that the model dynamics are not consistent with the features of the experimental data. In general, the dimensionality of the system governing the dynamics is $2L$ to account for the both a and b-lagged variables. As this dimension grows with L, we have to rely on numerical computations to establish the model properties for larger L. For $L = 2$ and $L = 3$, we derive general functional forms for the eigenvalues of the system and compute their absolute values in each treatment. Based on these computations, convergence is achieved in the **Sink** and **SaddleNeg** treatments only, as in the experimental data. For $L > 3$ we simulate the MAve(L) model and find that the converging properties are consistent with the data.

Further, the case $L \to \infty$, is studied analytically, using the principle of E-stability (see Evans and Honkapohja, 2001). The guesses of the MAve(L) model are given by vector $\bar{\boldsymbol{z}}_t = \sum_{s=1}^{L} \boldsymbol{z}^*_{t-s}/L$. As $L \to \infty$, we can re-index the model to have $\bar{\boldsymbol{z}}_t =$

[10] 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 in every period t, see Eq. (F.4). The MAve(L) model assigns smaller and smaller weights to the most recent observation, see Eq. (F.5), and so this model is known as the model with *decreasing* gain, which approximates recursive least squares learning.

[11] If the window L is larger than the available data at a given time period, then the average is computed over a shorter window of all available observations at the time period.

$\sum_{s=1}^{t-1} \boldsymbol{z}^*_{t-s}/t$, and study the case of $t \to \infty$. The following recursive relation holds

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

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

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

The local stability conditions of this system depend on the Jacobian matrix of map $\boldsymbol{F}$ at the fixed point, which is the PONE. This Jacobian matrix is $\mathbf{M} - \mathbf{I}$ and its eigenvalues are $m_{11} - 1$ and $m_{22} - 1$. This system is asymptotically stable if both of these values are negative. This holds only in two treatments, **Sink** and **SaddleNeg**. Thus the MAve(L) model with large L is consistent with the major features of our experimental data.

We illustrate the MAve(L) model dynamics via simulations performed for this model for $L = 2$, $L = 3$ and $L = 14$; see the left, middle and right panels in Fig. F.14, respectively. The simulations are initialized by the values for the first period only, and we take the averages in the first period of all groups in the corresponding treatment as these initial values. For the later periods, when $t > L$ equation (F.5) is used, whereas when $t \leq L$, all available observations are equally weighted, i.e., $\bar{\boldsymbol{z}}_t = (\boldsymbol{z}^*_{t-1} + \cdots + \boldsymbol{z}^*_1)/(t-1)$.

There is a visible discrepancy in the dynamics between the data and the model. When $L = 2$ and $L = 3$, convergence in the **Sink** and **SaddleNeg** treatments occurs in the model more quickly than in the experiment and in a more orderly way (e.g., without the regular oscillations in the **SaddleNeg** treatment). This is especially visible during the first 5 periods. Divergence in the **SaddlePos** and **Source** cases is also quicker, i.e., the dynamics reach the boundary in fewer steps in the model as compared with the data. When $L = 14$, the convergence path is even less similar to the data. Notice that despite a quick start, the simulations do not reach the PONE in the **Sink** and **SaddleNeg** treatments in the 15 periods of the experiment. The convergence turns out to be very slow for high L, because of the *decreasing* gain of the MAve(L) model. New observations get lower weight as time goes on and this delays the incorporation of new information. This analysis and simulations lead to the following conclusion.

Result 5. *The MAve(L) model with $L \geq 2$ reproduces features 2-5 of the data. However, with larger L convergence becomes much slower in the model than in the experiment.*

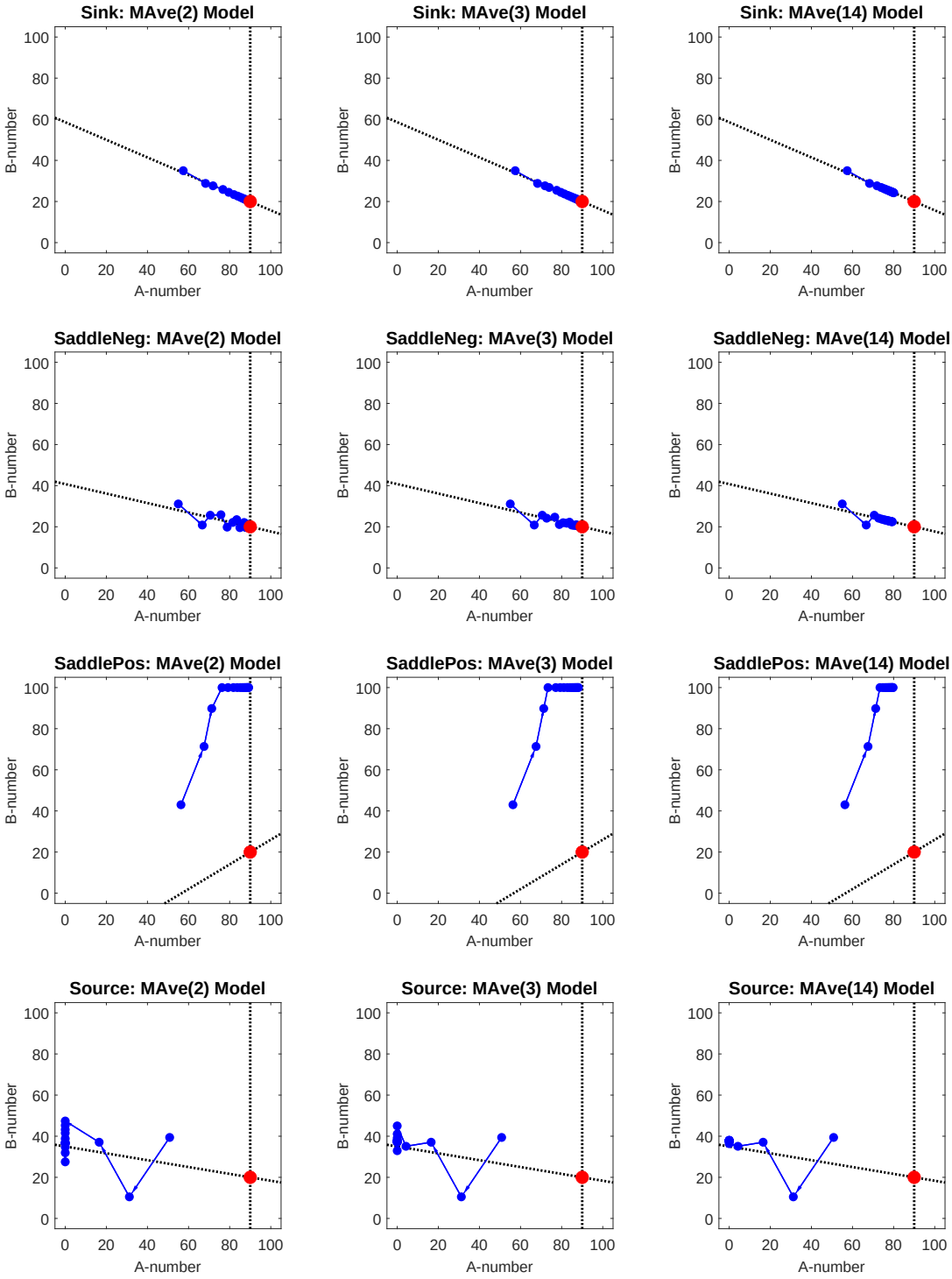

Figure F.14: Phase diagrams of $\boldsymbol{z}_t$-dynamics in the average models with 2 lags (*left panels*), 3 lags (*middle panels*) and 14 lags (*right panels*). For all three models, 15 periods are simulated with the initial point given by the a and b-averages of all first period guesses in the corresponding treatment. The dotted lines are the eigenvectors from the homogeneous level-1 model.

The fact that averaging of past targets generates convergence in the **SaddleNeg** but not in the **SaddlePos** treatment is intuitive. In the **SaddleNeg** treatment, the dynamics for the b-number are initially oscillating around the PONE, and averaging dampens these oscillations, leading to convergence. By contrast, in the **SaddlePos** treatment, deviations from PONE are all in the same direction, so that past averaging does not work.

F4 Additional Specifications and Estimation Results

F4.1 Quantal Response Equilibrium

The quantal response equilibrium (QRE) model introduced by McKelvey and Palfrey (1995) is a popular model to describe agents' actions in unrepeated games (Crawford et al., 2013). The model assumes that each subject has a noisy best response to the actions of others that are noisy as well.

We estimate the 'logit QRE', that is, the QRE model where noise is from a logistic distribution. Similarly to Camerer et al. (2004), we estimate the model separately on data for the first period only, and then on the pooled data for the rest of the periods. We suppress index t of a time period and consider an agent i playing the strategy $s_i^j = (a_i^j, b_i^j)$, where j is used as a counter of all possible strategies. In our experiment, participants choose both a and b numbers between 0 and 100 with up to 2 digits, resulting in $10,001$ strategies for each number and $10,001 \times 10,001$ strategies in total. According to (3), the profit of strategy s_i^j, given the strategies played by the others s_{-i}, is

$$\pi(s_i^j|s_{-i}) = \frac{500}{5 + |a_i^j - a^*| + |b_i^j - b^*|}\,, \tag{F.6}$$

where the targets a^* and b^* are treatment-specific linear functions of the averages of all a and b guesses that depend on s_{-i}.

Let us define the *(logit) quantal best response* of player i, $QBR_i(p_{-i}; \lambda)$ as a mixed strategy, p_i, with the probability of playing s_i^j given by

$$p_i(s_i^j) = \frac{\exp[\lambda \cdot \bar{\pi}(s_i^j|p_{-i})]}{\sum_{s_i'} \exp[\lambda \cdot \bar{\pi}(s_i'|p_{-i})]}\,, \tag{F.7}$$

where $\lambda \geq 0$ is the so-called *precision parameter* and $\bar{\pi}$ is the expected payoff of

playing s_i^j. The *logit QRE* is a mixed strategy profile p^* such that $p_i^* = QBR_i(p_{-i}^*; \lambda)$. Similar to Breitmoser (2012), we focus on the "principal" branch of the QRE. We start the numerical computation with the $\lambda = 0$ case, where the unique QRE is the uniform distribution over all strategies. At each step, we increase λ by 0.001 and compute the QRE distribution starting from the QRE p^* of the previous λ. Since the size of the strategy space is 10^6, to make the computations feasible, in computing the expected profit $\bar{\pi}$, we assume that the targets are derived from the expected values of the a and b numbers based on the distribution p. This assumption is reasonable for a sufficiently large number of players.

Parameter λ is estimated by maximizing the log-likelihood $\sum_i \sum_t \log p_{i,t}^*(\boldsymbol{z}_{i,t})$, where $\boldsymbol{z}_{i,t}$ is the vector of actual a and b guesses by individual i in period t in the experiment. To make the estimation computationally viable, we use a grid search for parameter λ.

F4.2 Noisy Introspection

The noisy introspection (NI) model introduced by Goeree and Holt (2004) combines the idea of iterative reasoning with noisy beliefs. Players believe that their opponents have noisier strategies, the opponents of opponents have even noisier strategies and so on. After several such iterations, the process converges to fully uninformed beliefs, i.e., to a uniform distribution over the strategy space. To model the increasing level of noise, let us define $\lambda_m = \lambda \cdot \mu^m$, where $0 \leq \mu < 1$ is the 'telescoping' parameter determining how fast the noise is increasing with each iteration of reasoning m. The NI model involves a mixed strategy p_0 defined iteratively from $p_m = QBR_m(p_{m+1}; \lambda_m)$, which is the right side of (F.7), where $\bar{\pi}$ is computed in the same way as for the QRE. The maximum m is set to 100.

Parameters λ and μ are estimated by maximizing the log-likelihood function $\sum_i \sum_t \log p_{0,i,t}(\boldsymbol{z}_{i,t})$, where as before $\boldsymbol{z}_{i,t}$ is the vector of actual individual guesses in the experiment. We again use a grid search for λ with a step size of 0.001 as well as for μ with a step size of 0.05.

F4.3 Experience-Weighted Attraction Model

Camerer and Ho (1999) proposed the Experience-Weighted Attraction (EWA) learning model. In this model, strategies are associated with attractions that reflect initial predispositions that are updated over time based on experience. The model incorpo-

rates both reinforcement and belief-based learning and has been widely used in the literature for modeling human behavior across a number of different games, including the BC game.

As there are $10,001$ strategies for each number in our 2DBC, for computational feasibility, we shall assume that the participants assess strategies for the a and b numbers separately. Let $s_{i,t} = (a_{i,t}, b_{i,t})$ be the actual strategy that participant i played in period t. When hypothetical strategies for the a-number are assessed in period t, we take $s_i^j = (a_i^j, b_{i,t})$ where a_i^j is indexed by j (with $j = 1, \ldots, 10,001$). Similarly, when hypothetical strategies for the b-number are assessed, $s_i^j = (a_{i,t}, b_i^j)$.

Let $A_{i,t}^j$ be an attraction level assigned by individual i to strategy s_i^j in period t. We adopt the parametrization from Camerer et al. (2002).[12] For any $t \geq 2$, the attraction level is defined recursively

$$A_{i,t}^j = \frac{\phi \cdot N_{t-1} \cdot A_{i,t-1}^j}{N_t} + \frac{[\delta + (1-\delta) \cdot I(s_i^j, s_{i,t})] \cdot \pi_i(s_i^j | s_{-i,t})}{N_t},$$
$$N_t = (1-\kappa) \cdot \phi \cdot N_{t-1} + 1,$$

where the indicator I takes the value of 1 when s_i^j is actually played at t, and 0 otherwise. Parameter δ determines the relative weight of the hypothetical payoff with respect to the actual payoff, ϕ is the decay rate of past observations, and κ is the accumulation rate of attractions that determines how fast strategies lock-in.[13] The payoff $\pi_i(s_i^j | s_{-i,t})$ is calculated according to (F.6) with $s_{-i,t}$ being the actual strategy profile played by the other players at t, and where, e.g., $b_i^j \equiv b_{i,t}$, if strategy s_i^j is about the a-number guess. Strategy s_i^j is then played by individual i in period $t+1$ with probability

$$p_{i,t+1}(s_i^j) = \frac{\exp[\lambda \cdot A_{i,t}^j]}{\sum_k \exp[\lambda \cdot A_{i,t}^k]}, \tag{F.8}$$

where λ is the sensitivity or precision parameter.

We initialize the attractions from the experimental data of period 1 separately for each session, following the approach of Camerer et al. (2002), p. 154, Eqs. (7) and

[12] This parametrization for the EWA model became more common in the literature. It differs from the original parametrization in Camerer and Ho (1999). In that paper, the authors use parameters δ and ϕ as we do here, but instead of parameter κ, they have $\rho = (1-\kappa)\phi$ measuring the rate of decay for experience.

[13] The parameter estimates for this model using our experimental data (Table 6 in Section 5) indicate the following. Participants weight their actual experiences higher than the hypothetical payoffs ($\delta = 0.65$) reflecting a higher reliance on reinforcement learning relative to belief-based learning. They forget past information relatively fast ($\phi = 0.33$), and they average past attractions instead of cumulating them ($\kappa = 0$), avoiding in this way, the lock-in effect.

(8). That is, we assign all strategies observed in period 1 with zero frequencies, the attractions $A^j_{i,1} = 0$. Then we match the actual frequencies of the played strategies with the initial attractions, $A^j_{i,1}$ so that (F.8) is satisfied.

The parameters are estimated by maximizing log likelihood

$$\sum_i \sum_t \big(\log p^a_{i,t}(a_{i,t}) + \log p^b_{i,t}(b_{i,t}) \big),$$

where $p_{i,t}$ for a and b-number are in (F.8) and $a_{i,t}$ and $b_{i,t}$ are the actual guesses in the experiment.

F4.4 Gill and Prowse model

Gill and Prowse (2016) propose a structural mixed level-k learning model. The model assumes the *fixed level-k types*, who maintain their level throughout the experiment, as well as the *learning types*, who increase their levels with experience. The individual first period choices in the experiment are used as the initial conditions of the model. A learning type $k \to k+1$ plays, at each period from $t = 2$ to 15, a mixed strategy assigning probability $\frac{15-t}{13}$ to the level-k choice and probability $\frac{t-2}{13}$ to the level-$k+1$ choice.

In what follows x denotes one of $10,001$ possible permissible values (from 0 to 100 with up to 2 digits) either for the a or for b-number. The Gill-Prowse model explicitly specifies the error structure so that the choices are sampled from discretized and truncated Student's t distributions. The probability of a specific choice x by participant i using level-k in period t is given by[14]

$$p_{i,t}(x|k) = \frac{\mathcal{T}(x; \mu_{i,t}(k), \sigma_t(\mu), \nu)}{\sum_{s=0}^{10,000} \mathcal{T}(s; \mu_{i,t}(k), \sigma_t(\mu), \nu)},$$

where $\mathcal{T}$ is the CDF of the Student's t distribution with mean $\mu_{i,t}(k)$ equal to the deterministic homogeneous level-k choice for the a or b-number as in Eq. (9) in Section 4.1 of the paper. Parameter ν denotes the degrees of freedom of this distribution, whereas the scale parameter is specified as

$$\sigma_t(\mu) = \exp\left(\alpha + \beta \cdot I_{\text{PONE}} - \delta \cdot \frac{t-2}{13} \right),$$

[14]The Gill-Prowse model treats choices $x = 100$ as special cases outside of level-k model. We do not need this since in our divergent treatments, choices of 100 may be consistent with the model.

where I is indicator showing whether $\mu_{i,t}(k)$ is within $\pm$ 1 of the PONE value. The time and choice-dependent variance allows us to model convergence in the experimental data. The variance tends to reduce over time and to be smaller near the PONE (i.e., $\delta > 0$, $\beta < 0$).

Let f_k and $f_{k\to k+1}$ denote the probabilities (fractions) of the fixed level-k types and the learning level-$k \to k+1$ types, respectively, and K is the highest level. Note that in some specifications we also include a fraction f_E of equilibrium players as a fixed type. The fractions of types as well as the parameters α, β, δ and ν are estimated by maximizing the log-likelihood[15]

$$\sum_i \log \left[\sum_k f_k \prod_{t=2}^{15} p_{i,t}(a_{i,t}, b_{i,t}|k) + \right.$$
$$\left. + \sum_{k<K} f_{k\to k+1} \prod_{t=2}^{15} \Big(\frac{15-t}{13} p_{i,t}(a_{i,t}, b_{i,t}|k) + \frac{t-2}{13} p_{i,t}(a_{i,t}, b_{i,t}|k+1) \Big) \right],$$

where the joint density $p_{i,t}(a_{i,t}, b_{i,t}|k) = p_{i,t}(a_{i,t}|k) p_{i,t}(b_{i,t}|k)$. This way the structural model assumes that any individual i uses the same level for making both the a and b-guesses in period t.

We report the parameter estimates for various Gill-Prowse model specifications in Table F.5.

F4.5 Models estimated for the first period choices

First period choices in our experiment are of special interest. We estimate the applicable models separately on period 1 data and compare their fit. All learning models are excluded since they require conditioning on the previous period data. The results of this analysis is shown in Table F.6 and the parameter estimates in Table F.7.

[15] To avoid numerical instabilities in estimation we set $\nu \geq 0.5$.

Specifications	**Parameter Estimates**					
	f_0	f_1	f_2	f_3	f_4	f_E
0-1 Learning	0.11 (0.04)	0.56 (0.05)				
0-2 Learning	0.11 (0.02)	0.23 (0.05)	0.07 (0.05)			
0-3 Learning	0.11 (0.02)	0.23 (0.05)	0.06 (0.05)	0.00 (0.00)		
0-4 Learning	0.11 (0.02)	0.23 (0.06)	0.06 (0.05)	0.00 (0.00)	0.00 (0.00)	
0-1 E Learning	0.11 (0.04)	0.54 (0.05)				0.01 (0.01)
0-2 E Learning	0.11 (0.02)	0.22 (0.05)	0.07 (0.05)			0.02 (0.01)
0-3 E Learning	0.11 (0.02)	0.21 (0.05)	0.06 (0.05)	0.00 (0.00)		0.02 (0.01)
0-4 E Learning	0.11 (0.02)	0.21 (0.05)	0.06 (0.04)	0.00 (0.00)	0.00 (0.00)	0.02 (0.01)
	$f_{0\to1}$	$f_{1\to2}$	$f_{2\to3}$	$f_{3\to4}$		
0-1 Learning	0.33 (0.03)					
0-2 Learning	0.23 (0.06)	0.36 (0.04)				
0-3 Learning	0.22 (0.06)	0.36 (0.05)	0.02 (0.01)			
0-4 Learning	0.22 (0.07)	0.36 (0.05)	0.02 (0.02)	0.00 (0.00)		
0-1 E Learning	0.33 (0.04)					
0-2 E Learning	0.23 (0.06)	0.35 (0.04)				
0-3 E Learning	0.23 (0.07)	0.35 (0.05)	0.02 (0.01)			
0-4 E Learning	0.23 (0.06)	0.35 (0.04)	0.02 (0.02)	0.00 (0.00)		
	α	β	δ	ν		
0-1 Learning	1.97 (0.17)	−0.63 (0.45)	3.38 (0.33)	0.50 (0.00)		
0-2 Learning	1.85 (0.17)	−0.58 (0.44)	3.39 (0.37)	0.50 (0.00)		
0-3 Learning	1.83 (0.17)	−0.59 (0.44)	3.39 (0.37)	0.50 (0.00)		
0-4 Learning	1.84 (0.19)	−0.59 (0.34)	3.39 (0.37)	0.50 (0.00)		
0-1 E Learning	1.89 (0.19)	−0.87 (1.47)	3.20 (0.32)	0.50 (0.00)		
0-2 E Learning	1.76 (0.19)	−0.91 (1.03)	3.21 (0.34)	0.50 (0.00)		
0-3 E Learning	1.75 (0.19)	−0.91 (1.02)	3.20 (0.34)	0.50 (0.00)		
0-4 E Learning	1.75 (0.17)	−0.91 (0.81)	3.20 (0.36)	0.50 (0.00)		

Table F.5: Parameter estimates of the Gill-Prowse model for various specifications.

Models		Out-of-sample RMSEs				
		Sink	**SadNeg**	**SadPos**	**Source**	**Overall**
PONE		31.48	33.76	36.06	41.45	35.88
QRE		20.71	24.16	23.50	22.18	22.68
NI		19.59	23.65	25.23	**21.90**	22.69
Level-k	0	22.08	25.66	**22.61**	23.06	23.39
	1	20.37	31.36	37.08	37.73	32.39
	2	21.59	33.69	47.12	47.36	38.94
	3	25.08	34.88	48.30	48.13	40.30
	4	26.66	33.80	49.26	54.24	42.49
Mixed	0 1	19.97	22.40	24.51	22.11	22.31
	0 1 2	19.97	22.40	24.51	22.11	22.31
	0 1 2 3	19.97	22.40	24.51	22.11	22.31
	0 1 2 3 4	19.97	22.40	24.51	22.11	22.31
	1 2	20.49	24.23	40.54	29.87	29.76
	0 E	19.74	23.61	21.70	24.03	22.34
	1 E	23.87	30.24	27.52	32.63	28.75
	0 1 E	**19.22**	**22.28**	22.84	22.92	**21.87**
	0 1 2 E	19.22	22.33	22.89	22.92	21.89
	0 1 2 3 E	19.26	22.42	22.88	22.92	21.92
	0 1 2 3 4 E	19.25	22.41	22.99	22.92	21.95
	1 2 E	24.27	26.76	30.57	23.00	26.31
CH-Poisson	0 1	19.97	22.40	24.51	22.11	22.31
	0 1 2	20.02	22.48	24.57	22.03	22.33
	0 1 2 3	20.03	22.49	24.57	22.03	22.34
	0 1 2 3 4	20.03	22.49	24.57	22.02	22.34

Table F.6: Performance of the different models using individual data for period 1 only. This table includes only those models from Table 5 of the paper that do not require conditioning on the previous period data. The models are compared in terms of the out-of-sample RMSE using the leave-one-out procedure. The smallest RMSE for each treatment and overall is shown in boldface. Parameter estimates for models with parameters are reported in Table F.7.

Models		Parameter Estimates					
		λ					
QRE		0.03 (0.00)					
		λ	μ				
NI		0.06 (0.00)	0.05 (0.00)				
		f_0	f_1	f_2	f_3	f_4	f_E
Mixed	0 1	0.76 (0.04)	0.24 (0.04)				
	0 1 2	0.76 (0.04)	0.24 (0.04)	0.00 (0.00)			
	0 1 2 3	0.76 (0.04)	0.23 (0.05)	0.00 (0.00)	0.01 (0.02)		
	0 1 2 3 4	0.76 (0.03)	0.24 (0.04)	0.00 (0.00)	0.00 (0.01)	0.00 (0.01)	
	1 2		0.66 (0.03)	0.34 (0.03)			
	0 E	0.80 (0.02)					0.20 (0.02)
	1 E		0.59 (0.03)				0.41 (0.03)
	0 1 E	0.68 (0.04)	0.18 (0.04)				0.14 (0.02)
	0 1 2 E	0.67 (0.03)	0.18 (0.04)	0.01 (0.01)	0.15 (0.03)		
	0 1 2 3 E	0.68 (0.03)	0.15 (0.06)	0.00 (0.01)	0.02 (0.03)		0.14 (0.03)
	0 1 2 3 4 E	0.68 (0.03)	0.16 (0.06)	0.00 (0.00)	0.02 (0.03)	0.01 (0.01)	0.14 (0.03)
	1 2 E		0.30 (0.04)	0.32 (0.03)			0.39 (0.03)
		τ					
CH-Poisson	0 1	0.32 (0.08)					
	0 1 2	0.27 (0.05)					
	0 1 2 3	0.27 (0.04)					
	0 1 2 3 4	0.27 (0.04)					

Table F.7: Parameter estimates of the models for period 1 guesses. Standard errors are reported in parentheses. The parameters of the model with the smallest overall RMSE are shown in boldface.

References

Breitmoser, Y. (2012). Strategic reasoning in p-beauty contests. *Games and Economic Behavior 75*(2), 555–569.

Camerer, C. and H. T. Ho (1999). Experience-weighted attraction learning in normal form games. *Econometrica 67*(4), 827–874.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory 104*(1), 137–188.

Camerer, C. F., T.-H. Ho, and J. K. Chong (2004). Behavioural game theory: thinking, learning and teaching. In *Advances in understanding strategic behaviour*, pp. 120–180. Springer.

Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature 51*(1), 5–62.

Evans, G. W. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Gill, D. and V. Prowse (2016). Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy 124*(6), 1619–1676.

Goeree, J. K. and C. A. Holt (2004). A model of noisy introspection. *Games and Economic Behavior 46*(2), 365–382.

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

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

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