Rationalist Experiments on War

Supplementary Materials

APPENDIX 1

Experimental Approach

Fearon's (1995) rationalist explanations for war are mechanisms of bilateral conflict that can be formally generalized and experimentally tested. I model the mechanisms and test them experimentally. The experiments isolate the strategic mechanisms in a controlled setting, testing if actual human decision-making converge to or diverge from the equilibrium predictions in the models.

While experiments may have high internal validity, their external validity is an important concern. One limitation is that that we cannot and should not scale up the stakes in the experiments to involve the life and death of real human beings, as decisions for wars do. However, it is important to note that formal models of war do not rely on "stake effects." The models are driven by specific payoff structures, which *are* replicated in the experiments. It is also important to consider if small stakes tend to under- or over-estimate the treatment effect.[1] Take the present study as an example: Being "reneged upon" in a commitment-problem situation leads to the loss of dollars in the context of the experiment, whereas the equivalent loss may be measured in terms of life and death in international politics. If we believe people are less trusting and more sensitive to the commitment problem (the treatment variable) in the latter case, then the treatment effect isolated in the lab may in fact be an under-estimation of the true treatment effect in actual international crises.

[1] I am grateful to Dustin Tingley for a conversation on this point.

The second limitation is that the subject pool is not a representative sample of national leaders. It is very difficult to obtain such a sample. It should be emphasized, however, that the behavioral predictions from Fearon's (1995) models are not specific to national leaders, but general to all decision-makers. While we cannot fully eliminate the possibility that certain findings may be affected by the subject samples, the data do provide real behavioral information on decision-making given our variables of interest.

Experiments replicate the ceteris-paribus condition often assumed in scientific theories but rarely achievable with observational data. Hence, carefully designed experiments provide special advantages for testing theoretical mechanisms in political science. Yet it is also important to conduct field verifications of the experimental results, so that we can better assess the substantive significance of the relevant mechanisms. A field verification exercise is conducted here with a historical case study, which is summarized in the paper and detailed in Appendix 5.

APPENDIX 2

Crisis Bargaining Game

I analyze the incentive structures implemented in the experiment to extract the theoretical predictions condition by condition. The game tree is diagrammed in Figure 1 in the paper.

Public Information with/without Enforcement

Suppose that there is no private information and all payoffs are known to the players. In particular, both players know that $c_A = c_B$.

In the enforcement condition, A cannot change its Stage-1 demand in Stage 2, thus $x_1 = x_2$. In the subgame perfect equilibrium A will demand $x_1 = x_2 = v - w_{1B} + c_B$ given B's reservation level ($w_{1B} - c_B$) at Stage 1. Based on the values in the experiment, B will accept the demand $x_1 = x_2$ since $v - x \geq w_{1B} - c_B$.[2] Hence:

> **Prediction 1:** War will be avoided if there is enforcement in the public-information condition.

In the no-enforcement condition, A can change its Stage-1 demand in Stage 2. In Stage 1, B is choosing between x_1 and war. Since there is no enforcement, A's x_1 is not credible and B will only consider x_2 in Stage 1. A will demand $x_2 = v - w_{2B} + c_B$ in Stage 2, based

[2] That B will not fight if the utility of war equals the utility of agreement is assumed.

on the amount that makes B just willing to accept. By backward induction, B will choose war in Stage 1 in the subgame perfect equilibrium since $w_{1B} - c_B > v - x_2$ in Stage 2. Thus:

> **Prediction 2:** War will occur with certainty if there is no enforcement in the public-information condition.

Private Information with/without Enforcement

Suppose there is private information on the costs of war c_A and c_B, and that c_A and c_B are drawn from a discrete uniform distribution on integers [0, v].

In the enforcement condition, $x_1 = x_2$ and B will choose war if $w_{1B} - c_B > v - x_1$. Hence, in any equilibrium the probability of war will be $\Pr(w_{1B} - c_B > v - x_1)$. Specifically, B will choose war if $5 - c_B > 10 - x_1$, and in any equilibrium the probability of war will be $\Pr(c_B < x_1 - 5)$. As such, A's expected utility for demanding x_1 is $u_A(x_1) = \Pr(c_B < x_1 - 5)(5 - c_A) + (1 - \Pr(c_B < x_1 - 5))(x_1)$. The experiment restricted c_A and c_B to be integers drawn from a discrete uniform distribution on [0, 4]. By direct proof, computing the maximized $u_A(x_1)$ in all combinations of c_A and c_B shows that the risk of war is always positive. The result and intuition are similar to that in Fearon's (1995, 410-11) take-it-or-leave-it game. A faces a trade-off between the size of its demand x_1 and the risk of war generated by the unknown c_B, with a higher demand leading to a better potential bargain but also a higher risk of war. Thus:

Prediction 3: There is always a positive risk of war in the private-information condition with enforcement.

In the no-enforcement condition, the equilibrium outcome will converge to the subgame perfect equilibrium outcome analyzed in the public-information case based on the values defined in the experiment. This is because B, who only considers x_2 in Stage 1, will choose war since the payoff for war in Stage 1 ($5 - c_B = 3$) will never be less than the payoff for agreement ($10 - x_2$) in Stage 2, given that the value of x_2 demanded by A will be based on B's reservation level ($3 - c_B \leq 3$), with $c_B \geq 0$.[3] Thus:

Prediction 4: War will occur with certainty in the private-information condition without enforcement.

[3] The case of war being avoided due to the possibility of $c_B = 0$ is trivial.

APPENDIX 3

Experimental Instructions

The experiment is programmed and conducted with z-Tree (Fischbacher 2007). Subjects viewed the instructions on their computer screens. The instructions are reproduced fully as follows.

Welcome:

Welcome to the experiment!

The experiment will take one hour. If you follow the instructions and make good decisions, you might earn a considerable amount of money. Hence it is important that you read the instructions very carefully. All the money you earn is yours to keep, and will be paid to you, in cash, at the end of the experiment. Your confidentiality is assured.

Please do not communicate with other players during the experiment. If you have questions or need assistance, raise your hand and a monitor will come to you.

You should NOT look at the decisions of others, or talk or laugh or exclaim aloud in the experiment. You will be warned if you violate the rule the first time. If you violate the rule a second time, you will be asked to leave and you will not be paid.

Overview:

The experiment is divided into 5 SCENARIOS with a total of 17 ROUNDS. Each round is independent. The experiment begins with Scenario 1 and ends with Scenario 5.

In the experiment, you will be randomly divided into groups of 2 players. Your opponent will change at random after every round. You will NOT be matched with the same opponent twice in the same scenario.

Earnings:

Your dollar earnings for the experiment are determined as follows.

The computer will randomly choose 9 rounds out of the 17 rounds. Then it will sum up your total point earnings in all 9 chosen rounds. Each round has a total possible value of 10 points.

We will pay you

$0.50 PER POINT

That is half a dollar, or fifty cents, for EVERY SINGLE POINT you earn. For example, 20 points will give you $10.00 in cash.

The more points you win, the more money you receive.

Transition:

Let us begin the experiment.

Role:

In the experiment, you are a national leader facing an international crisis.

In this crisis, your country is bargaining with another country for a valuable Prize.

As a national leader, you will want to get the BEST DEAL for your country.

Pairing:

You will be randomly assigned as either Country A or Country B in each round.

Your opponent will change at random after every round. You will NOT be matched with the same opponent twice in the same scenario.

Transition:

Scenario 1 will start in a few seconds.

SCENARIOS 1 – 2

Payoffs:

You are a national leader bargaining with another country over a valuable Prize.

Both countries know that:

- The Prize is worth **10 points**
- If there is a war, both countries will pay the **costs of war**

[For the public-information group]

The cost of war is PUBLIC information:

- **You know your Opponent's cost of war**
- **Your Opponent knows your cost of war**

[For the private-information group]

The cost of war is PRIVATE information:

- **You do NOT know your Opponent's cost of war**
- **Your Opponent does NOT know your cost of war**

Costs:

To generate the costs of war, the computer will assign one of the values {0, 1, 2, 3, 4} to you.

Then, it will assign one of the values {0, 1, 2, 3, 4} to your Opponent.

Once generated, the cost of war will be FIXED for Scenarios 1 and 2.

Costs:

The computer has generated the costs of war.

The cost is FIXED for all rounds in Scenarios 1 and 2.

[For the public-information group]

Your Cost of War = **2 points**
Opponent's Cost of War = **2 points**

[For the private-information group]

Your Cost of War = **2 points**
Opponent's Cost of War = **Unknown**

War Costs:

War is always costly.

Whenever war is triggered, **you lose 2 points.**

[For the public-information group]

Whenever war is triggered, **your opponent loses 2 points.**

[For the private-information group]

Whenever war is triggered, **your opponent loses [unknown].**

Power Shift:

The game has 2 stages:

STAGE 1

In Stage 1, Country A and Country B are **EQUALLY POWERFUL.**

If they fight a war, each country will be able to seize 50% of the prize for itself. Hence, Country A gets 5 points and Country B gets 5 points.

But because war is costly, both countries will also LOSE POINTS based on its own cost of war.

STAGE 2

In Stage 2, Country A becomes **MORE POWERFUL** than Country B.

If they fight a war, Country A will be able to seize 70% of the prize for itself. Hence, Country A gets 7 points and Country B gets 3 points.

But because war is costly, both countries will also LOSE POINTS based on its own cost of war.

Stages:

The game has 2 stages. Here is how the game works:

STAGE 1:

Country A suggests how much of the prize it wants for itself.

- The Prize is worth 10 points. Country A can demand 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 points of the Prize.
- Country A's suggestion in Stage 1 is not implemented until it is confirmed in Stage 2.

Country B agrees or disagrees

- If Country B agrees, you PROCEED TO STAGE 2.
- If Country B disagrees, war is triggered and THE GAME ENDS HERE.

STAGE 2:

Country A confirms the number of points it wants.

If you are playing in the **ENFORCEMENT CONDITION:**

- The computer WILL NOT ALLOW Country A to make any changes in Stage 2.
- Hence, the amount which Country A suggested in Stage 1 WILL BE THE SAME amount that is confirmed in Stage 2.

If you are playing in the **NO-ENFORCEMENT CONDITION:**

- The computer WILL ALLOW Country A to make changes in Stage 2.
- Hence, the amount which Country A suggested in Stage 1 MAY OR MAY NOT BE THE SAME amount that is confirmed in Stage 2.

Country B agrees or disagrees

PAYOFFS:

If Country B agrees in both Stages 1 and 2,

- Country A will get the number of points it demanded.
- Country B will get what is left of the Prize (i.e. 10 points minus the points Country A has taken).
- Both countries DO NOT LOSE POINTS in the costs of war.

*If Country B disagrees (war) in **Stage 1,***

- Country A gets **5 points minus its cost of war.** Country B gets **5 points minus its cost of war.**

*If Country B disagrees (war) in **Stage 2,***

- Country A gets **7 points minus its cost of war.** Country B gets **3 points minus its cost of war.**

Note:

Note that:

- Country B can wage war in Stage 1 or Stage 2.
- Both countries can avoid the costs of war if they can reach an agreement in Stage 2.

Questions:

To ensure that you have read the instructions carefully, here are a few questions:

(1) If war is triggered in Stage 1, each country gets 5 points minus its cost of war. Given your cost of war, how many points do you get if war is triggered in Stage 1? Please type your answer here: ___

(2) If war is triggered in Stage 2, Country A gets 7 points minus its cost of war and Country B gets 3 points minus its cost of war. If you are COUNTRY B, how many points do you get if war is triggered in Stage 2? ___

(3) You will be randomly assigned as either Country A or Country B in each round. Guess the % chance (probability) that you will be assigned as Country A. ___

(4) One of the values {0, 1, 2, 3, 4} is selected as your Opponent's cost of war. Guess the % chance (probability) that your Opponent will get the value 1. The value 2? The value 3? ___

Flag:

You are currently in Scenario [1/2].

Prelude:

You are in Scenario [1/2]. Here is the scenario summary:

STAGE 1:

Country A suggests how much of the prize it wants for itself.

Country B agrees (go to Stage 2) **or disagrees** (war is triggered).

STAGE 2:

Country A confirms how much of the prize it wants for itself.

Country B agrees (bargain is made) **or disagrees** (war is triggered).

PAYOFFS:

If bargain is made,

- Country A gets what it confirmed in Stage 2.
- Country B gets what is left of the Prize.
- Both countries avoid the costs of war.

*If war is triggered in **Stage 1,***

- Country A gets **5 points minus its cost of war.**

- Country B gets **5 points minus its cost of war.**

*If war is triggered in **Stage 2,***

- Country A gets **7 points minus its cost of war.**
- Country B gets **3 points minus its cost of war.**

Group Assignment:

There is **[ENFORCEMENT / NO ENFORCEMENT]** in this scenario.

- The computer [WILL NOT ALLOW / WILL ALLOW] Country A to make changes in Stage 2.

- Hence, the amount which Country A suggested in Stage 1 [will be the SAME / MAY OR MAY NOT BE THE SAME] amount that is confirmed in Stage 2.

Country Assignment:

In this round, you are randomly assigned as a leader in Country [A/B].

Treatment Summary:

There is [ENFORCEMENT / NO ENFORCEMENT].

- Country A [**cannot / can**] **make changes** in Stage 2.
- The amount Country A suggested in Stage 1 [will be the SAME / MAY OR MAY NOT BE THE SAME] amount confirmed in Stage 2.

The cost of war is [PRIVATE / PUBLIC] information.

[Private-information group]

- Your Cost of War = **2 points** and Opponent's Cost of War = **Unknown.**
- You do NOT know your Opponent's cost of war
- Your Opponent does NOT know your cost of war

[Public-information group]

- Your Cost of War = **2 points** and Opponent's Cost of War = **2 points.**
- You know your Opponent's cost of war
- Your Opponent knows your cost of war

Actual Play:

[STAGE 1]

[Country A (decision)]

You are Country A.

Suggest how much of the prize you want by choosing ONE of the possibilities below.

A gets 10, B gets 0
A gets 9, B gets 1
A gets 8, B gets 2
A gets 7, B gets 3
A gets 6, B gets 4
A gets 5, B gets 5
A gets 4, B gets 6
A gets 3, B gets 7
A gets 2, B gets 8
A gets 1, B gets 9
A gets 0, B gets 10

You [CAN / CANNOT] CHANGE the amount in Stage 2.

Your Cost of War = 2 points. Opponent's Cost of War = [2 points / Unknown].

THIS STAGE – Both countries are equally powerful.
If war is triggered, you get **5 points minus your cost of war.**
Country B gets **5 points minus its cost of war.**

NEXT STAGE – You are more powerful.
If war is triggered, you get **7 points minus your cost of war.**
Country B gets **3 points minus its cost of war.**

[Country A (waiting)]

Please wait while Country B makes a decision.

[Country B (waiting)]

Please wait while Country A makes a decision.

[For Country B (decision)]

Country A has suggested the following:

"A gets [], B gets []"

If you agree, you proceed to Stage 2 in which Country A will confirm its demand.

Country A [CAN / CANNOT] CHANGE the amount in Stage 2.

If you disagree, war is triggered.

Your Cost of War = 2 points. Opponent's Cost of War = [2 points / Unknown].

THIS STAGE – Both countries are equally powerful.
If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

NEXT STAGE – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< AGREE – Go to Stage 2 >
< DISAGREE – Go to War >

[If Country B disagrees]

Country B disagreed. War is triggered.

[STAGE 2]

[Country A (decision)] [Enforcement group]

You are Country A.

In Stage 1, you suggested the following:

"A gets [], B gets []"

You cannot change this decision. Click < CONFIRM > to continue.

< CONFIRM >

[Country A (decision)] [No-enforcement group]

You are Country A.

In Stage 1, you suggested the following:

"A gets [], B gets []"

You can change or confirm the suggested amount by choosing ONE of the possibilities below:

A gets 10, B gets 0
A gets 9, B gets 1
A gets 8, B gets 2
A gets 7, B gets 3
A gets 6, B gets 4
A gets 5, B gets 5
A gets 4, B gets 6
A gets 3, B gets 7
A gets 2, B gets 8
A gets 1, B gets 9
A gets 0, B gets 10

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

THIS STAGE – You are more powerful.
If war is triggered, you get **7 points minus your cost of war.**
Country B gets **3 points minus its cost of war.**

[For Country B (decision)]

Country A has confirmed the following:

"A gets [], B gets []"

If you agree, the bargain is made.

If you disagree, war is triggered.

Your Cost of War = 2 points. Opponent's Cost of War = [2 points / Unknown].
\

THIS STAGE – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< AGREE – Bargain is Made >
< DISAGREE – Go to War >

Outcome:

[If Country B agrees]

Country B agreed.

Country A gets the points it demanded.

Country B gets what is left of the Prize.

[If Country B disagrees]

Country B disagreed. War is triggered.

Country A gets 7 points minus its cost of war.

Country B gets 3 points minus its cost of war.

Transition [After Round 5]:

Scenario 1 is over. Scenario 2 will start in a few seconds.

Transition [Start of Round 6]:

Scenario 2 is similar to Scenario 1 except for a change in the ENFORCEMENT CONDITION.

Repeat Screens from "Prelude" to "Outcome" in Rounds 2-10.

SCENARIO 3

Players are randomly divided into the enforcement and no-enforcement groups.

Transition:

Scenario 2 is over. Scenario 3 will start in a few seconds.

Comparison:

One major difference between Scenario 3 and the earlier scenarios is:

COUNTRY B CAN DECIDE TO WAGE WAR AT ANY TIME

Payoffs:

You are a national leader bargaining with another country over a valuable Prize.

Both countries know that:

- The Prize is worth 10 points

- If there is a war, both countries will pay the costs of war

- Cost of war is **2 points for all countries.** This cost is fixed throughout the remaining rounds. All countries know that the cost of war is 2 points for all countries.

War Costs:

War is always costly.

Whenever war is triggered, **you lose 2 points.**

Whenever war is triggered, **your opponent loses 2 points.**

Group Assignment:

There is **[ENFORCEMENT / NO ENFORCEMENT]** in this scenario.

- The computer [WILL NOT ALLOW / WILL ALLOW] Country A to make any changes in Stage 2.

- Hence, the amount which Country A suggested in Stage 1 [will be the SAME / MAY OR MAY NOT BE THE SAME] amount that is confirmed in Stage 2.

Stages:

The game has 2 stages. Here is how the game works:

STAGE 1:

Country A suggests how much of the prize it wants for itself. The Prize is worth 10 points. Country A can demand 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 points of the Prize.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to go to Stage 2.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

STAGE 2:

Country A confirms the number of points it wants.

- Since there is [ENFORCEMENT / NO ENFORCEMENT], the computer [WILL NOT / WILL] ALLOW Country A to make changes in Stage 2.
- Hence, the amount which Country A suggested in Stage 1 [WILL BE THE SAME / MAY OR MAY NOT BE THE SAME] amount that is confirmed in Stage 2.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to the bargain.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

PAYOFFS:

If Country B agrees in both Stages 1 and 2,

- Country A will get the number of points it confirmed in Stage 2.
- Country B will get what is left of the Prize (i.e. 10 points minus the points Country A has taken).
- Both countries DO NOT LOSE POINTS in the costs of war.

*If Country B disagrees (war) in **Stage 1,***

- Country A gets **5 points minus its cost of war.** Country B gets **5 points minus its cost of war.**

*If Country B disagrees (war) in **Stage 2,***

- Country A gets **7 points minus its cost of war.** Country B gets **3 points minus its cost of war.**

Comparison:

To summarize, one major difference between Scenario 3 and the earlier scenarios is:

COUNTRY B CAN DECIDE TO WAGE WAR AT ANY TIME

Note:

Note that:

- Country B can wage war at any time.

- Both countries can avoid the costs of war if they can reach an agreement in Stage 2.

Prelude:

You are in Scenario 3. Here is the scenario summary:

STAGE 1:

Country A suggests how much of the prize it wants for itself.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to go to Stage 2.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

STAGE 2:

Country A confirms how much of the prize it wants for itself.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to the bargain.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

PAYOFFS:

If bargain is made,

- Country A gets what it confirmed in Stage 2.
- Country B gets what is left of the Prize.
- Both countries avoid the costs of war.

Country B can trigger war at any time:

If war is triggered in **Stage 1,**

- Country A gets **5 points minus its cost of war.**
- Country B gets **5 points minus its cost of war.**

If war is triggered in **Stage 2,**

- Country A gets **7 points minus its cost of war.**
- Country B gets **3 points minus its cost of war.**

Flag:

You are currently in Scenario 3.

Country Assignment:

In this round, you are randomly assigned as a leader in Country [A/B].

Treatment Summary:

There is [ENFORCEMENT / NO ENFORCEMENT].

- Country A [**cannot / can**] **make changes** in Stage 2.
- The amount Country A suggested in Stage 1 [will be the SAME / MAY OR MAY NOT BE THE SAME] amount confirmed in Stage 2.

The cost of war is PUBLIC information.

- Your Cost of War = **2 points** and Opponent's Cost of War = **2 points.**
- You know your Opponent's cost of war
- Your Opponent knows your cost of war

Actual Play:

[STAGE 1]

[Country A (decision)]

You are Country A.

Suggest how much of the prize you want by choosing ONE of the possibilities below.

A gets 10, B gets 0
A gets 9, B gets 1
A gets 8, B gets 2
A gets 7, B gets 3
A gets 6, B gets 4
A gets 5, B gets 5
A gets 4, B gets 6
A gets 3, B gets 7
A gets 2, B gets 8
A gets 1, B gets 9
A gets 0, B gets 10

You [CAN / CANNOT] CHANGE the amount in Stage 2.

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

Country B can trigger war at any time.

THIS STAGE – Both countries are equally powerful.
If war is triggered, you get **5 points minus your cost of war.**
Country B gets **5 points minus its cost of war.**

NEXT STAGE – You are more powerful.
If war is triggered, you get **7 points minus your cost of war.**
Country B gets **3 points minus its cost of war.**

[Country A (waiting)]

Please wait while Country B makes a decision.

[Country B (waiting)]

Country A has not made its decision yet.

You can either WAIT to see Country A's demand, or DISAGREE without seeing Country A's demand.

To wait, just do nothing until Country A has made its decision. To disagree, click the button below.

THIS STAGE – Both countries are equally powerful.
If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

NEXT STAGE – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

There is [ENFORCEMENT / NO ENFORCEMENT] in this scenario.

< DISAGREE – Go to War >

[For Country B (decision)]

Country A has suggested the following:

"A gets [], B gets []"

If you agree, you proceed to Stage 2 in which Country A will confirm its demand.

Country A [CAN / CANNOT] CHANGE the amount in Stage 2.

If you disagree, war is triggered.

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

THIS STAGE – Both countries are equally powerful.
If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

NEXT STAGE – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< AGREE – Go to Stage 2 >
< DISAGREE – Go to War >

[STAGE 2]

[Country A (decision)] [Enforcement group]

You are Country A.

In Stage 1, you suggested the following:

"A gets [], B gets []"

You cannot change this decision. Click < CONFIRM > to continue.

< CONFIRM >

[Country A (decision)] [No-enforcement group]

You are Country A.

In Stage 1, you suggested the following:

"A gets [], B gets []"

You can change or confirm the suggested amount by choosing ONE of the possibilities below:

A gets 10, B gets 0
A gets 9, B gets 1
A gets 8, B gets 2
A gets 7, B gets 3
A gets 6, B gets 4
A gets 5, B gets 5

A gets 4, B gets 6
A gets 3, B gets 7
A gets 2, B gets 8
A gets 1, B gets 9
A gets 0, B gets 10

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

Country B can trigger war at any time.

THIS STAGE – You are more powerful.
If war is triggered, you get **7 points minus your cost of war.**
Country B gets **3 points minus its cost of war.**

[Country B (waiting)]

Country A has not made its decision yet.

You can either WAIT to see Country A's demand, or DISAGREE without seeing Country A's demand.

To wait, just do nothing until Country A has made its decision. To disagree, click the button below.

THIS STAGE – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

There is [ENFORCEMENT / NO ENFORCEMENT] in this scenario.

< DISAGREE – Go to War >

[For Country B (decision)]

Country A has confirmed the following:

"A gets [], B gets []"

If you agree, the bargain is made.

If you disagree, war is triggered.

Your Cost of War = 2 points. Opponent's Cost of War = [2 points / Unknown].

THIS STAGE – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< AGREE – Bargain is Made >
< DISAGREE – Go to War >

Outcome:

[If Country B agrees]

Country B agreed.

Country A gets the points it demanded.

Country B gets what is left of the Prize.

[If Country B disagrees]

Country B disagreed. War is triggered.

Country A gets 7 points minus its cost of war.

Country B gets 3 points minus its cost of war.

Repeat Screens from “Prelude” to “Outcome” in Rounds 12-15

SCENARIO 4

Players are randomly divided into a "30-seconds" group and a "60-seconds" group.

Transition:

Scenario 3 is over. Scenario 4 will start in a few seconds.

Comparison:

One major difference between Scenario 3 and Scenario 4 is:

PAYOFFS FOR WAR DEPEND ON THE TIMER CLOCK

Payoffs:

You are a national leader bargaining with another country over a valuable Prize.

Both countries know that:

- The Prize is worth 10 points
- If there is a war, both countries will pay the costs of war
- Cost of war is 2 points for all countries. This cost is fixed throughout the remaining rounds. All countries know that the cost of war is 2 points for all countries.

War Costs:

War is always costly.

Whenever war is triggered, **you lose 2 points.**

Whenever war is triggered, **your opponent loses 2 points.**

Scenario Condition I:

There is **NO ENFORCEMENT** in this scenario.

- The computer WILL ALLOW Country A to make changes in Stage 2.
- Hence, the amount which Country A suggested in Stage 1 MAY OR MAY NOT BE THE SAME amount that is confirmed in Stage 2.

Scenario Condition II:

In this scenario:

COUNTRY B CAN DECIDE TO WAGE WAR AT ANY TIME.

Scenario Condition III:

In this scenario, the payoffs for war depend on the **TIMER CLOCK**.

UNTIL THE [30TH /60TH] SECOND, Country A and Country B are **EQUALLY POWERFUL**. If war is triggered,

- Country A gets 5 points minus its cost of war.
- Country B gets 5 points minus its cost of war.

AFTER THE [30TH /60TH] SECOND, Country A becomes **MORE POWERFUL**. If war is triggered,

- Country A gets 7 points minus its cost of war.
- Country B gets 3 points minus its cost of war.

Elaboration I:

Unlike earlier scenarios, the payoffs for war do **NOT** depend on whether you are in Stage 1 or Stage 2.

- As long as Country B disagrees (war) **AT OR BEFORE the [30TH /60TH] SECOND**, Country A gets 5 points minus its cost of war and Country B gets 5 points minus its cost of war. It DOES NOT MATTER whether war occurs in Stage 1 or Stage 2.

- As long as Country B disagrees (war) **AFTER the [30TH /60TH] SECOND**, Country A gets 7 points minus its cost of war and Country B gets 3 points minus its cost of war. It DOES NOT MATTER whether war occurs in Stage 1 or Stage 2.

The Timer Clock **does NOT stop ticking** until the end of the round. It resets only at the end of the round. It **does NOT reset** for each stage.

Stages:

The game has 2 stages. Here is how the game works:

STAGE 1:

Country A suggests how much of the prize it wants for itself. The Prize is worth 10 points.

Country A can demand 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 points of the Prize.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to go to Stage 2.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

STAGE 2:

Country A confirms the number of points it wants.

- Since there is NO ENFORCEMENT, the computer WILL ALLOW Country A to make changes in Stage 2.
- Hence, the amount which Country A suggested in Stage 1 MAY OR MAY NOT BE THE SAME amount that is confirmed in Stage 2.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to the bargain.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

PAYOFFS:

If Country B agrees in both Stages 1 and 2,

- Country A will get the number of points it confirmed in Stage 2.
- Country B will get what is left of the Prize (i.e. 10 points minus the points Country A has taken).
- Both countries DO NOT LOSE POINTS in the costs of war.

If Country B disagrees (war) ***at or before the [30^{th} / 60^{th}] second,***

- Country A gets **5 points minus its cost of war.** Country B gets **5 points minus its cost of war.**

If Country B disagrees (war) ***after the [30^{th} / 60^{th}] second,***

- Country A gets **7 points minus its cost of war.** Country B gets **3 points minus its cost of war.**

Comparison:

To summarize, one major difference between Scenario 3 and Scenario 4 is:

Payoffs for war depend on the TIMER CLOCK

Note:

Note that:

- Country B can wage war at any time.
- Both countries can avoid the costs of war if they can reach an agreement in Stage 2.

Prelude:

You are in Scenario 4. Here is the scenario summary:

STAGE 1:

Country A SUGGESTS how much of the prize it wants for itself.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to go to Stage 2.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

STAGE 2:

Country A CONFIRMS how much of the prize it wants for itself.

Country B has 3 options:

- WAIT to see Country A's demand, and then AGREE to the bargain.
- WAIT to see Country A's demand, and then DISAGREE (go to war).
- DISAGREE (go to war) without seeing Country A's demand.

PAYOFFS:

If bargain is made,

- Country A gets what it confirmed in Stage 2.
- Country B gets what is left of the Prize.
- Both countries avoid the costs of war.

Country B can trigger war at any time:

If war is triggered in ***at or before the [30^{th} / 60^{th}] second,***

- Country A gets **5 points minus its cost of war.**

- Country B gets **5 points minus its cost of war.**

If war is triggered in ***after the [30th / 60th] second,***

- Country A gets **7 points minus its cost of war.**
- Country B gets **3 points minus its cost of war.**

Reminder:

Remember:

Payoffs for War Depend on the Timer Clock Regardless of Whether You are in Stage 1 or 2.

The Timer Clock does NOT stop ticking until the end of the round.
It does NOT reset for each stage.

Flag:

You are currently in Scenario 4.

Country Assignment:

In this round, you are randomly assigned as a leader in Country [A/B].

Treatment Summary:

There is NO ENFORCEMENT.

- Country A **can make changes** in Stage 2.
- The amount Country A suggested in Stage 1 MAY OR MAY NOT BE THE SAME amount confirmed in Stage 2.

The cost of war is PUBLIC information.

- Your Cost of War = **2 points** and Opponent's Cost of War = **2 points.**
- You know your Opponent's cost of war
- Your Opponent knows your cost of war

Actual Play:

[STAGE 1]

[Country A (decision)]

TIMER CLOCK: [sec]

You are Country A.

Suggest how much of the prize you want by choosing ONE of the possibilities below.

A gets 10, B gets 0
A gets 9, B gets 1
A gets 8, B gets 2
A gets 7, B gets 3
A gets 6, B gets 4
A gets 5, B gets 5
A gets 4, B gets 6
A gets 3, B gets 7
A gets 2, B gets 8
A gets 1, B gets 9
A gets 0, B gets 10

You CAN CHANGE the amount in Stage 2.

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

Country B can trigger war at any time.

AT/BEFORE [30TH / 60TH] SECOND – Both countries are equally powerful.
If war is triggered, you get **5 points minus your cost of war.**
Country B gets **5 points minus its cost of war.**

AFTER [30TH / 60TH] SECOND – You are more powerful.
If war is triggered, you get **7 points minus your cost of war.**
Country B gets **3 points minus its cost of war.**

[Country A (waiting)]

Please wait while Country B makes a decision.

[Country B (waiting)]

Country A has not made its decision yet.

You can either WAIT to see Country A's demand, or DISAGREE without seeing Country A's demand.

To wait, just do nothing until Country A has made its decision. To disagree, click the button below.

AT/BEFORE [30TH / 60TH] SECOND – Both countries are equally powerful.

If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

AFTER [30TH / 60TH] SECOND – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

There is NO ENFORCEMENT in this scenario.

< DISAGREE – Go to War >

[For Country B (decision)]

Country A has suggested the following:

"A gets [], B gets []"

If you agree, you proceed to Stage 2 in which Country A will confirm its demand.

Country A CAN CHANGE the amount in Stage 2.

If you disagree, war is triggered.

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

AT/BEFORE [30TH / 60TH] SECOND – Both countries are equally powerful.
If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

AFTER [30TH / 60TH] SECOND – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< AGREE – Go to Stage 2 >
< DISAGREE – Go to War >

[STAGE 2]

[Country A (decision)]

You are Country A.

In Stage 1, you suggested the following:

"A gets [], B gets []"

You can change or confirm the suggested amount by choosing ONE of the possibilities below:

A gets 10, B gets 0
A gets 9, B gets 1
A gets 8, B gets 2
A gets 7, B gets 3
A gets 6, B gets 4
A gets 5, B gets 5
A gets 4, B gets 6
A gets 3, B gets 7
A gets 2, B gets 8
A gets 1, B gets 9
A gets 0, B gets 10

Your Cost of War = 2 points. Opponent's Cost of War = 2 points.

Country B can trigger war at any time.

AT/BEFORE [30TH / 60TH] SECOND – Both countries are equally powerful.
If war is triggered, you get **5 points minus your cost of war.**
Country B gets **5 points minus its cost of war.**

AFTER [30TH / 60TH] SECOND – You are more powerful.
If war is triggered, you get **7 points minus your cost of war.**
Country B gets **3 points minus its cost of war.**

[Country B (waiting)]

Country A has not made its decision yet.

You can either WAIT to see Country A's demand, or DISAGREE without seeing Country A's demand.

To wait, just do nothing until Country A has made its decision. To disagree, click the button below.

AT/BEFORE [30TH / 60TH] SECOND – Both countries are equally powerful.
If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

AFTER [30TH / 60TH] SECOND – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< DISAGREE – Go to War >

[For Country B (decision)]

Country A has confirmed the following:

"A gets [], B gets []"

If you agree, the bargain is made.

If you disagree, war is triggered.

Your Cost of War = 2 points. Opponent's Cost of War = [2 points / Unknown].

AT/BEFORE [30TH / 60TH] SECOND – Both countries are equally powerful.
If war is triggered, Country A gets **5 points minus its cost of war.**
You get **5 points minus your cost of war.**

AFTER [30TH / 60TH] SECOND – Country A is more powerful.
If war is triggered, Country A gets **7 points minus its cost of war.**
You get **3 points minus your cost of war.**

< AGREE – Bargain is Made >
< DISAGREE – Go to War >

Outcome:

[If Country B agrees]

Country B agreed.

Country A gets the points it demanded.

Country B gets what is left of the Prize.

[If Country B disagrees]

Country B disagreed. War is triggered.

Country A gets 7 points minus its cost of war.

Country B gets 3 points minus its cost of war.

APPENDIX 4

Supplementary Figures and Tables

Figure 2 in the main text shows the evolution in the incidence of war over 10 rounds. Figure A1 decomposes the data to reveal trends across different information conditions. In particular, the introduction of a commitment problem in the public-information condition caused the war incidence to jump from 33% (Round 5) to 67% (Round 6), while its removal caused the incidence of war to fall from 56% to 11%. In the private-information condition, war incidence increased from 11% to 67% when the commitment problem was introduced, and fell from 71% to 43% when it was removed. The data also show that the treatment effect is generally sharper in later rounds than in earlier rounds. This observation is consistent with the experimental literature. In game-theoretic experiments with multiple rounds, outcomes in the initial rounds are often more distant from the equilibrium predictions due to the subjects' unfamiliarity with the newly-introduced strategic environment. An interesting question is why do bargains fail in the enforcement condition in the beginning rounds. It turns out that 67%, 100% and 80% of the bargaining failures in Rounds 1, 2 and 3 respectively were caused by reservation-level offers, without which the treatment effect might have been sharper. In theory, Player A will make an offer to Player B based on the latter's reservation level in Stage 1, and B will accept if the offer is equal or greater than its reservation level. Of course, since the reservation level is the same as the war payoff, B is actually indifferent between the two options. Hence, reservation-level offers to B (3 points) will always run a risk of war. Since subjects are randomly assigned as either A or B in each round, the practical risks of reservation-level offers will only be apparent to all players after the first few rounds.

Figure A1: Incidence of War (By Information Conditions)

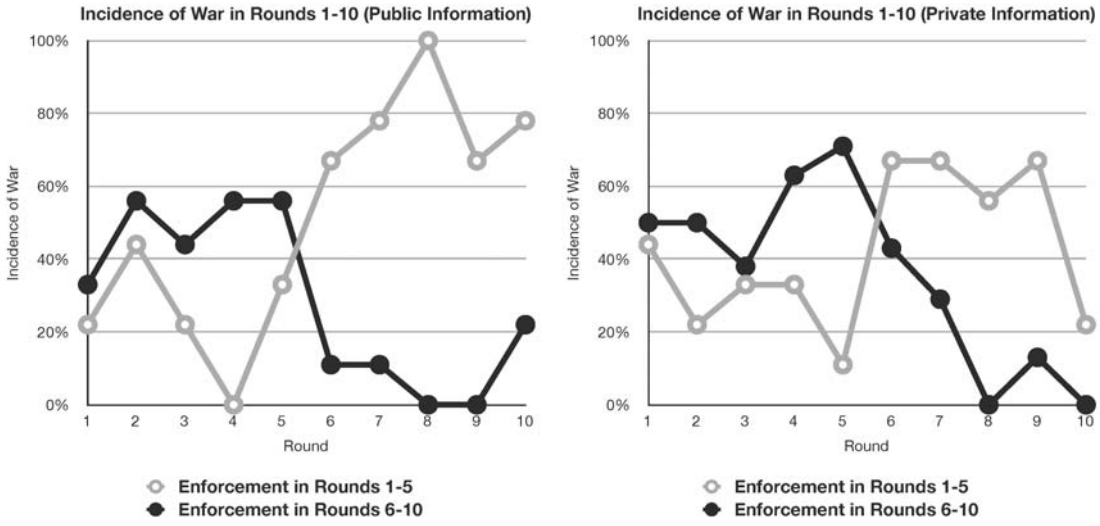

Table A1. Logit Estimates of Determinants for the Decision for War: Rounds 11-15

	(1)	(2)	(3)	(4)	(5)	(6)
Enforcement	-3.459*** (0.661)	-3.254*** (0.626)	-3.256*** (0.676)	-4.199*** (0.756)	-3.778*** (0.732)	-3.832*** (0.822)
Initial offer		-0.141 (0.266)	-0.142 (0.261)		0.058 (0.249)	0.049 (0.245)
Risk preference			-0.003 (0.230)			-0.114 (0.254)
Constant	0.975 (0.302)**	1.331 (1.313)	1.350 (1.755)	0.663 (0.702)	0.332 (1.470)	0.994 (1.997)
Round and session fixed-effects	No	No	No	Yes	Yes	Yes
Prob>Chi^2	0.000	0.000	0.000	0.000	0.001	0.000
Log-likelihood	-64.016	-58.767	-58.766	-56.110	-52.826	-52.664
Pseudo-R^2	0.366	0.322	0.322	0.445	0.391	0.393
Observations	151	137	137	151	137	137

Notes: *** $p \leq 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$. Player B can decide for war in these rounds without seeing Player A's initial offer. In parentheses are robust standard errors corrected for clustering at the subject level. Round and session dummies are used to control for round and session fixed effects.

APPENDIX 5

Case Analysis:

Japan's Calculations on the Eve of the Pacific War

I conduct a case study to investigate the historical realism of the commitment problem.[4] I focus my case study on three specific questions:

1) Does the commitment-problem logic operate in a real-world crisis triggered by an impending power shift?

2) How does the logic affect the decision process leading to war?

3) Does the case reveal important omissions in our theoretical model?

These questions must be addressed with a case that allows the decision process to be studied in close detail.[5] This requires a case with a sharp power shift as well as reliable archival records that detail the decision process. The case of Japan's deliberations on the eve of the Pacific War satisfies both requirements. First, the case is supported by archival evidence that is remarkably rich and reliable. For most wars, formal records of top-level decision-making are either unavailable or censored, forcing scholars to work with lower-

[4] I focus on the commitment problem to test whether the positive finding in the experiment also holds in historical crises. It is difficult to isolate the private-information effect with observational data: all historical crises have occurred in environments with private information, and it is hard to compare and measure variations across different dimensions of the information environment with historical data.

[5] This objective differs from using case studies to generalize whether the explanatory variable (commitment problem) has systematic correlation with the outcome (war decision). See, e.g., George and Bennett (2005), Gerring (2007), and Barnes and Weller (2011). Causation was experimentally evaluated in the earlier sections. Potential confounders and selection biases were eliminated by the random assignment of experimental treatments, allowing for a clean test of the theoretical mechanism.

level sources or postwar retrospections. This case is a rare exception that allows the decision process to be traced with high precision.

Second, the case includes a shock that created an unambiguous perception of an impending power shift.[6] The shock was generated by the U.S. economic blockade on Japan with a total oil embargo starting in August 1941. The embargo left Japan with no doubt about its impending decline, as Japan depended on foreign oil imports to sustain its war machine. This is sufficient for my purpose – which is simply to separate out two scenarios across which the magnitude of the explanatory variable was sharply changed, and compare the policy calculations before and after in close detail.[7]

To facilitate the replicability of the analysis, the (1) dependent variable, (2) explanatory variable, (3) study timeframe and (4) evidential boundaries must be defined. (1) My *dependent variable* is not war, but the intermediate decisions or arguments made by leaders for or against war. Thus, the case contains multiple observations within the study timeframe before and after the oil shock. The shock connects to the (2) *explanatory variable* – the perception of a future vulnerability that the opponent cannot commit not to exploit. The (3) *study timeframe* is between July 1941 just before the oil embargo to December 1941 when war was declared. My analysis focuses on the calculations of

[6] The increase in U.S. military preparations during 1941 had also alerted Japan to a potential power shift in the future. However, the U.S. military preparations were continuous in nature rather than a single discrete shock. Thus they did not create in Japan a perception of power shift that was as sudden and as unambiguous as what the oil embargo did in July-August 1941.

[7] The complexity of history seldom offers the convenience of exogeneity. Exogeneity arguments are easily challenged, as there is usually a long history of interactions between two states prior to a crisis that led them to war. This case study does not assume exogeneity, but it narrows the study timeframe and treats the shock in the Japanese case as a first approximation for analysis. Of course, the US-Japan conflict leading to the Pacific War did not suddenly start in July 1941 – it started as early as in 1931 if not earlier – but a tightened timeframe allows analysis to be focused and tractable.

Japanese leaders in response to the impending power shift created by the U.S. embargo. The inquiry is of a limited nature and does not cover the full historical processes leading to the Pacific War.[8] The purpose is not to explain the Pacific War in entirety or test the power of different theories, but simply to address the three questions stated above.

I restrict (4) my *data boundaries* to the records of official meetings between top leaders – the "imperial conferences" that approved key national decisions and the "liaison conferences" that ironed out differences among the cabinet ministers and military chiefs. The data include the four imperial conferences preceding the Pacific War and the 38th to 74th liaison conferences between the imperial conferences. The meeting notes have been translated (Ike 1967) and are accessible to researchers who wish to check the interpretations. Non-official remarks by top leaders and comments by lower-level officials are excluded, although they appear consistent with the findings here. It is relatively easy to infer a commitment problem retrospectively from the features of a case, or to select compelling quotes from historical personalities to show the relevance of the commitment problem. But the test is stronger if we examine the complete records of the actual meetings where decisions for war were debated among top leaders, and find that the commitment-problem logic was explicitly and repeatedly articulated.

[8] For a comprehensive historical narrative, see Iriye (1987).

Finding 1: Japan's decision for war involved strategic calculations that repeatedly applied the commitment-problem logic.

The commitment-problem logic is as follows: It is better to fight earlier than later, because in the later period the power balance shifts against one's favor, and the opponent cannot commit not to exploit its advantage in the future. The power shift was significantly triggered by the U.S. oil embargo that began on 1 August 1941.

Evidence from the September imperial conference supports Finding 1. The strategic logic was repeatedly articulated to justify a specific deadline for war. The logic fronted the opening statement of the Prime Minister, who commenced the proceedings, as well as the statements made by the second and third speakers at the conference.[9] For instance, the second speaker, Navy Chief of Staff Nagano Osami, warned: "A number of vital military supplies, including oil, are dwindling day by day. This will cause a gradual weakening of our national defense, and lead to a situation in which, if we maintain the status quo, the capacity of our Empire to act will be reduced in the days to come.... [I]t would be very dangerous for our Empire to remain idle and let the days go by."[10]

There was consensus that Japan's relative power would rapidly decline over time. Thus Japan must prepare to fight earlier than later. Otherwise, its political and military situation would only become worse later. Arguments along the same logic were repeated at the conference. Stemming the relative decline was crucial, since the U.S. could not

[9] Records of Imperial Conference, 6 September 1941, in Ike 1967, 138-41.
[10] Ibid, 139.

credibly commit not to exploit a weaker Japan in the future: “Even if we should make concessions to the United States by giving up part of our national policy for the sake of a temporary peace, the United States, its military position strengthened, is sure to demand more and more concessions on our part; and ultimately our Empire will have to lie prostrate at the feet of the United States.”[11]

Distrust of Western powers frequently surfaced in the imperial and liaison conferences. When strategic situations change, even formal agreements could be reneged. For example, when the German-Soviet war started, President of the Privy Council Hara Yoshimichi told the imperial conference in July: “Some people say that it would be improper for Japan to attack the Soviet Union in view of the Neutrality Pact; but the Soviet Union is notorious for her habitual acts of betrayal. If we were to attack the Soviet Union, no one would regard it as treachery.”[12] In August, Foreign Minister Toyoda Teijiro highlighted the warning from the German Ambassador: “You will be tricked by the United States, and negotiations will be drawn out; you had better break them off and avoid being tricked.”[13] The reference materials for the September imperial conference emphasized: “In short, military force should be used promptly if there is no prospect of diplomatic success. It is expected that the United States and Great Britain will try to delay us with diplomatic negotiations. We must be careful not to be inveigled into this trap.”[14]

---

[11] Reference Materials for Answering Questions at the Imperial Conference on 6 September Regarding “The Essentials for Carrying Out the Empire's Policies,” in Ike 1967, 152.

[12] Records of Imperial Conference, 2 July 1941, in Ike 1967, 87.

[13] Records of the 49th Liaison Conference, 30 August 1941, in Ike 1967, 126.

[14] Reference Materials for Answering Questions at the Imperial Conference on 6 September Regarding “The Essentials for Carrying Out the Empire's Policies,” in Ike 1967, 155.

What led to the urgent deterioration in relative power? Director of the Planning Board Suzuki Teiichi explained that because of the economic blockade, "our Empire's national power is declining day by day. Our liquid fuel stockpile, which is the most important, will reach bottom by June or July of next year …."[15] To the question "Why have we set the last ten days of October as a tentative deadline for war preparations?" the answer was:

> We need not repeat that at present oil is the weak point of our Empire's national strength and fighting power .... As time passes, our capacity to carry on war will decline, and our Empire will become powerless militarily. Meanwhile the naval and air forces of the United States will improve remarkably as time goes on …. Hence the passing of time not only means that we will face more difficulties in military operations, but also means that the increasing military preparedness of the United States Navy will surpass the naval power of our Empire after next autumn, and that we will finally be forced to surrender to the United States and Great Britain without a fight.[16]

Speed was the essence of the game. Navy Chief of Staff Nagano explained that the burdens of a prolonged war were best addressed if Japan could "seize the enemy's important military areas and sources of materials quickly at the beginning of the war, making our operational position tenable and at the same time obtaining vital materials from the areas now under hostile influence."[17] Victory depended on the success of the first stage of military operations, which in turn depended on three factors: "first, to decide quickly to commence hostilities in view of the realities of our fighting capacity and theirs; second, to take the initiative rather than to allow them to do so; third, to consider the meteorological conditions in the operational areas in order to make operations

[15] Records of Imperial Conference, 6 September 1941, in Ike 1967, 148.

[16] Reference Materials for Answering Questions at the Imperial Conference on 6 September Regarding "The Essentials for Carrying Out the Empire's Policies," in Ike 1967, 154.

[17] Records of Imperial Conference, 6 September 1941, in Ike 1967, 139-40.

easier."[18] Delaying the decision for war was equivalent to denying the Japanese military the critical advantages necessary for surviving or winning the war.

The September imperial conference marked the point at which Japan fixed a deadline for war. But after September, there were two more imperial conferences before the outbreak of war in December. Was the final decision for war influenced by a different set of calculations? To assess this possibility, I examined the records of the two imperial conferences after September. I found that the same strategic logic persisted. In the November conference, Army Chief of Staff Sugiyama emphasized: "[T]he ratio of armament between Japan and the United States will become more and more unfavorable to us as time passes; and particularly, the gap in air armament will enlarge rapidly. Moreover, [American war preparations] will make rapid progress. Also, [the] joint defensive capability of [United States, Great Britain, the Netherlands and China] will be rapidly increased … Thus it would be very disadvantageous for us to delay; and it is to be feared that it might become impossible for us to undertake offensive operations."[19] Prime Minister Tojo Hideki said at the close of the conference: "Two years from now we will have no petroleum for military use. Ships will stop moving … I fear that we would become a third-class nation after two or three years if we just sat tight."[20] In the December imperial conference, however, the proceedings no longer revolved around the arguments for war, since the deadline for war had been formalized. The die was already

[18] Ibid, 140.
[19] Ibid, 225-6.
[20] Ibid, 238.

cast.[21] The conference focused on the war preparations.[22] It ended with an imperial sanction to declare war against the United States.

The commitment-problem logic was activated by an impending power shift triggered by the U.S. oil embargo. Thus, we should expect a discernible difference in the strategic calculations at the imperial conferences *before* and *after* the embargo. I found that this was indeed the case.

At the July imperial conference before the embargo, the focus was on plans to establish the Greater East Asia Co-prosperity Sphere and settle the war in China expeditiously, as well as the question of war with the Soviet Union. However, "[i]n carrying out the plans outlined above … our Empire will not be deterred by the possibility of being involved in a war with Great Britain and the United States"; thus Japan should also prepare for that possibility.[23] Army Chief of Staff Sugiyama argued that "in carrying out various measures for the solution of the Northern Problem [against the Soviets]," Japan should keep to its "basic position of always being prepared for war with Great Britain and the United States, since the attitude of these countries toward Japan cannot be viewed with optimism."[24]

---

[21] President of the Privy Council Hara Yoshimichi opened the questions-and-answers session noting that "[w]e are discussing a very grave subject; but it was previously taken up by [the last] Imperial Conference, and every step that could be taken has been taken. Therefore, I have nothing in particular to add." Records of Imperial Conference, 1 December 1941, in Ike 1967, 279.

[22] Ibid, 271-8.

[23] "Outline of National Policies in View of the Changing Situation", 2 July 1941, in Ike 1967, 78.

[24] Records of Imperial Conference, 2 July 1941, in Ike 1967, 81. See also Navy Chief of Staff Nagano's comments in ibid, 81-2. The conference was aware that Japan might provoke Britain and the U.S. into war, but it was a negative scenario which Japan should try its best to avoid. Ibid, 83, 87, 88-9.

At this point, while Japanese leaders acknowledged the possibility of a war with Britain and the U.S., their belligerence did not center on the U.S. The conference heard no specific plan to attack the U.S., and there was no clear articulation of the strategic logic that was to be repeatedly emphasized in later imperial conferences.[25] In September, the situation changed. The oil embargo created a rapid and impending power decline, and the logic of the commitment problem gripped the minds of the decision-makers. Here, the commitment-problem theory successfully predicts both the content of the archival evidence as well as the *change* in the pattern of the evidence.

How was the agenda set for the September imperial conference? Of major importance was the September 3 liaison conference between Prime Minister Konoye Fumimaro, Foreign Minister Toyoda Teijiro, War Minister Tojo Hideki, and the military chiefs. At this seven-hour conference, Japanese leaders debated the policy proposals that were subsequently tabled at the imperial conference on September 6. The justifications for war were consistent with the logic of the commitment problem. Navy Chief of Staff Nagano, whose navy first drafted the policy proposals, began the meeting stating that "[i]n various respects the Empire is losing materials: that is, we are getting weaker. By contrast, the enemy is getting stronger. With the passage of time, we will get increasingly weaker, and we won't be able to survive.... Although I am confident that at the present time we have a

[25] Iriye (1987, 150) pointed out that "[a]lthough war with the combined ABCD powers had been envisaged for some time, as of early August there had been no comprehensive master plan."

chance to win a war, I fear that this opportunity will disappear with the passage of time."[26]

As seen in the cited examples, the element of speed was a key emphasis: Japan must make a swift decision because the situation was changing quickly against its favor. Hence, the September conference agreed that war preparations should be completed by late October, and the November conference fixed the deadline for war in early December. It appears that the strategic justifications for war were not driven only by the impending power shift, but also by the perceived *speed* of the shift. This provides the next finding:

> **Finding 2:** The perceived speed of shift in relative power was important in Japan's calculations. This factor was omitted in the model.

How does the speed of the power shift affect decisions for war? In a utility-based mechanism, the loss in the expected utility for war accelerates as the speed of decline in relative power increases. For the declining state, waiting costs increase as the speed of shift increases. We may term this a utility-based effect of the power shift. In the case of the Pacific War, evidence of the utility-based effect can be seen from much of the evidence presented above for Findings 1 and 2. However, the speed of the power shift may also create another effect that is not captured in the utility-based mechanism. A rapid power shift promotes hastened or truncated diplomacy.[27] Negotiations may be more likely to fail under the tight time pressure of a truncated bargaining timeframe. This

[26] Records of Liaison Conference, 3 September 1941, in Ike 1967, 130-1.
[27] See Van Evera 1999, 83-5.

suggests a behavioral mechanism that connects from the utility-based effect in the first mechanism. Due to the utility effect, states truncate the bargaining process with a tight timeline, which increases the risk of bargaining failure. We may call this the "truncation effect" of the power shift.

> **Finding 3:** The speed of shift in relative power made Japan truncate the bargaining timeframe.

The impending power shift led to the decision for a tight deadline for war, which translated into a tight timeline for diplomacy. The tight timeline severely restricted the possibility of a successful negotiation. This fact was pointed out repeatedly. For example, Foreign Minister Togo stated at the November imperial conference that "the situation is becoming more and more critical every day, and negotiations with the United States are very much restricted by the time element; consequently, to our regret, there is little room left for diplomatic maneuvering. Moreover, the conclusion of a Japanese-American understanding would necessitate great speed in negotiations, partly because of the time required for domestic procedures on the American side …. The prospects of achieving an amicable settlement in the negotiations are, to our deepest regret, dim."[28] Remarks that the tight timeline made diplomacy difficult were also made in various liaison conferences between the September and December imperial conferences.[29]

[28] Records of Imperial Conference, 5 November 1941, in Ike 1967, 214.
[29] See, e.g., Records of Liaison Conferences, in Ike 1967, 167-84, 199-207, 239-47.

The case evidence shows the truncation of the bargaining timeframe due to the perceived speed of shift in relative power. But it does not show if there is a truncation *effect* on the decision for war. Since we cannot observe counterfactuals, we cannot know if a hypothetical exogenous removal of the tight deadline for diplomacy would have led to a peaceful bargain between the U.S. and Japan. The problem of inference is compounded as the truncation effect theoretically follows the utility effect of power shift, and bargaining truncation often surfaces in conjunction with the utility effect. It is hard to parse out cleanly the truncation effect with historical data to show that the truncation of the bargaining timeframe is not merely incidental, but is in fact causally influential in the decision for war.

To circumvent the constraints in the observational data, I replicate the truncation mechanism in a laboratory setting and evaluate its significance experimentally. I begin by incorporating Findings 2 and 3 into the experimental setup. The basic idea is to make the war payoffs time-dependent. I conduct the experiment in the public-information/no-enforcement condition with the interaction modified. In Stage 1 of the game, Player A decides on its demand $x_1 \in [0, 10]$, where x_1 is an integer. At the same time, Player B decides whether to wait for the demand (and subsequently accept or reject it) or to wage a war. If B accepts x_1, the game enters Stage 2 in which A decides its demand $x_2 \in [0, 10]$. At the same time, Player B decides whether to wait for the demand (and subsequently accept or reject it) or to wage a war. If B accepts x_2, the game ends with payoff (x_2, 10 - x_2). In both periods, if B chooses to fight before or at the T-th second (where T is a positive integer), the game ends with payoff (5 - c_A, 5 - c_B), but if B chooses to fight after

the *T*-th second, the game ends with payoff (7 - c_A, 3 - c_B), where c_i is the cost of war for State *i*. Hence, the payoff shift is avoided if a bargain is made before the *T*-th second.[30]

This is a bargaining experiment that probes whether the speed of change in relative bargaining power affects the risk of war. In this experiment, the payoffs for war are time-dependent. In the control condition, the war payoff (5 - c_A, 5 - c_B) applies until the 60th second and the war payoff (7 - c_A, 3 - c_B) applies after the 60th second. In the treatment condition, the duration with war payoff (5 - c_A, 5 - c_B) is reduced to 30 seconds. The specific timings are calibrated based on simulated plays, with 30 seconds providing the time pressure to elicit the truncation effect, and 60 seconds providing some time buffer to hold off the truncation effect. Subjects were randomly assigned to treatment and control conditions. They were randomly assigned as either Player A or Player B and they played with a randomly assigned opponent. The observable implications are straightforward: If the truncation effect does not apply, we should not expect the war outcomes to differ between the treatment and control groups; but if it does, we should.

This was a one-round bargaining experiment (Round 16) that followed the previous experiment, with the same experimental setting and payment scheme. Hence, when the subject reached this experiment, he or she should be familiar with the structure of the bargaining game and be sufficiently prepared to handle the additional complexity. Before playing, subjects read instructions that highlighted the differences with the previous game

[30] As such, the predicted outcome differs from that of the earlier public-information model with no enforcement. A's offers will be based on B's reservation level, which is always higher before or at the *T*-th second than after the *T*-th second. Hence, B will not wait beyond the *T*-th second because A will always make a worse offer to B after the *T*-th second. Knowing this, A will make and confirm its offer based on B's reservation level before the *T*-th second, and B will accept. War is thus avoided.

(see Appendix 3). Meanwhile, the typical tradeoffs of using the same subject pool were reduced by experimental design: the stranger-matching protocol eliminated potential reciprocity effects; the use of a random payment mechanism and the revelation of earnings at the end of the session reduced potential endowment effects; and the restriction of each session to less than an hour reduced the likelihood of experimental fatigue. Nevertheless, it should be noted that this experiment is a single-trial preliminary probe without the full infrastructure for a decisive experimental test. Its purpose is simply to serve as an initial detection test for the truncation effect.

Finding 4: The truncation effect is likely to exist.

In the treatment group where a tight timeline was imposed to elicit the truncation effect, the percentage of war outcomes was 88% compared to 56% in the control group where the timeline was doubled to buffer the truncation effect. The difference in war outcomes across the two groups is significant (n = 35: Mann-Whitney test, p = 0.0350; two-tailed test of proportion, p = 0.0324) with dyadic binary outcomes as observations.

However, Finding 4 is not a decisive result. I set up a robustness test using Session 1, which ran the experiment for three rounds. The statistical significance behind Finding 4 is based on analysis that combined the single-round outcomes from Sessions 2 and 3 with the first-round outcomes from Session 1. First-round outcomes are most comparable to the single-round outcomes, since they were similarly generated by subjects playing the game for the first time. Hence, combining the single-round outcomes with the first-round

outcomes from Session 1 provides the most comparable combination of observations. Nonetheless, Finding 4 is not robust unless it passes an additional hurdle: that every possible combination of outcomes yields statistically significant *p*-values. This is not the case. When single-round outcomes in Sessions 2 and 3 are combined with second-round and third-round outcomes in Session 1, the differences in war outcomes across the two groups (76% in control group against 67% in treatment group with second-round outcomes in Session 1; and 82% in control group against 61% in treatment group with third-round outcomes in Session 1) are consistent with Finding 4 but statistically insignificant ($n = 35$: Mann-Whitney test, $p = 0.5271$ and $p = 0.1706$ respectively; two-tailed test of proportion, $p = 0.5211$ and $p = 0.1644$ respectively). On the whole, this preliminary probe gives us grounds to suspect that the truncation effect exists. But Finding 4 should be treated as indicative rather than decisive. More experiments are necessary to test the robustness of this particular result.