Consistent Depth of Reasoning in Level-k Models

David J. Cooper
Enrique Fatas
Antonio J. Morales
Shi Qi

Online Appendix

Contents:

A Experimental Instructions

A.1 General Instructions (role: participant 1)

The purpose of this experiment is to study how individuals make decisions in certain environments. The instructions are easy to understand and if you follow them, your will receive an amount of money at the end of the experiment, in a fully confidential way, meaning that no participant will get to know the payoffs to the other participants (note that this implies that no participant will know your payoff from this experiment). You may ask questions at any time by raising your hand. Any communication between participants is forbidden and subjected to immediate exclusion from the experiment.

In this experiment, you will be taking 20 independent decisions, grouped in five blocks of four decisions each. Although all participants will be facing the same 20 decisions, half of them will play the role of participant 1 and the other half will play the role of participant 2. Your role has been randomly assigned depending on the cubicle number you are in; in this experiment, you will be taking 20 independent decisions as participant 1.

All the decisions share the same logic. You will be required to choose a multiple of 10 between 110 and 200 (both included). The number you have chosen will be matched with the number chosen by a randomly chosen participant 2. Your payoffs will depend upon the two chosen numbers using a procedure that we will explain in due time. Notice that you will never know the identity of the participant with which you have been matched, neither the other way around.

At the beginning of each block, we will explain the procedure by which the chosen numbers determine your payoffs. We will use the same procedure for the four decisions within a block. Although all decisions within a block share the same logic, each decision has its own specificities that will be explained in detail. For a better understanding of the procedure by which payoffs are linked to chosen numbers for each decision, we will handle you a payoff table displaying all possible combinations of chosen numbers and the corresponding payoffs, and we will explain to you how the payoffs have been obtained from the chosen numbers using the block procedure.

It is important that you understand how the payoff tables are generated, because at the end of the experiment, one of the twenty decisions will be randomly chosen and the corresponding payoff table together with your chosen number and the number chosen by the participant with which you have been matched will be used to compute your payoff from this experiment. All the payoffs in the payoff table are expressed in ECU. The exchange rate used in this experiment is €1=10 ECU.

Before proceeding to the first block, we will show you a payoff table. Notice that all the payoffs displayed in this payoff table are randomly generated integers from the interval [0, 200]. This means that, contrary to what will happen in blocks 1 to 5, there is no inner logic behind the payoff table. Because there are 10 multiples of 10 between 110 and 200, this payoff table, as well as all the payoff tables that you will receive in this experiment, is composed of 10 rows and 10 columns. Also, and unlike in this payoff table, the decisions that you make along this experiment may imply negative payoffs. You can always avoid losses by choosing appropriate decisions but notice that any negative payoff will be compensated with the payoffs you will receive from answering two questionnaires at the end of the experiment.

		Participant 2									
		110	120	130	140	150	160	170	180	190	200
Participant 1	110	137	30	7	183	148	122	148	114	145	128
		56	43	112	107	132	56	169	163	194	71
	120	142	107	167	147	19	90	71	112	108	112
		14	176	18	77	12	67	77	34	9	112
	130	38	128	111	182	191	18	183	159	24	75
		166	64	46	157	9	200	111	32	43	34
	140	188	200	52	83	63	52	89	132	143	117
		108	156	130	50	69	163	97	102	45	145
	150	45	115	48	83	136	36	133	126	129	30
		101	73	46	159	113	27	5	191	116	200
	160	64	5	186	11	117	98	122	35	185	10
		2	45	170	8	171	25	125	50	29	80
	170	167	80	19	19	167	178	196	179	41	43
		120	187	130	49	169	110	175	28	195	64
	180	168	147	162	15	51	9	180	60	160	126
		21	9	193	37	76	56	72	23	184	154
	190	173	183	18	49	181	32	83	192	153	8
		75	6	139	56	121	136	187	135	24	75
	200	77	86	88	2	195	49	84	59	188	103
		47	176	31	1	199	107	70	174	94	42

The rows of the payoff table display the numbers that can be chosen by **Participant 1**, shaded in gray; and the columns display the numbers that can be chosen by **Participant 2**, shaded in white. A pair composed of a number chosen by a Participant 1 and a number chosen by a Participant 2 determines one cell of the payoff table. And each cell contains two numbers: the payoff to **Participant 1**, shaded in gray, and the payoff to **Participant 2**, shaded in white.

For a better understanding of how to interpret the numbers displayed in the payoff table, we will solve some examples. We will choose random numbers for a participant 1 and a participant 2 using these 10 numbered balls, from 1 to 10. The numbers will be used to compute the payoff to each participant using the payoff table projected on the white screen. Pay attention to the way in which the payoffs to the participants are associated to their chosen numbers. When we finish the examples, we will randomly choose one more number for a participant 1 and a participant 2 and you will have to answer correctly the following questionnaire for you to take part in the experiment.

Question 1:

Participant 1: Number ________ Participant 1 gets ________ ECUS

Participant 2: Number ________ Participant 2 gets ________ ECUS

Question 2:

Participant 1: Number ________ Participant 1 gets ________ ECUS

Participant 2: Number ________ Participant 2 gets ________ ECUS

A.2 Instructions for the Five Classes of Games

Here we include the instructions for the five classes of games, corresponding to those sessions with the order 1/2/3/4/5

A.2.1 Block 1: Imperfect Price Competition

Remember that all your decisions are numbers which are a multiple of 10 between 110 and 120. Participants 1 choose numbers shaded in gray and participants 2 choose numbers shaded in white. In each cell, the payoffs to participants 1 and 2 are shaded in gray and white, respectively.

In the four decisions within this block, the payoffs to the participants will depend upon the smaller chosen number by the two participants (between 110 and 200), and on a percentage that will be varying along the four decisions (either 20% or 80%). In the first two decisions, the percentage will be the same for the two participants, while in the remaining two decisions, percentages will be different to each other.

Let's see an example. The first payoff table that you have been handed in refers to the first decision. The participant with the smaller chosen number receives as many ECU as her chosen number, while the other participant (if her chosen number is larger) receives 20% of the smaller chosen number. For example, if participant 1 chooses 180 and participant 2 chooses 160, the smaller number is 160 and therefore, participant 2 gets 160 ECU while participant 1 gets 20% of 160, that is, 32 ECU. You can check in the payoff table that if participants 1 and 2 choose 180 and 160 respectively, the payoffs to participants 1 and 2 are 32 and 160, respectively. If both participants had chosen the same number (for example, 160), each of them would have obtained 50% of that number, plus half of the percentage. Because in the first decision both participants have the same percentage (20%), each player would get an additional 10%. If you check the payoff table, you will see that if both participants choose 160, their payoffs would be 96 (80+16).

We will now explain some more examples using the payoff table for decision 1, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

Note that the four decisions in this block differ again in the percentage that it is used to compute the payoff to participants. The only difference between decision 2 and decision 1 is that a participant receives as many ECU as her chosen number if her chosen number is the smaller number; 80% of the smaller number if her chosen number is not the smaller number; and half of her chosen number plus 40% of that number (rather than 10%), if both participants chose the same number.

For example, if participant 1 chooses 150 and participant 2 chooses 160, the smaller number is chosen by participant 1, who receives 150 ECU; while participant 2 gets 80% of 150 (120 ECU). If you check the payoff table, you will see that if participants 1 and 2 choose 150 and 160 respectively, their payoffs are 150 and 120, respectively.

If both participants had chosen the same number (for example, 150), each participant would have received 50% of the chosen number (75), plus half of the percentage they are endowed with. As in decision 2, both participants are endowed with percentage 80%, both participants would receive an additional 40%. The payoff table shows that if both participants choose 150, the payoff to them is 135 (75 + 60).

We will now explain more examples using the payoff table for decision 2, as it is currently projected on the white screens in the lab. Please, pay attention to the explanation. When we finish the explanation, we will ask you to choose a number for decision 2, although you can revise your choice before we collect the papers.

In the two last decisions, unlike the first two decisions, the percentages that the participants are endowed with are not equal to each other. In decision 3, participant 1 gets 20% of the smaller number if her chosen number is larger, while participant 2 gets 80%. In decision 4, these percentages are 80% for participant 1 and 20% for participant 2.

For example, if in decision 3 participant 1 has chosen 120 and participant 2 has chosen 140, the smaller number is the one chosen by participant 1, who receives 120 ECU, while participant 2 gets 80% of that amount (96 ECU). The payoff table shows that if participants 1 and 2 choose 120 and 140 respectively, their payoffs are 120 and 96, respectively.

If in decision 3, participant 1 had chosen the number 140 and participant 2 had chosen 120, then the smaller number would correspond now to participant 2, who would get 120 ECU (as much as participant 1 in the previous example). But now, participant 1 would only get 20% of that number (24 ECU). The payoff table confirms that if participants 1 and 2 choose 140 and 120, respectively, their payoffs would be 24 and 120, respectively.

If both players had chosen the same number (for example, 150), each participant would have obtained 50% of that number (75), plus half of their percentage. Because in decision 3, participant 1 is endowed with a percentage of 20%, he would get an additional 10% (15). Because in decision 3, participant 2 is endowed with a percentage of 80%, he would get an additional 40% (60). If you check the payoff table, you will find that if both participants choose 150, the payoffs are 90 (75 + 15) and 135 (75 + 60) for participant 1 and 2, respectively.

We will now be explaining more examples using the payoff table for decision 3, and then we will use the payoff table for decision 4. Please, pay attention to the explanation. When we finish the explanation, we will ask you to make your choices for decisions 3 and 4, although you me revise them before we collect the papers.

A.2.2 Block 2: Minimum Coordination Game

Remember that all your decisions are numbers which are a multiple of 10 between 110 and 120. Participants 1 choose numbers shaded in gray and participants 2 choose numbers shaded in white. In each cell, the payoffs to participants 1 and 2 are shaded in gray and white, respectively.

In the four decisions within this block, the payoffs to the participants depend again on the smaller number chosen by the two participants (between 110 and 200) and on a percentage that will be varying along the four decisions (again, 20% and 80%). Sometimes, this percentage will be the same for the two participants (In the first two decisions), while in others it will be different (in the last two decisions).

The logic of block 1 and block 2 differ in that now, the two participants receive the smaller number (of the two chosen), minus a percentage (20% or 80%) of the number they have chosen. Consider the following example. The first payoff table that you have been given refers to the first decision, where both participants receive the smaller of the two chosen number minus 20% of the number they have chosen (both participants are endowed with the same percentage). For example, if participant 1 chooses 180 and participant 2 chooses 160, the smaller number (160) is chosen by participant 2, and the payoffs to both participants are this number (160) minus 20% of the number they chose: 36 ECU for participant 1 (20% of 180) and 32 ECU for participant 2 (20% of 160). You can check in the payoff table that if participants 1 and 2 choose 180 and 160 respectively, their payoffs are 124 (160-36) and 128 (160-32), respectively.

If both participants had chosen the same number (for example, 160), each of them would have received this number (160), minus 20% of 160 (32). If you check the payoff table, you will see that if both participants choose 160, their payoffs are 128 (160-32).

We will now explain some more examples using the payoff table for decision 1, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers. Note that the four decisions in this block differ again in the percentage that it is used to compute the payoff to participants. The only difference between decision 2 and 1 is that a participant receives as many ECU as the smaller number, minus 80% of the number she has chosen. For example, if participant 1 chooses 150 and participant 2 chooses 120, the smaller number is the number chosen by participant 2 (120), and the payoff to participant 1 is 0 (120-120), while participant 2 gets 24 (120-96). If you check the payoff table, you will see that if participants 1 and 2 choose 150 and 120, respectively, their payoffs are 0 and 24, respectively.

If both players had chosen the same number (for example, 120), then each participant would have received this number (120), minus 80% of this number (96), that is, 24. If you check the payoff table, you will see that if both players choose 120, their payoffs is 24.

We will now explain some more examples using the payoff table for decision 2, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

For the next two decisions, unlike the first two ones, the percentages that are used to compute the participants' payoffs, are not equal for the two participants. In Decision 3, the percentage for participant 1 is 20% and the percentage for participant 2 is 80%. In Decision 4, the percentages are 80% for participant 1 and 20% for participant 2.

For example, if in Decision 3 participant 1 chooses 120 and participant 2 chooses 140, the smaller number is the one chosen by participant 1 (120), and the payoffs to participant 1 are 120 minus 20% of 120 (24), that is, 96. Participant 2 receives 120 minus 80% of her chosen number (80% of 140 is 112), that is, 8. The payoff table confirms that if participants 1 and 2 choose 120 and 140 respectively, their payoffs are 96 and 8 respectively.

If in Decision 3, participant 1 had chosen 140 and participant 2 had chosen 120, the smaller number would be the one chosen by participant 2 (120), and the payoff to participant 2 would be 120 minus 80% of 120 (96), that is, 24. Participant 1 would get 120 minus 20% of her chosen number (20% of 140 is 28), that is, 92. The payoff table shows that if participants 1 and 2 choose 140 and 120, respectively, their payoffs are 92 and 24, respectively.

If both players had chosen the same number (150), each of them would have received 150 minus the corresponding percentage of 150 (30 for participant 1 and 120 for participant 2); that is, participants 1 and 2 would have obtained 120 and 30, respectively. If you check the payoff table, you will find that if both participants choose 150, their payoff are 120 (150-30) and 30 (150-120).

We will now be explaining more examples using the payoff table for decision 3, and then we will use the payoff table for decision 4. Please, pay attention to the explanation. When we finish the explanation, we will ask you to make your choices for decisions 3 and 4, although you can revise them before we collect the papers.

A.2.3 Block 3: Travelers' Dilemma

Remember that all your decisions are numbers which are a multiple of 10 between 110 and 120. Participants 1 choose numbers shaded in gray and participants 2 choose numbers shaded in white. In each cell, the payoffs to participants 1 and 2 are shaded in gray and white, respectively.

In the four decisions within this block, the payoffs to the participants depend again on the smaller number chosen by the two participants (between 110 and 200) and on a quantity that will be varying along the four decisions (20 and 80). Sometimes, this quantity will be the same for the two participants (as in the first two decisions), while in others it will be different (as in the last two decisions).

The logic of this block resembles the logic in block 2, because the payoffs to the two participants also depends on the smaller number (of the two chosen). But in this block, the participant who has chosen the smaller number receives as many ECU as the smaller number plus a quantity, while the participant who has chosen the larger number receives as many ECU as the smaller number minus a quantity. If both participants choose the same number, then there is no additional quantity to be added/subtracted, and the payoffs to the participants are simply their chosen number.

Consider the following example. The first payoff table that you have been given refers to the first decision, where both participants receive the smaller of the two chosen number and the quantity to be added/subtracted is the same for the two participants, 20. For example, if participant 1 chooses 180 and participant 2 chooses 160, the smaller number (160) is chosen by participant 2, whose payoff is 180 (160+20) and the payoff to participant 1 is 140 (160-20). You can check in the payoff table that if participants 1 and 2 choose 180 and 160 respectively, their payoffs are 140 (160-20) and 180 (160+20), respectively.

If both participants had chosen the same number (for example, 160), each of them would have received this number (160). If you check the payoff table, you will find that if both participants choose 160, their payoffs are 160.

We will now explain some more examples using the payoff table for decision 1, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

Note that the four decisions in this block differ again in the quantities that are used to compute the payoffs to participants. The only difference between decision 2 and 1 is that a participant receives as many ECU as the smaller number, and 80 if added/subtracted depending on her chosen number being the smaller number. For example, if participant 1 chooses 150 and participant 2 chooses 120, the smaller number is the number chosen by participant 2 (120), and the payoff to participant 1 is 40 (120-80), while participant 2 gets 200 (120+80). If you check the payoff table, you will see that if participants 1 and 2 choose 150 and 120, respectively, their payoffs are 40 and 200, respectively.

If both players had chosen the same number (for example, 120), then each participant would have received this number (120). If you check the payoff table, you will see that if both players choose 120, their payoffs is 120.

We will now explain some more examples using the payoff table for decision 2, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

For the next two decisions, unlike the first two ones, the quantities that are added/subtracted to compute participants' payoffs, are not equal across participants. In Decision 3, the quantity for participant 1 is 20 and the quantity for participant 2 is 80. In Decision 4, these quantities are 80 for participant 1 and 20 for participant 2.

For example, if in Decision 3 participant 1 chooses 120 and participant 2 chooses 140, the smaller number is the one chosen by participant 1 (120), and the payoffs to participant 1 are 120 plus 20, that is, 140. Participant 2 receives 120 minus 80, that is, 40. The payoff table confirms that if participants 1 and 2 choose 120 and 140 respectively, their payoffs are 140 and 40 respectively.

If in Decision 3, participant 1 had chosen 140 and participant 2 had chosen 120, the smaller number would be the one chosen by participant 2 (120), and the payoff to participant 2 would be 120 plus 80, that is, 200. Participant 1 would get 120 minus 20, that is, 100. The payoff table shows that if participants 1 and 2 choose 140 and 120, respectively, their payoffs are 100 and 200, respectively.

If both players had chosen the same number (150), each of them would have received 150. If you check the payoff table, you will find that if both participants choose 150, their payoffs are 150.

We will now be explaining more examples using the payoff table for decision 3, and then we will use the payoff table for decision 4. Please, pay attention to the explanation. When we finish the explanation, we will ask you to make your choices for decisions 3 and 4, although you can revise them before we collect the papers.

A.2.4 Block 4: The "11-20" Game

Remember that all your decisions are numbers which are a multiple of 10 between 110 and 120. Participants 1 choose numbers shaded in gray and participants 2 choose numbers shaded in white. In each cell, the payoffs to participants 1 and 2 are shaded in gray and white, respectively.

In the four decisions within this block, the payoffs to the participants depend on their chosen number (between 110 and 200) and on a quantity (20 or 80). Sometimes, this quantity will be the same for the two participants (as in the first two decisions), while in others it will be different (as in the last two decisions).

Unlike block 3, in this block each participant will receive as many ECU as their chosen number plus an additional quantity (20 or 80), if her number is exactly ten units below the number chosen by the other participant.

Consider the following example. The first payoff table that you have been given refers to the first decision. For example, if participant 1 chooses 180 and participant 2 chooses 160, participant 1 gets 180 and participant 2 gets 160. There is no additional quantity to be added because no participant has undercut the other by 10 units. You can check in the payoff table that if participants choose 180 and 160, then their payoffs are 180 and 160 respectively.

If participant 2 had chosen 170, then participant 1 would have got 180 and participant 2 would have received 190 (170+20). If you check the payoff table, you will find that if participant 1 and 2 choose 180 and 170, respectively, then their payoffs are 180 and 190 respectively.

We will now explain some more examples using the payoff table for decision 1, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

Note that the four decisions within this block differ again in the quantities that are used to compute the payoffs to participants. The only difference between decision 2 and 1 is that a participant receives as many ECU as her chosen number plus 80 if this number is exactly 10 units below the number chosen by the other participant. For example, if participant 1 chooses 150 and participant 2 chooses 120, each participant gets her chosen number: the payoff to participant 1 is 150, while participant 2 gets 120. If you check the payoff table, you will see that if participants 1 and 2 choose 150 and 120, respectively, their payoffs are 150 and 120, respectively.

If participant 1 had chosen 110 and participant 2 had chosen 120, the payoff to participant 1 would be her number, 110, plus 80 because he had undercut participant 2 by exactly 10 units. If you check the payoff table, you will see that if participant 1 chooses 110 and participant 2 chooses 120, then the payoffs to participants 1 and 2 are 190 (110+80) and 120 respectively.

We will now explain some more examples using the payoff table for decision 2, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

For the next two decisions, unlike the first two ones, the quantities that are added if undercutting by 10 units is achieved, are not equal across participants. In Decision 3, the quantity for participant 1 is 20 and the quantity for participant 2 is 80. In Decision 4, these quantities are 80 for participant 1 and 20 for participant 2.

For example, if in Decision 3 participant 1 chooses 120 and participant 2 chooses 130, then participant 1 gets her chosen number, 120, plus 20; that is 140, and participant 2 gets her chosen number 130. The payoff table confirms that if participants 1 and 2 choose 120 and 130 respectively, their payoffs are 140 and 130 respectively.

If in Decision 3, participant 1 had chosen 130 and participant 2 had chosen 120, then participant 1 would get her chosen number, 120, and participant 2 would get 200 (120+80). The payoff table shows that if participants 1 and 2 choose 130 and 120, respectively, their payoffs are 130 and 200, respectively.

If both players had chosen the same number (150), each of them would have received 150. If you check the payoff table, you will find that if both participants choose 150, their payoffs are 150.

We will now be explaining more examples using the payoff table for decision 3, and then we will use the payoff table for decision 4. Please, pay attention to the explanation. When we finish the explanation, we will ask you to make your choices for decisions 3 and 4, although you can revise them before we collect the papers.

A.2.5 Block 5: The All-pay Auction

Remember that all your decisions are numbers which are a multiple of 10 between 110 and 120. Participants 1 choose numbers shaded in gray and participants 2 choose numbers shaded in white. In each cell, the payoffs to participants 1 and 2 are shaded in gray and white, respectively.

In the four decisions within this block, the payoffs to the participants depend on their chosen number (between 110 and 200) and on a quantity (20 or 80). Sometimes, this quantity will be the same for the two participants (as in the first two decisions), while in others it will be different (as in the last two decisions).

Unlike in the previous blocks, in this block each participant will always receive a fix amount of 110 ECU, minus her chosen number plus an additional quantity (20 or 80), if her number is larger than the number chosen by the other participant. If both participants choose the same number, then each participant gets half of her additional quantity.

Consider the following example that uses the payoff table from Decision 1. For example, if participant 1 chooses 170 and participant 2 chooses 160, the larger number is from participant 1 (170). Participant 1 receives the fixed amount (110), minus her chosen number (170), plus the additional quantity (20); her payoff is therefore -40 (110-170+20). Participant 2 receives the fixed amount (110) minus her chosen number (160), that is, -50. You can check in the payoff table that if participants choose 180 and 160, then their payoffs are -40 and -50 respectively.

If both participants had chosen the same number 160, then both participants would have received the fixed amount (110), minus the chosen number (160) plus half of the additional quantity (20). That is, the payoff to each participant would have been -40. If you check the payoff table, you will find that if both participants choose 160, their payoffs are -40 and -40 respectively.

We will now explain some more examples using the payoff table for decision 1, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

Note that the four decisions within this block differ again in the quantities that are used to compute the payoffs to participants. The only difference between decision 2 and 1 is that a participant receives the additional quantity 80 if her chosen number is larger than the number chosen by the other participant. For example, if participant 1 chooses 150 and participant 2 chooses 120, the larger number is the one chosen by participant 1 (150). Participant 1 receives the fixed amount (110), minus her chosen number (150) plus the additional quantity (80): 40. Participant 2 receives the fixed amount (110), minus her chosen number (120), that is, -10. If you check the payoff table, you will see that if participants 1 and 2 choose 150 and 120, respectively, their payoffs are 40 and -10, respectively.

If participant 2 would have chosen a larger number (160), then she would receive the fixed amount (110), minus her chosen number (160), plus the additional quantity (80); that is, 30. Participant 1 would receive -40 (110-150). If you check the payoff table, you will find that if participant 1 chooses 150 and participant 2 chooses 160, their payoffs are -40 and 30 respectively.

We will now explain some more examples using the payoff table for decision 2, which is projected on the white screens. Please, pay attention to the explanation. When we finish explaining the examples, we will ask you to make your choice, although you can revise your decision before we collect the papers.

For the next two decisions, unlike the first two ones, the additional quantities that are added are not equal across participants. In Decision 3, the quantity for participant 1 is 20 and the quantity for participant 2 is 80. In Decision 4, these quantities are 80 for participant 1 and 20 for participant 2.

For example, if in Decision 3 participant 1 chooses 120 and participant 2 chooses 130, the larger number is the one chosen by participant 2 (130). Then the payoff to participant 2 is the fixed amount (110), minus her chosen number (130) plus the additional quantity (80); that is, 60. Participant 1 receives the fixed amount (110) minus her chosen number (120), that is, -10. The payoff table confirms that if participants 1 and 2 choose 120 and 130 respectively, their payoffs are -10 and 60 respectively.

If in Decision 3, participant 1 had chosen 130 and participant 2 had chosen 120, then the larger number would be the one chosen by participant 1 (130). Then the payoff to participant 1 would be the fixed amount (110), plus the additional quantity (20), minus her chosen number (130); that is, 0. The payoff to participant 2 would be the difference between the fixed amount (110) and her chosen number (120), that is, -10. The payoff table shows that if participants 1 and 2 choose 130 and 120, respectively, their payoffs are 0 and -10, respectively.

We will now be explaining more examples using the payoff table for decision 3, and then we will use the payoff table for decision 4. Please, pay attention to the explanation. When we finish the explanation, we will ask you to make your choices for decisions 3 and 4, although you can revise them before we collect the papers.

B Frequency Table

We break down subjects' choices by class of game and by payoff parameters within class in the following table. The payoff parameters are given from the subject's point of view. For example, the rows labeled Class 1, LH refers to data from Class 1, Imperfect Price Competition, where the subject's own payoff parameter is low ($\alpha_i = 20$) and his rival's payoff parameter is high ($\alpha_j = 80$).

Table B: Summary of Choices

Game		Choice									
		110	120	130	140	150	160	170	180	190	200
Class 1, LL	Freq.	95	36	19	20	17	7	8	4	4	14
	Perc.	42.4%	16.1%	8.5%	8.9%	7.6%	3.1%	3.6%	1.8%	1.8%	6.3%
Class 1, LH	Freq.	79	39	16	26	22	9	10	9	8	6
	Perc.	35.3%	17.4%	7.1%	11.6%	9.8%	4.0%	4.5%	4.0%	3.6%	2.7%
Class 1, HL	Freq.	27	18	16	22	23	21	18	26	22	31
	Perc.	12.1%	8.0%	7.1%	9.8%	10.3%	9.4%	8.0%	11.6%	9.8%	13.8%
Class 1, HH	Freq.	17	12	18	13	26	18	21	25	37	37
	Perc.	7.6%	5.4%	8.0%	5.8%	11.6%	8.0%	9.4%	11.2%	16.5%	16.5%
Class 2, LL	Freq.	4	3	11	14	25	14	13	16	17	107
	Perc.	1.8%	1.3%	4.9%	6.3%	11.2%	6.3%	5.8%	7.1%	7.6%	47.8%
Class 2, LH	Freq.	21	10	30	11	24	19	10	14	7	78
	Perc.	9.4%	4.5%	13.4%	4.9%	10.7%	8.5%	4.5%	6.3%	3.1%	34.8%
Class 2, HL	Freq.	51	23	48	26	19	9	11	2	3	32
	Perc.	22.8%	10.3%	21.4%	11.6%	8.5%	4.0%	4.9%	0.9%	1.3%	14.3%
Class 2, HH	Freq.	85	26	34	19	17	7	3	3	0	30
	Perc.	38.0%	11.6%	15.2%	8.5%	7.6%	3.1%	1.3%	1.3%	0.0%	13.4%
Class 3, LL	Freq.	25	12	11	10	14	13	16	24	47	52
	Perc.	11.2%	5.4%	4.9%	4.5%	6.3%	5.8%	7.1%	10.7%	21.0%	23.2%
Class 3, LH	Freq.	36	21	12	13	20	16	21	23	32	30
	Perc.	16.1%	9.4%	5.4%	5.8%	8.9%	7.1%	9.4%	10.3%	14.3%	13.4%
Class 3, HL	Freq.	101	25	18	11	9	13	16	8	11	12
	Perc.	45.1%	11.2%	8.0%	4.9%	4.0%	5.8%	7.1%	3.6%	4.9%	5.4%
Class 3, HH	Freq.	113	25	21	11	16	11	3	6	5	13
	Perc.	50.5%	11.2%	9.4%	4.9%	7.1%	4.9%	1.3%	2.7%	2.2%	5.8%
Class 4, LL	Freq.	0	0	2	0	2	2	5	17	63	133
	Perc.	0.0%	0.0%	0.9%	0.0%	0.9%	0.9%	2.2%	7.6%	28.1%	59.4%
Class 4, LH	Freq.	1	5	6	3	2	3	7	21	45	131
	Perc.	0.5%	2.2%	2.7%	1.3%	0.9%	1.3%	3.1%	9.4%	20.1%	58.5%
Class 4, HL	Freq.	1	1	3	4	6	4	7	28	132	38
	Perc.	0.5%	0.5%	1.3%	1.8%	2.7%	1.8%	3.1%	12.5%	58.9%	17.0%
Class 4, HH	Freq.	2	1	0	1	3	5	19	57	100	36
	Perc.	0.9%	0.5%	0.0%	0.5%	1.3%	2.2%	8.5%	25.5%	44.6%	16.1%
Class 5, LL	Freq.	196	15	7	3	0	1	1	0	0	1
	Perc.	87.5%	6.7%	3.1%	1.3%	0.0%	0.5%	0.5%	0.0%	0.0%	0.5%
Class 5, LH	Freq.	192	17	10	1	1	1	0	1	0	1
	Perc.	85.7%	7.6%	4.5%	0.5%	0.5%	0.5%	0.0%	0.5%	0.0%	0.5%
Class 5, HL	Freq.	81	100	18	7	7	3	2	4	1	1
	Perc.	36.2%	44.6%	8.0%	3.1%	3.1%	1.3%	0.9%	1.8%	0.5%	0.5%
Class 5, HH	Freq.	90	67	19	11	11	3	8	5	6	4
	Perc.	40.2%	29.9%	8.5%	4.9%	4.9%	1.3%	3.6%	2.2%	2.7%	1.8%

C Technical Details of Econometric Models

In this section, we provide the technical details of the baseline level-k model described in the manuscript. We use superscript $C \in \{1, 2, 3, 4, 5\}$ to denote the classes of games. We use $x \in \{110, 120, ..., 200\}$ to denote a subject's action in a game. A generic subject's own payoff parameter is denoted by $\alpha_i \in \{20, 80\}$ and his rival's payoff parameter is denoted by $\alpha_j \in \{20, 80\}$. The payoff function of subject i in Class C, $\pi^C(x_i, x_j|\alpha_i)$, depends on her own choice x_i, her rival's choice x_j, and her own payoff parameter in a game α_i. The payoff function does not depend on the rival's payoff parameter α_j for any of the games we study.

The baseline model, as specified in the main text, allows for five different types of subjects: Level-0, Level-1, Level-2, Pure-mixing, and Semi-mixing. We denote the types by subscripts $\{0, 1, 2, M, S\}$, respectively. Level-0 types are non-strategic, and make choices based on an exogenous probability distribution p_0. For example, if they use a uniform distribution, then the probability distribution over actions is $p_0 = \{0.1, 0.1, ..., 0.1\}$.

Level-1 types believe all other subjects are Level-0 types, and take the probability distribution p_0 as given. The expected payoff of a Level-1 subject choosing own effort x_i in class C given her own payoff parameter α_i is $E\pi_1^C(x_i|\alpha_i, \alpha_j)$. Because a Level-1 subject expects all subjects are level-0, and she believes that the rival's action is not strategic and does not depend on α_j, we drop α_j from level-1 player's expected payoff function. Then,

$$E\pi_1^C(x_i|\alpha_i) = \sum_{x_j \in \{110,120,...,200\}} \pi^C(x_i, x_j|\alpha_i) \cdot p_0(x_j)$$

The probability of choosing an action $x_i \in \{110, 120, ..., 200\}$ is:

$$p_1^C(x_i|\alpha_i) = \frac{\exp(\lambda E\pi_1^C(x_i|\alpha_i))}{\sum_{k \in \{110,120,...,200\}} \exp(\lambda E\pi_1^C(k|\alpha_i))}$$

We incorporate a noise term into subjects' decision making. Except for level-0 types, all types use a logit rule. The parameter λ, giving the sensitivity of subjects to differences in expected payoffs, governs the amount of noise in subjects' decisions. If $\lambda = 0$, subjects' choices are uniformly distributed over the ten available options. As λ increases, choices become more sensitive to differences in expected payoffs. As $\lambda \to \infty$, the distribution of choices converges to deterministic expected payoff maximization. For the baseline model, the value of λ is assumed to be the same for all types.

Similarly, Level-2 types believe all other subjects are Level-1 types, and take their rival's probability distribution over actions $p_1^C(x_j|\alpha_j)$ as given. The expected payoff of a level-2 type subject choosing action x_i in class C given her own payoff parameter α_i and her rival's payoff parameter α_j is

$$E\pi_2^C(x_i|\alpha_i,\alpha_j) = \sum_{x_j \in \{110,120,\ldots,200\}} \pi^C(x_i,x_j|\alpha_i) \cdot p_1^C(x_j|\alpha_j).$$

The probability of choosing x_i is then:

$$p_2^C(x_i|\alpha_i,\alpha_j) = \frac{\exp(\lambda E\pi_2^C(x_i|\alpha_i,\alpha_j))}{\sum_{k\in\{110,120,\ldots,200\}} \exp(\lambda E\pi_2^C(k|\alpha_i,\alpha_j))}$$

Unlike a Level-1 type, a Level-2 type's choice probability depends on both players' payoff parameters.

Pure-mixing types (Type M) randomly draw a level in the beginning of each game. They use level-1 reasoning with probability θ_1, level-2 with probability θ_2, and level-0 with probability $1-\theta_1-\theta_2$. In a given game of class C, the choice probability of a Type M is $p_M^C(x_i|\alpha_i,\alpha_j) = \theta_1 p_1^C(x_i|\alpha_i) + \theta_2 p_2^C(x_i|\alpha_i,\alpha_j) + (1-\theta_1-\theta_2)p_0(x_i)$.

Semi-mixing types (Type S) randomly draw a level for each *class* of games, but use the same level of reasoning for all games within a class. The mixing probabilities $\{\theta_0,\theta_1,\theta_2\}$ are the same as those of the pure-mixing types. After the levels are drawn, a semi-mixing type's probability distribution over actions in a game from class C is either $p_1^C(x|\alpha_i)$, or $p_2^C(x|\alpha_i,\alpha_j)$, or $p_0(x)$, depending on the level drawn for the class of games.

A given subject i makes a sequence of 20 choices, four choices in each class of games. To simplify notation, we use h to denote high payoff parameter $\alpha = 80$, and l to denote low payoff parameter $\alpha = 20$. We use $\chi_i^C = \{x_{i,hh}^C, x_{i,hl}^C, x_{i,lh}^C, x_{i,ll}^C\}$ to denote any four-tuple of choices in Class C, for any $C \in \{1,2,3,4,5\}$. Similarly, we denote the choice probability of level-1 type with $\alpha_i = 80$ as $p_{1,h}^C(x)$ and the choice probability with $\alpha_i = 20$ as $p_{1,l}^C(x)$. The choice probability of level-2 type with own payoff parameter $\alpha_i = 80$ and rival payoff parameter $\alpha_j = 80$ can be written as $p^C_{2,hh}(x)$. Probabilities $p^C_{2,hl}(x)$, $p^C_{2,lh}(x)$ and $p^C_{2,ll}(x)$ are analogously defined. In addition, Pure-mixing types' probability distributions $p^C_{M,hl}(x)$, $p^C_{M,hl}(x)$, $p^C_{M,lh}(x)$ and $p^C_{M,ll}(x)$ are analogously defined. Extending the notation in a natural way, let $\tilde{p}^C_l(\chi^C)$ to denote the probability of any given sequence of four actions χ^C in class $C \in \{1,2,3,4,5\}$ for a type $l \in \{0,1,2,M,S\}$. Using the model specified above, we specify $\tilde{p}^C_l(\chi)$ for the five types of subjects:

- Level-0 Type – $\tilde{p}^C_0(\chi) = p_0(x_{hh})p_0(x_{hl})p_0(x_{lh})p_0(x_{ll})$;
- Level-1 Type – $\tilde{p}^C_1(\chi) = p^C_{1,h}(x_{hh})p^C_{1,h}(x_{hl})p^C_{1,l}(x_{lh})p^C_{1,l}(x_{ll})$;
- Level-2 Type – $\tilde{p}^C_2(\chi) = p^C_{2,hh}(x_{hh})p^C_{2,hl}(x_{hl})p^C_{2,lh}(x_{lh})p^C_{2,ll}(x_{ll})$;
- Pure-mixing Type M – $\tilde{p}^C_M(\chi) = p^C_{M,hh}(x_{hh})p^C_{M,hl}(x_{hl})p^C_{M,lh}(x_{lh})p^C_{M,ll}(x_{ll})$;
- Semi-mixing Type S – $\tilde{p}^C_S(\chi) = \theta_1\tilde{p}^C_1(\chi) + \theta_2\tilde{p}^C_2(\chi) + (1-\theta_1-\theta_2)\tilde{p}^C_0(\chi)$.

The probability of observing a subject who is type $l \in \{0,1,2,M,S\}$ choosing such a sequence $\xi = \{\chi^1,\chi^2,\chi^3,\chi^4,\chi^5\} = \{x^1_{hh}, x^1_{hl}, ..., x^5_{lh}, x^5_{ll}\}$ is $\prod_{C=1}^5 \tilde{p}^C_l(\chi^C)$

Let the probability weight of a subject being Type l to be w_l, and $w_0 = 1 - \sum_{l\in\{1,2,M,S\}} w_l$. Therefore the probability of observing a subject choosing this sequence ξ, is

$$\Omega(\xi) = \sum_{l=0,1,2,M,S} w_l \cdot \prod_{C=1}^{5} \tilde{p}^C_l(\chi^C)$$

For each subject, we observe a sequence of 20 choices, one for each game played. Let $\mathbb{X}$ denote the space of all possible sequences of 20 actions, and let ξ_n denote a particular sequence observed of subject n, so $\xi_n \in \mathbb{X}$. We are treating the sequence of actions $\xi_n = \{\chi^1_n,\chi^2_n,\chi^3_n,\chi^4_n,\chi^5_n\} = \{x^1_{n,hh}, x^1_{n,hl}, ..., x^5_{n,lh}, x^5_{n,ll}\}$ of each subject as one observation, so our formulation of the likelihood function already takes into consideration the possible correlation of actions by the same subject across different classes of games.

We construct the likelihood of observing each 20-tuple by first calculating the likelihood for each type, based on the choice probabilities described above, and then using w_0, w_1, w_2, w_M and w_S to calculate a weighted average of the likelihoods. The log likelihood function can be written as

$$\mathbb{L}(\Theta) = \sum_{n=1}^{224} \log(\Omega(\xi_n|\Theta))$$

where Θ denotes the set of parameters to be estimated.

C.1 Models in Table 6

Model 1 allows all five types of individuals described in the preceding section. This is the baseline model.

Model 2 is restricted to allow only consistent types, so $w_M = w_S = 0$, and $w_0 = 1 - w_1 - w_2$. This is the "Consistent Types only" Level-k model.

Model 3 is restricted to allow only the mixing types (Type M and Type S), so $w_0 = w_1 = w_2 = 0$ and $w_S = 1 - w_M$. This is the "Mixing Types only" model.

C.2 Models in Table 7

The first column repeats Model 1, the baseline model.

We estimated an alternative model allowing for two additional pure-mixing types, those who randomly draw a type between Level 1 and Level 0 (Type $M, 10$) and those who randomly draw a type between Level 2 and Level 0 (Type $M, 20$). The choice probability of these two additional mixing types are:

$$\begin{aligned} p^C_{M,10}(x_i|\alpha_i,\alpha_j) &= \frac{\theta_1}{1+\theta_1} p^C_1(x_i|\alpha_i) + \frac{1}{1+\theta_1} p_0(x_i) \\ p^C_{M,20}(x_i|\alpha_i,\alpha_j) &= \frac{\theta_2}{1+\theta_2} p^C_1(x_i|\alpha_i) + \frac{1}{1+\theta_2} p_0(x_i) \end{aligned}$$

Here, the probabilities θ_1 and θ_2 are the same as those for the other mixing types. The estimated weight on these two types is essentially zero.

Model 4 relaxes the assumption that level-0 type choice probability is uniformly distributed. Conditional on being a Level 0 type, there is a probability γ_{safe} of being the "safe" type, who would choose the maximin choice; there is a probability γ_{coop} of being the "cooperative" type, who would choose to maximize payoffs subject to both players making the same choices. According to our game designs, the cooperative choice is 200 in Classes 1, 2, 3, 4 and 110 in Class 5, and the safe choice is 110 in Classes 1, 2, 3 and 5 and 200 in Class 4. In addition to the consistent Level 1 types (with a probability weight ω_1) and the consistent Level 2 types (with a probability weight ω_2), there are three consistent Level-0 types in the model: (1) Level-0 types who draw their choices from a uniform distribution (with a probability weight $(1-\omega_1-\omega_2-\omega_M-\omega_S)\times(1-\gamma_{safe}-\gamma_{coop})$); (2) Level-0 types who always choose the safe choice (with a probability weight $(1-\omega_1-\omega_2-\omega_M-\omega_S)\times\gamma_{safe}$); (3) Level-0 types who always choose the cooperative choice (with a probability weight $(1-\omega_1-\omega_2-\omega_M-\omega_S)\times\gamma_{coop}$).The pure mixing types (with a probability weight ω_M) draw one of these five types at the beginning of each game and the semi-mixing types (with a probability weight ω_S) draw one of the five types at the beginning of each class of games. The mixing types have a probability θ_1 of being Level 1, a probability of θ_2 of being Level 2, a probability of $(1-\theta_1-\theta_2)\times\gamma_{safe}$ of being the "safe" Level-0 type, a probability of $(1-\theta_1-\theta_2)\times\gamma_{coop}$ of being the "cooperative" Level-0 type, and a probability of $(1-\theta_1-\theta_2)\times(1-\gamma_{safe}-\gamma_{coop})$ of being the uniform Level-0 type.

Model 5 adds an additional consistent type. Level 3 types believe all other subjects are level-2 types, and form expectation of payoffs as

$$E\pi_3^C(x_i|\alpha_i,\alpha_j)=\sum_{x_j\in\{110,120,\ldots,200\}}\pi^C(x_i,x_j|\alpha_i)\cdot p_2^C(x_j|\alpha_j,\alpha_i).$$

The probability of choosing x is then:

$$p_3^C(x_i|\alpha_i,\alpha_j)=\frac{\exp(\lambda E\pi_3^C(x_i|\alpha_i,\alpha_j))}{\sum_{k\in\{110,120,\ldots,200\}}\exp(\lambda E\pi_3^C(k|\alpha_i,\alpha_j))}$$

Both Type M and Type S subjects are mixing between 4 different types (Level-0, Level-1, Level-2, and Level-3). With probability θ_1, mixing types behave consistent with level-1; with probability θ_2, they behave consistent with level-2; with probability θ_3, they behave consistent with level-3; and with probability $1-\theta_1-\theta_2-\theta_3$, they behave consistently with level-0. The probability of being a consistent Level-3 type is w_3.

Model 6 is the Cognitive Hierarchy model, in which level-0 and level-1 remain the same as in our baseline model. However, level-2 types' reasoning has an increased level of strategic depth compared to our baseline model. Instead of believing all other players are level-1 types, level-2 thinking takes into accounts the existence of both level-1 types and level-0 types. They use Bayes' rule to update the probability of meeting a Level-1 type, where

$$\sigma_1 = \frac{w_1 + (w_M + w_S)\theta_1}{w_1 + w_0 + (w_M + w_S)(1 - \theta_2)}.$$

So the probability of meeting a level-0 type is $1 - \sigma_1$. Level-2 reasoning applies to Level-2 Type, as well as Type M and Type S when they are drawn to be level-2 in a particular game.

C.3 Models in Table 8

The first column repeats Model 1, the baseline model.

Model 7 uses expected payoff premiums to predict subject mixing probability. To construct expected payoff premiums, we first construct the expected payoffs for each level. We use the probability distribution generated endogenously by the model as the distribution of a subject's own actions conditioned on their levels. We use the empirical probability distribution observed in the data to get the distribution of actions used by a subject's rival. In other words, we calculate the expected payoff for a subject as if they behave exactly as our model predicts (conditioned on the level) and face a randomly drawn rival from the population.

We use the empirical choice probability $\tilde{p}^C(x_j|\alpha_j, \alpha_i)$ as the probability over rival's choice x_j. We use the model choice probabilities $p_0(x)$, $p_1^C(x|\alpha_i)$, and $p_2^C(x|\alpha_i, \alpha_j)$ as a player's own choice probability. For a Type M individual, the expected payoff in a given class of game C, with a pair of payoff parameters α_i and α_j, is

$$\begin{aligned}
E\pi_0^C(\alpha_i, \alpha_j) &= \sum_{x_i \in \{110,120,\ldots,200\}} p_0^C(x_i) \sum_{x_j \in \{110,120,\ldots,200\}} \pi^C(x_i, x_j|\alpha_i) \cdot \tilde{p}^C(x_j|\alpha_j, \alpha_i) \\
E\pi_1^C(\alpha_i, \alpha_j) &= \sum_{x_i \in \{110,120,\ldots,200\}} p_1^C(x_i|\alpha_i) \sum_{x_j \in \{110,120,\ldots,200\}} \pi^C(x_i, x_j|\alpha_i) \cdot \tilde{p}^C(x_j|\alpha_j, \alpha_i) \\
E\pi_2^C(\alpha_i, \alpha_j) &= \sum_{x_i \in \{110,120,\ldots,200\}} p_2^C(x_i|\alpha_i, \alpha_j) \sum_{x_j \in \{110,120,\ldots,200\}} \pi^C(x_i, x_j|\alpha_i) \cdot \tilde{p}^C(x_j|\alpha_j, \alpha_i)
\end{aligned}$$

We use the following mixing probabilities for Type M, where

$$\tilde{\theta}_1^C(\alpha_i, \alpha_j) = \bar{\theta}_1 + \mu_1 \cdot (E\pi_1^C(\alpha_i, \alpha_j) - E\pi_0^C(\alpha_i, \alpha_j));$$
$$\tilde{\theta}_2^C(\alpha_i, \alpha_j) = \bar{\theta}_2 + \mu_2 \cdot (E\pi_2^C(\alpha_i, \alpha_j) - E\pi_1^C(\alpha_i, \alpha_j)).$$

Then using a logit transformation, we restrict the mixing probabilities to be between 0 and 1:

$$\theta_1^C(\alpha_i, \alpha_j) = \frac{\exp(\tilde{\theta}_1^C(\alpha_i, \alpha_j))}{1 + \exp(\tilde{\theta}_1^C(\alpha_i, \alpha_j)) + \exp(\tilde{\theta}_2^C(\alpha_i, \alpha_j))};$$
$$\theta_2^C(\alpha_i, \alpha_j) = \frac{\exp(\tilde{\theta}_2^C(\alpha_i, \alpha_j))}{1 + \exp(\tilde{\theta}_1^C(\alpha_i, \alpha_j)) + \exp(\tilde{\theta}_2^C(\alpha_i, \alpha_j))}.$$

For Type S subjects, we calculate the average expected payoff of a level-l ($l \in \{0, 1, 2\}$) subject in a game of class C as

$$E\Pi_l^C = (E\pi_l^C(80, 80) + E\pi_l^C(80, 20) + E\pi_l^C(20, 80)E\pi_l^C(20, 20))/4$$

we use the same $\bar{\theta}_1$, $\bar{\theta}_2$, μ_1, and μ_2, so the mixing probabilities of a Type S player are

$$\tilde{\theta}_1^C = \bar{\theta}_1 + \mu_1 \cdot (E\Pi_1^C - E\Pi_0^C);$$
$$\tilde{\theta}_2^C = \bar{\theta}_2 + \mu_2 \cdot (E\Pi_2^C - E\Pi_1^C).$$

Then using a logit transformation, we restrict the mixing probabilities to be between 0 and 1:

$$\theta_1^C = \frac{\exp(\tilde{\theta}_1^C)}{1 + \exp(\tilde{\theta}_1^C) + \exp(\tilde{\theta}_2^C)};$$
$$\theta_2^C = \frac{\exp(\tilde{\theta}_2^C)}{1 + \exp(\tilde{\theta}_1^C) + \exp(\tilde{\theta}_2^C)}.$$

Models 8 and 9 explore the relationship between cognitive ability, as indicated by Raven scores (RPM), and the consistency of a subject's depth of reasoning. Model 8 allows the mixing probabilities that subject n uses *if he is a mixing type (M or S)* to depend on her Raven score R_n, where

$$\tilde{\theta}_{1,n} = \bar{\theta}_1 + \mu_1 \cdot R_n;$$
$$\tilde{\theta}_{2,n} = \bar{\theta}_2 + \mu_2 \cdot R_n.$$

Then using a logit transformation, we restrict the mixing probabilities to be between 0 and 1:

$$\theta_{1,n} = \frac{\exp(\tilde{\theta}_{1,n})}{1 + \exp(\tilde{\theta}_{1,n}) + \exp(\tilde{\theta}_{2,n})};$$
$$\theta_{2,n} = \frac{\exp(\tilde{\theta}_{2,n})}{1 + \exp(\tilde{\theta}_{1,n}) + \exp(\tilde{\theta}_{2,n})}.$$

Model 9 allows the probability of types to depend on subject n's raven score R_n. Specifically, it affects the combined weight on both mixing types: $w_{mix} = w_M + w_S$. In particular,

$$\tilde{w}_{n,mix} = \phi + \varphi \cdot R_n.$$

Then using a logit transformation, so the weight w_{mix} is between 0 and 1:

$$w_{n,mix} = \frac{\exp(\tilde{w}_{n,mix})}{1 + \exp(\tilde{w}_{n,mix})}.$$

By extension, the total weight on being a consistent type is $1 - w_{mix}$. *Conditional on being a consistent type*, let the probability of being a Level-1 type be δ_1, the probability of being a Level-2 type be δ_2, and the probability of being a Level-0 type be $1 - \delta_1 - \delta_2$. Also conditional on being a mixing type, the probability of being a Type M is δ_M. Then the unconditional probabilities of subject i can be constructed as follows:

$$w_{1,i} = (1 - w_{i,mix})\delta_1; \qquad w_{2,i} = (1 - w_{i,mix})\delta_2$$
$$w_{M,i} = w_{i,mix}\delta_M; \qquad w_{S,i} = w_{i,mix}(1 - \delta_M)$$
$$w_{0,i} = (1 - w_{i,mix})(1 - \delta_1 - \delta_2)$$

C.4 Models in Table 10

The first two columns repeats Model 1 and Model 2.

Model 10 has four types: Level 1, Level 2, Level 0 and Type M, and does not have Type S. Model 10 also does not allow Type M to mix level 0. In other words, the choice probability of

Type M is:

$$p_M^C(x_i|\alpha_i,\alpha_j) = \theta_1 p_1^C(x_i|\alpha_i) + (1-\theta_1)p_2^C(x_i|\alpha_i,\alpha_j)$$

Model 11 models an Ambiguity Aversion type (AA type). The AA types choose between being a Level 1 Type and a Level 2 Type. An AA type knows his rival is either Type 0 or Type 1, and assigns a probability $\tau_1 \in [0,1]$ to the other's type being Level 1, and a probability $\tau_0 = 1-\tau_1$ to the other's type being Level 0.

The AA types understand that four scenarios are possible:

1. She is Level 1, and her rival is Level 0, then her payoff is $\tilde{\pi}_{1,0} = \sum_{x_i} E\pi_1(x_i)\cdot p_1(x_i)$;
2. She is Level 1, but her rival is Level 1, then her payoff is $\tilde{\pi}_{1,1} = \sum_{x_i} E\pi_2(x_i)\cdot p_1(x_i)$;
3. She is Level 2, but her rival is Level 0, then her payoff is $\tilde{\pi}_{2,0} = \sum_{x_i} E\pi_1(x_i)\cdot p_2(x_i)$;
4. She is Level 2, and her rival is Level 1, then her payoff is $\tilde{\pi}_{2,1} = \sum_{x_i} E\pi_2(x_i)\cdot p_2(x_i)$;

In scenarios 2 and 3, she made a mistake. Let $\underline{\pi} = \min\{\tilde{\pi}_{1,0}, \tilde{\pi}_{1,1}, \tilde{\pi}_{2,0}, \tilde{\pi}_{2,1}\}$. We define a subject's value $v_{k,l} = \tilde{\pi}_{k,l} - \underline{\pi}$, where $k=1,2$ is her own type, and $l=0,1$ is her rival's type.

An AA type assign a mixing probability $\hat{\theta}_k$ to her own type k. An AA type solve the following optimization problem to find the optimal mixing probability $\hat{\theta}_1$, where $\hat{\theta}_2 = 1-\hat{\theta}_1$:

$$\max_{\hat{\theta}_1\in[0,1]} \left(\sum_{l\in\{0,1\}} \tau_l \left(\sum_{k\in\{1,2\}} \hat{\theta}_k v_{k,l}\right)^{\eta}\right),$$

where $\eta \in (0,1)$ is the ambiguity aversion parameter.

The choice probability of an AA Type is $p_{AA}(x_i) = \hat{\theta}_1(\eta)p_1(x_i) + (1-\hat{\theta}_1(\eta))p_2(x_i)$. We can use this to construct the likelihood function. This resembles the pure-mixing type in our original specification. Compare to the Pure-Mixing types, we use an ambiguity aversion parameter (η) instead of a couple of mixing parameters (θ_1 and θ_2). The model structure is less flexible, so the fit is slightly worse.

Model 12 makes a minor change to Model 11. In choosing a level, Model 11 assumes that the AA types anticipate the noises both in their own decisions and in other's decisions. The modified

Model 12 assumes that the AA types, when choosing a level, believe that they will maximize perfectly while still anticipating noise in others' choices.

To be more precise, the AA types believe that the choice probability $\hat{p}_k(x_j) = 100\%$ for the x_j that maximizes level k ($k = 1, 2$) player's payoff, and $\hat{p}_k(x_i) = 0\%$ for all other x_i. This probability is different from the actual choice probability with noise, $p_k(x_j)$, in the baseline model and Model 11. The rest of the model specification is identical to that of Model 11.

C.5 Alternative Cognitive Hierarchy and Higher Depths of Reasoning

In the main manuscript, we have considered Model 5 - a model incorporating level-3 types, and Model 6 - a variant of the cognitive hierarchy (CH) model of Camerer, Ho, and Chong (2004). In our version of the CH model, level-2 types take into account that both level-1 types and level-0 types exist, and use Bayes rule to generate beliefs about the likelihood of being matched with a level-1 type.

Table C.1 shows results from several extensions of the CH model and a model of higher depths of reasoning. In Table C.1, for comparison purposes, Columns 1, 2 and 4 repeat the Models 1, 6, and 5, respectively. Model C1 allows the CH types' beliefs of $\sigma_{2,1}$ to be fitted directly from data instead of imposing rational expectations.

Model C2 adds Level-3 CH types, who takes into accounts the existence of level-0, level-1 and level-2 types. The Level-3 CH types' beliefs of $\sigma_{3,1}$ and $\sigma_{3,2}$ are fitted directly from data. Allowing for level-3 types and/or fitting the beliefs over lower levels rather than imposing rational expectations improve the models ability to fit the data but have no impact on our main qualitative conclusions.

Model C3 builds on Model 5 and adds Level 4 types, who believe all other subjects are level-3 types. The expected payoffs and choice probabilities are similarly defined in Model 5. Both Type M and Type S subjects are mixing between 5 different types (Level-0, Level-1, Level-2, Level-3, and Level-4). With probability θ_1, mixing types behave consistent with level-1; with probability θ_2, they behave consistent with level-2; with probability θ_3, they behave consistent with level-3; with probability θ_4, they behave consistent with level-4; and with probability $1 - \theta_1 - \theta_2 - \theta_3 - \theta_4$, they behave consistently with level-0. The probability of being a consistent Level-4 type is w_4. The estimation results of Model C3, presented in Table C.1, are almost identical with those of

Model 5. Adding types with even higher depths of reasoning does not meaningfully change model estimation.

Table C.1. Alternative Cognitive Hierarchy Models

	Model 1 Baseline	Model 6 Cognitive Hierarchy	Model C1 Cog. Hier. Fitted	Model 5 Model with Level-3	Model C2 Cog. Hier. w/ Level-3	Model C3 Model with Level-4
w_1	0.097***	0.079**	0.071**	0.053*	0.034	0.053*
	(0.036)	(0.035)	(0.035)	(0.031)	(0.027)	(0.031)
w_2	0.000	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-	-
w_M	0.431***	0.283***	0.223***	0.319***	0.242***	0.319***
	(0.065)	(0.062)	(0.053)	(0.062)	(0.051)	(0.062)
w_S	0.462***	0.632***	0.704***	0.619***	0.717***	0.619***
	(0.065)	(0.025)	(0.061)	(0.067)	(0.057)	(0.067)
θ_1	0.560***	0.470***	0.360***	0.503***	0.371***	0.503***
	(0.023)	(0.068)	(0.053)	(0.024)	(0.041)	(0.024)
θ_2	0.153***	0.217***	0.325***	0.091***	0.197***	0.091***
	(0.017)	(0.034)	(0.049)	(0.017)	(0.037)	(0.017)
λ	0.175***	0.165***	0.163***	0.175***	0.183***	0.175***
	(0.010)	(0.008)	(0.007)	(0.008)	(0.009)	(0.008)
$\sigma_{2,1}$		0.635	0.345***		0.351***	
		Bayesian	(0.064)		(0.077)	
w_3				0.000	0.000	0.000
				-	-	-
θ_3				0.119***	0.122***	0.060
				(0.015)	(0.015)	(0.558)
$\sigma_{3,1}$					0.086	
					(0.061)	
$\sigma_{3,2}$					0.781***	
					(0.083)	
w_4						0.000
						-
θ_4						0.059
						(0.557)
L.L.	$-8,201.187$	$-8,192.194$	$-8,185.691$	$-8,135.189$	$-8,119.272$	$-8,135.189$
AIC	16,416.375	16,398.388	16,387.384	16,288.378	16,262.543	16,292.378
BIC	16,440.256	16,422.270	16,414.678	16,319.083	16,303.483	16,329.906

Notes: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively. Log Likelihood is abbreviated as L.L.

C.6 Models of Alternative Mixing Specifications

Table C.2 includes three additional variant models of alternative specifications of mixing types. The first column repeats Model 1, the baseline model.

TABLE C.2. ALTERNATIVE MIXING SPECIFICATIONS

	MODEL 1 BASELINE	MODEL C4 RESTRICTED NO S TYPES	MODEL C5 DIFF. MIX. PROB. FOR S TYPES	MODEL C6 DIFF. MIX. PROB. FOR CLASSES 4-5
w_1	0.097***	0.115***	0.075***	0.083***
	(0.036)	(0.031)	(0.029)	(0.031)
w_2	0.000	0.000	0.000	0.000
	-	-	-	-
w_M	0.431***	0.855***	0.476***	0.624***
	(0.065)	(0.033)	(0.051)	(0.056)
w_S	0.462***	0	0.449***	0.263***
	(0.065)	FIXED	(0.023)	(0.019)
θ_1	0.560***	0.558***	0.602***	0.394***
	(0.023)	(0.019)	(0.049)	(0.055)
θ_2	0.153***	0.164***	0.232***	0.125***
	(0.017)	(0.014)	(0.031)	(0.027)
λ	0.175***	0.190***	0.180***	0.179***
	(0.010)	(0.010)	(0.010)	(0.008)
$\theta_{S,1}$			0.497***	
			(0.035)	
$\theta_{S,2}$			0.041***	
			(0.017)	
θ'_1				0.792***
				(0.031)
θ'_2				0.144***
				(0.028)
LOG LIKELIHOOD	$-8,201.187$	$-8,283.478$	$-8,167.299$	$-8,072.132$
AIC	16,416.375	16,578.955	16,352.599	16,162.263
BIC	16,440.256	16,599.425	16,383.303	16,192.968

NOTES: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% respectively.

Model C4 in the second column allows the consistent types and type M, not type S: $w_S = 0$, and $w_0 = 1 - w_1 - w_2 - w_M$.

Model C5 allows Type M and Type S to have different mixing probabilities. In particular, the probability of Type M being level-k $k \in \{0, 1, 2\}$ is θ_k, and the probability of Type S being level-k $k \in \{0, 1, 2\}$ is $\theta_{S,k}$. Unlike the baseline model, we do not make any ex-ante restrictions on $\theta_{S,k} = \theta_k$.

Model C6 relaxes the baseline model's assumption that the mixing probabilities are the same across different classes of games. In particular, Model C5 allows different mixing probabilities for Classes 4 and 5 (θ'_k for Classes 4 and 5). We do not make any ex-ante restrictions on how θ'_k compares to θ_k for Classes 1, 2 and 3.

In summary, we find that inconsistent types are predominant in all of the alternative models in Table C.2. This supports the idea that inconsistent types is a robust feature of our empirical setting.

C.7 Model of Learning and Fatigue

Changes in depth of reasoning across classes of games could reflect feedback-free learning rather than inconsistency. On the other hand, subjects might become fatigued by the end of the experiment, leading to less sophisticated behavior. Both these possibilities might affect our results. To address these two concerns, we estimated a version of the baseline model that allowed for both learning and fatigue. In Table C.3, the first column repeats Model 1, the baseline model. The second column presents the results of Model C7 - the learning and fatigue model.

In the alternative setup of Model C7, we consider a "Learning" type and a "Fatigue" type. The mixture weight of the Learning type is ω_{Learn}, and the mixture weight of the Fatigue type is $\omega_{Fatigue}$. Both types start by drawing a depth of reasoning (Level-0, Level-1, or Level-2) for the first class of games played. Note that this is not necessarily Class 1, as different sessions saw Classes 1 - 3 in different orders. The order of classes are reported in Table 2 of the manuscript.

In the first class they play, both types have a probability θ_1 of being Level 1, a probability θ_2 of being Level 2, and a probability $(1-\theta_1-\theta_2)$ of being Level 0. For subsequent classes, the Learning types who are not in Level 2 can learn and have a higher depth of reasoning with a probability ρ_{Learn}. The learning types can only become more sophisticated in reasoning. Once a learning type becomes Level 1, she can only be Level 1 or higher in the subsequent classes. A learning type stays Level 2 in all subsequent classes once she becomes Level 2. The opposite happens for the Fatigue types. At the beginning of each new class of games, Fatigue types who are not already at Level 0 move down a level (Level-1 to Level-0 or Level-2 to Level-1) with a probability $\rho_{Fatigue}$. A fatigue type stays Level 0 in all subsequent classes once she becomes Level 0.

Table C.3. Learning and Fatigue Model

	Model 1 Baseline	Model C7 Learning & Fatigue
w_1	0.097***	0.067*
	(0.036)	(0.035)
w_2	0.000	0.000
	-	-
w_M	0.431***	0.468***
	(0.065)	(0.067)
w_S	0.462***	0.357***
	(0.065)	(0.078)
θ_1	0.560***	0.559***
	(0.023)	(0.022)
θ_2	0.153***	0.159***
	(0.017)	(0.017)
λ	0.175***	0.180***
	(0.010)	(0.011)
w_{Learn}		0.108**
		(0.053)
$w_{Fatigue}$		0.000
		-
ρ_{Learn}		0.170***
		(0.056)
$\rho_{Fatigue}$		0.459
		(24.155)
Log Likelihood	$-8,201.187$	$-8,199.186$
AIC	$16,416.375$	$16,420.372$
BIC	$16,440.256$	$16,457.900$

Notes: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Results of Model C7 show that adding learning and fatigue types does little to improve the model fit - note that the BIC goes up rather than down. The model detects a small fraction of learning types, but does not detect the presence of any fatigue types. The overall impact on our conclusions is slight. Some of the inconsistency that the baseline model identifies as being due to semi-mixed types may reflect learning without feedback. Consistent with the original manuscript, fatigue does not appear to play an important role in our results.

C.8 Rational Expectations and Nash Equilibrium Types

We fit two variations of the baseline model considering possible Nash Equilibrium behaviors. In the first, Model C8, we add an additional type that responds to the empirical distribution (Rational Expectations). Beliefs are set equal to the observed distribution over actions in each of the actions, and actions are then a noisy best response to beliefs. The second variant, Model C9, adds a type who believes that other agents are playing according to the Nash equilibrium. For games with a unique equilibrium, full weight is put on this equilibrium. Many of the games have multiple equilibria in these cases, beliefs split the weight equally between the various equilibria. For some variants of the 11 - 20 game, the equilibrium is in mixed strategies rather than pure strategies. Beliefs are consistent with the mixed strategy equilibrium in these cases.

Table C.4. Rational Expectation and Nash Equilibrium Types

	Model 1 Baseline	Model C8 Rational Expectations	Model C9 Nash Equil. Types
w_1	0.097***	0.098***	0.097***
	(0.036)	(0.037)	(0.036)
w_2	0.000	0.000	0.000
	-	-	-
w_M	0.431***	0.427***	0.431***
	(0.065)	(0.066)	(0.065)
w_S	0.462***	0.464***	0.462***
	(0.065)	(0.065)	(0.065)
θ_1	0.560***	0.559***	0.560***
	(0.023)	(0.023)	(0.023)
θ_2	0.153***	0.153***	0.153***
	(0.017)	(0.018)	(0.017)
λ	0.175***	0.174***	0.175***
	(0.010)	(0.010)	(0.010)
w_{RE}		0.002	
		(0.007)	
w_{NE}			0.000
			-
Log Likelihood	$-8,201.187$	$-8,201.092$	$-8,201.187$
AIC	$16,416.375$	$16,418.184$	$16,418.375$
BIC	$16,440.256$	$16,445.478$	$16,445.668$

Notes: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively.

In both of these models, we use $\tilde{p}_g^C$ to denote the empirical effort distribution or the belief weights on the Nash Equilibrium in Class C and Game $g \in \{hl, hh, lh, ll\}$. We denote the additional type - the NE types, whose expected payoff is

$$E\pi_{g,NE}^C(x_i) = \sum_{x_j \in \{110,120,...,200\}} \pi_g^C(x_i, x_j) \cdot \tilde{p}_g^C(x_j).$$

The probability of choosing an action $x_i \in \{110, 120, ..., 200\}$ is

$$p_{g,NE}^C(x_i) = \frac{\exp(\lambda E\pi_{g,NE}^C(x_i))}{\sum_{k \in \{110,120,...,200\}} \exp(\lambda E\pi_{g,NE}^C(k))}.$$

The probability of the NE type choosing any given sequence of actions $\chi = \{x_{hh}, x_{hl}, x_{lh}, x_{ll}\}$ in a particular class C is

$$\tilde{p}_{NE}^C(\chi) = p_{hh,NE}^C(x_{hh}) p_{hl,NE}^C(x_{hl}) p_{lh,NE}^C(x_{lh}) p_{ll,NE}^C(x_{ll}).$$

The probability of observing this type of subject choosing a sequence of 20 actions is $\prod_{C=1}^5 \tilde{p}_{NE}^C(\chi^C)$. The mixture weight of the NE types is ω_{NE}.

The estimation results are in Table C.4. The first column repeats Model 1, the baseline model. The second column presents the rational expectations model - Model C8. The weight on the rational expectations type is not statistically significant and the AIC and BIC go up rather than down, indicating that the fit is worse once we account for the increase in the number of parameters. The third column of the table above shows the results of the model with a Nash equilibrium type - Model C9. The result is a corner solution with no weight on the new type. Obviously, the fit is not improved by adding this type.

It is not surprising that the Nash equilibrium type receives no weight given the types of games used in our design. All five classes were taken from previous papers where part of the motivation was that Nash equilibrium did a poor job of explaining behavior. For example, in justifying use of the 11 20 game (our Class 4), Arad and Rubinstein (2012) point to the fact that level-k reasoning is quite natural while the Nash equilibrium is not.[1] Likewise, Goeree and Holt (2005) explicitly

[1] Using level-k reasoning is very natural: The games payoffs are described explicitly by the best-response function, a characteristic that triggers iterative reasoning. It is hard to think of plausible alternative decision rules for this

mention Nash equilibriums lack of predictive power for the minimum coordination game (our Class 2).[2] By design, these are not games where Nash equilibrium is expected to have much explanatory power.

C.9 Models of Alternative Noise Distributions

In the baseline model, the mixture model uses the logit noise distribution which remains the same for all subjects throughout the games. As a robustness test, we fit two additional models of alternative noise specifications. The first model, Model C10, uses the three-parameter t-distribution (Gill and Prowse, 2016), with mean μ, scale σ, and degrees of freedom ν. We set the mean to be the effort choice that maximizes the expected payoffs of subjects and estimate the scale and degrees of freedom parameters.

Same as in the Baseline model, Level-0 types in Model C10 are non-strategic and make choices based on an exogenous probability distribution p_0. Level-1 types believe all other subjects are Level-0 types, and take the probability distribution p_0 as given. Similarly, Level-2 types believe all other subjects are Level-1 types, and take their rival's probability distribution over actions $p_1^C(x_j|\alpha_j)$ as given. The expected payoff of a Level $k = 1, 2$ subject choosing action x_i in class C given her own payoff parameter α_i and her rival's payoff parameter α_j is $E\pi_k^C(x_i|\alpha_i, \alpha_j)$, similarly defined as in the baseline model.

The probability of subject i in level $k = 1, 2$ choosing an action $x_i \in \{110, 120, ..., 200\}$ is:

$$p_k^C(x_i|\alpha_i) = \frac{f(x_i; \mu_k(\alpha_i, \alpha_j)), \sigma, \nu)}{\sum_{z\in\{110,120,...,200\}} f(z; \mu_k(\alpha_i, \alpha_j)), \sigma, \nu)},$$

where $f(\cdot; \mu, \sigma, \nu)$ is the density function of the three-parameter t-distribution with mean μ, scale σ, and degrees of freedom ν. We set the mean $\mu_k(\alpha_i, \alpha_j)$ of the t-distribution to be the effort choice that maximizes the expected payoff of subject i of level $k = 1, 2$.

The second model, Model C11, allows for switching between two logit error parameters. At the beginning of every game, a subject draws the logit error parameters λ_1 with probability $\zeta \in [0, 1]$,

game. In particular, there is no pure-strategy Nash equilibrium [the game has a non-obvious symmetric mixed strategy equilibrium], and the game lacks dominated strategies.

[2]Some theorists argue that coordination game experiments are useless for game theory because the Nash equilibrium and its refinements have no predictive power in this case.

and draws λ_2 with probability $1-\zeta$. For example, let $p^C_{k,hh}(x_{hh}|\lambda_g)$ be the choice probability of a Level k individual with $(\alpha_i, \alpha_j) = (80, 80)$ in the game hh and draws a logit parameter λ_g $(g = 1, 2)$. Then, the choice probability is $p^C_{2,hh}(x_{hh}) = \zeta \cdot (p^C_{k,hh}(x_{hh}|\lambda_1) + (1-\zeta)(p^C_{k,hh}(x_{hh}|\lambda_2)$. We similarly define $p^C_{k,hl}(x_{hl})$, $p^C_{k,lh}(x_{lh})$, and $p^C_{k,ll}(x_{ll})$. The rest of the model is identical to the baseline model.

TABLE C.5. MODELS OF ALTERNATIVE NOISE DISTRIBUTIONS

	MODEL 1 BASELINE	MODEL C10 3-PARAMETER T-DISTRIBUTION	MODEL C11 TWO LOGIT PARAMETERS
w_1	0.097***	0.094***	0.148***
	(0.036)	(0.034)	(0.048)
w_2	0.000	0.022**	0.000
	-	(0.010)	-
w_M	0.431***	0.120**	0.300***
	(0.065)	(0.068)	(0.066)
w_S	0.462***	0.691***	0.548***
	(0.065)	(0.043)	(0.065)
θ_1	0.560***	0.484***	0.542***
	(0.023)	(0.024)	(0.028)
θ_2	0.153***	0.196***	0.199***
	(0.017)	(0.019)	(0.021)
λ	0.175***		
	(0.010)		
σ		10.529***	
		(1.714)	
ν		1.563**	
		(0.793)	
λ_1			0.082***
			(0.007)
λ_2			0.351***
			(0.029)
ζ			0.479***
			(0.048)
LOG LIKELIHOOD	$-8,201.187$	$-8,593.631$	$-8,163.985$
AIC	$16,416.375$	$17,203.262$	$16,345.970$
BIC	$16,440.256$	$17,230.556$	$16,376.675$

NOTES: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively.

The estimation results are in Table C.5. The first column repeats Model 1, the baseline model.

The second column presents the model with t-distribution - Model C10. The model with the 3-parameter t-distribution puts significant weight on both mixing types, but relatively more on the semi-mixing type than the baseline model. However, Model C10 fits quite a bit worse than the baseline model. Intuitively, the expected payoff maximizing action is the model outcome for both error structures, but the logistic distribution is more sensitive to the effect of changes in the payoff structure on play of other actions. Given the poor fit, we put little weight on this in evaluating our conclusions.

The third column presents the model with two logit error parameters - Model C11. Allowing the error parameter to vary between games leads to a slightly better fit (AIC and BIC are both lower) and somewhat changes the weights on types: compared to the baseline model, there is more weight on Level-1 and semi-mixing types, less on pure mixing types). None of these changes are large and none affect our qualitative conclusions.

C.10 Models of Alternative Level-0 Specifications

In the baseline model, Level-0 types are non-strategic and randomly choose effort choices according to a uniform distribution. One concern is that Level-0 strategies or beliefs can vary across games. In the manuscript, we consider an alternative specification of Level-0 types in Model 4 from Table 7. Model 4 has three Level-0 Types: (1) Level-0 types who draw their action from a uniform distribution; (2) Level-0 types who always choose the safe choice; (3) Level-0 types who always choose the cooperative choice.[3]

In the appendix, we fit two additional models. The first model, Model C12, is almost identical to Model 4 described above, except for the following detail. Rather than having three types of Level-0 players, there is only one type. This type uses a distribution that is a weighted average of the safe choice, and the cooperative choice, and the uniform distribution. The weights are γ_{safe}, γ_{coop}, and $1 - \gamma_{safe} - \gamma_{coop}$, respective. The mixture weight parameters are estimated from data.

[3] According to our game designs, the cooperative choice is 200 in Classes 1, 2, 3, 4 and 110 in Class 5, and the safe choice is 110 in Classes 1, 2, 3 and 5 and 200 in Class 4.

Specifically, the Level-0 type distribution of effort choice for each class of games is:

$$\begin{aligned} p_0^C &= \gamma_{safe} \cdot [1, 0, ..., 0] + \gamma_{coop} \cdot [0, .., 0, 1] + (1 - \gamma_{safe} - \gamma_{coop}) \cdot [0.1, 0.1, ..., 0.1] \ \ \forall C = 1, 2, 3; \\ p_0^4 &= (\gamma_{safe} + \gamma_{coop}) \cdot [0..., 0, 1] + (1 - \gamma_{safe} - \gamma_{coop}) \cdot [0.1, 0.1, ..., 0.1]; \\ p_0^5 &= (\gamma_{safe} + \gamma_{coop}) \cdot [1, 0..., 0] + (1 - \gamma_{safe} - \gamma_{coop}) \cdot [0.1, 0.1, ..., 0.1]. \end{aligned}$$

The second model, Model C13, allows for an extremely flexible specification of Level-0 types. For each of the five classes of games, the Level-0 effort choices are drawn from a discretized beta distribution. The distribution yields the uniform distribution if parameters $\alpha = \beta = 1$. For other parameter values of α and β, it is possible to get either peaked distributions or u-shaped distributions. Separate values of α and β are estimated for each class of games.

The estimation results are in Table C.6. The first column repeats Model 1, the baseline model. The second column repeats Model 4 in Table 7 of the manuscript – the model with three pure Level-0 types. The third column reports Model C12 with Level-0 types being the weighted average of of the safe choice, and the cooperative choice, and the uniform distribution.

The fourth column reports Model 13 with Level-0 types following beta-distributions.The auxiliary table below, Table C.6A, lists the estimates for the beta distribution shape parameters for each of the five classes.

Comparing Model C12 with the closely related model in the paper (Model 4, Table 7), the fit is significantly improved by what seems like a minor change. The model puts 96.1% of the weight on the two mixing types, versus 91% in Model 4. However, the weight shifts from the pure mixing type to the semi-mixing type. The difference does not have much to do with the choices of pure Level-0 types, as neither model puts substantial weight on this type (the respective weights on pure Level-0 types are .016 and .033). Likewise, the difference does not come from pure mixing types, who make the same decisions in both models when drawing Level-0, holding parameter θ_1, θ_2, and λ fixed. Therefore, the difference must come from the semi-mixing types. Under Model 4 in the paper, a semi-mixed type draws one of the three Level-0 types (uniform, safe, or cooperative) at the beginning of a class of games, and must play according to this type for all four games in the class. Model C12 does not impose this restriction. In other words, compared to Model 4 in the paper, Model C12 captures additional hard-to-explain inconsistencies in behavior within classes of games by allowing semi-mixing types to switch between the three Level-0 types. The approach is a purely mechanical way of picking up the switches and lacks any theoretical justification.

Table C.6. Models of Alternative Level-0 Specifications

	Model 1 Baseline	Model 4 Non-uniform Level-0	Model C12 Non-uniform Level-0 - Wgt. Avg.	Model C13 Beta Dist. Level-0
w_1	0.097***	0.075*	0.000	0.000
	(0.036)	(0.041)	-	-
w_2	0.000	0.000	0.006	0.005
	-	-	(0.013)	(0.010)
w_M	0.431***	0.447***	0.170***	0.163***
	(0.065)	(0.075)	(0.034)	(0.059)
w_S	0.462***	0.463***	0.791***	0.793***
	(0.065)	(0.068)	(0.036)	(0.061)
θ_1	0.560***	0.595***	0.475***	0.572***
	(0.023)	(0.021)	(0.023)	(0.028)
θ_2	0.153***	0.117***	0.200***	0.141***
	(0.017)	(0.015)	(0.021)	(0.026)
λ	0.175***	0.183***	0.072***	0.080***
	(0.010)	(0.015)	(0.003)	(0.003)
γ_{safe}		0.105***	0.429***	
		(0.0011)	(0.019)	
γ_{coop}		0.000	0.288***	
		-	(0.011)	
Log Likelihood	$-8,201.187$	$-8,097.555$	$-7,805.972$	$-7,657.595$
AIC	16,416.375	16,213.109	15,629.945	15,349.190
BIC	16,440.256	16,243.814	15,660.650	15,407.188

Notes: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table C.6A. Beta Distribution Shape Parameters

	Class 1	Class 2	Class 3	Class 4	Class 5
α	0.251***	0.021*	0.179***	1.444***	0.039***
	(0.033)	(0.012)	(0.030)	(0.086)	(0.011)
β	0.388***	0.023*	0.178***	0.452***	0.665***
	(0.031)	(0.014)	(0.026)	(0.031)	(0.064)

Notes: Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively.

By fitting a beta-distribution for every class of games, Model C13 similarly improves the fit compared to Model 4 in the paper. The logic is the similar to that behind Model C12. The estimated beta-distribution roughly approximates a U-shaped distribution of choices for the Level-0 types in Classes 1 - 3 by putting large amounts of weight on the safe and cooperative choices. In fact, Model C13 allows even more flexibility than Model C12, putting weight on choices other than the end points and varying the relative weight on the two end points across different classes of games. For example, Model C13 puts 75% of the weight for Level-0 types on 200, while Model C13 reduces this to 43% by shifting weight to choices 170 190. This added flexibility allows Model C13 to fit the data considerably better than Model 12. The underlying logic is unchanged; Model C13 improves over the baseline model by better capturing switches within classes of games. Ultimately, all of the models put substantial weight on the pure mixing and semi-mixing types. The more flexible models are improving on the fit by better capturing switches within classes of games, but do so in an entirely mechanical fashion.

References

Arad, A., and Rubinstein, A. "The 11–20 Money Request Game: A Level-k Reasoning Study." *American Economic Review* 102, no. 7 (2012): 3561–3573.

Goeree, J. K., and Holt, C. A. "An Experimental Study of Costly Coordination." *Games and Economic Behavior* 51, no. 2 (2005): 349–364.