Collusion Among Adversaries

Matt Malis

Political Economy of Conflict Workshop
Vanderbilt University
4 March 2025

Slides: `mattmalis.github.io/slides`

The Puzzle of Provocation

Intuitive logic of provocation:

- A wants conflict, but wants B to attack first
- A says or does something (essentially costless) to "provoke" B
- B attacks, fighting ensues

The Puzzle of Provocation

Intuitive logic of provocation:

- A wants conflict, but wants B to attack first
- A says or does something (essentially costless) to "provoke" B
- B attacks, fighting ensues

Historical examples:

- Franco-Prussian War, 1870
  - Ems Dispatch: "effect of a red rag upon the Gallic bull"

The Puzzle of Provocation

Intuitive logic of provocation:

- A wants conflict, but wants B to attack first
- A says or does something (essentially costless) to "provoke" B
- B attacks, fighting ensues

Historical examples:

- Franco-Prussian War, 1870
  - Ems Dispatch: "effect of a red rag upon the Gallic bull"
- South Korean martial law episode, 2024
  - "provoke the North's attack at the NLL [Northern Limit Line]"

The Puzzle of Provocation

Intuitive logic of provocation:

- A wants conflict, but wants B to attack first
- A says or does something (essentially costless) to "provoke" B
- B attacks, fighting ensues

Historical examples:

- Franco-Prussian War, 1870
  - Ems Dispatch: "effect of a red rag upon the Gallic bull"
- South Korean martial law episode, 2024
  - "provoke the North's attack at the NLL [Northern Limit Line]"
- US entry into WWII; Gulf of Tonkin; Iraq invasion

The Puzzle of Provocation

Puzzle:

- How can it be both in A's interest to *provoke* B...
  - ...and in B's interest to *be provoked*?

The Puzzle of Provocation

Puzzle:

- How can it be both in A's interest to *provoke* B...
  - ...and in B's interest to *be provoked*?
- How can costless communication between adversaries be informative and influential?

Overview

This paper:

- Formal model of diplomacy and domestic politics

Overview

This paper:

- Formal model of diplomacy and domestic politics
- Demonstrates how *provocation* can be:
  - in the interest of both leaders
  - achieved through cheap-talk communication (private or public)

Overview

This paper:

- Formal model of diplomacy and domestic politics
- Demonstrates how *provocation* can be:
  - in the interest of both leaders
  - achieved through cheap-talk communication (private or public)
- Intuition:
  - A and B collude against A's domestic opposition, to extract a divisible benefit

Overview

This paper:

- Formal model of diplomacy and domestic politics
- Demonstrates how *provocation* can be:
  - in the interest of both leaders
  - achieved through cheap-talk communication (private or public)
- Intuition:
  - A and B collude against A's domestic opposition, to extract a divisible benefit
  - B making the first move toward conflict can be:
    - advantageous to B, for security reasons
    - advantageous to A, for domestic political reasons

Overview

This paper:

- Formal model of diplomacy and domestic politics
- Demonstrates how *provocation* can be:
  - in the interest of both leaders
  - achieved through cheap-talk communication (private or public)
- Intuition:
  - A and B collude against A's domestic opposition, to extract a divisible benefit
  - B making the first move toward conflict can be:
    - advantageous to B, for security reasons
    - advantageous to A, for domestic political reasons
  - communication allows them to coordinate their actions to realize these benefits

Overview

Outline:

- Model without communication (i.e. without provocation)
- Model with private communication
- Extension: public communication
- Cases

Game setup, without communication

Three players:

- leader A
- domestic audience D (in A's country)
- leader/state B (unitary actor)

Game setup, without communication

Three players:

- leader A
- domestic audience D (in A's country)
- leader/state B (unitary actor)

Sequence:

1. A and B types drawn by nature, observed privately

Game setup, without communication

Three players:

- leader A
- domestic audience D (in A's country)
- leader/state B (unitary actor)

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$

Game setup, without communication

Three players:

- leader A
- domestic audience D (in A's country)
- leader/state B (unitary actor)

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

Game setup, without communication

Three players:

- leader A
- domestic audience D (in A's country)
- leader/state B (unitary actor)

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$

Game setup, without communication

Three players:

- leader A
- domestic audience D (in A's country)
- leader/state B (unitary actor)

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A and B simultaneously: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Interpretations

Main interpretation:

- D as voter, or ruling coalition member
- can retain or remove A as leader

Interpretations

Main interpretation:

- D as voter, or ruling coalition member
- can retain or remove A as leader

Alternative interpretation 1 (formally equivalent):

- D as civilian leader, A as military

Interpretations

Main interpretation:

- D as voter, or ruling coalition member
- can retain or remove A as leader

Alternative interpretation 1 (formally equivalent):

- D as civilian leader, A as military

Alternative interpretation 2 (some formal changes):

- D as legislature, can authorize war or not

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A (or A') and B: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict ($z = 1$) or not ($z = 0$)
   - build arms; mobilize troops; seek external support; first strike
   - action that improves B's conflict payoffs, worsens A's
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader ($r = 1$) or replace her ($r = 0$)
4. A (or A') and B: cooperate ($d_i = 0$) or defect ($d_i = 1$)

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type
   - $Pr(x = \tilde{H}|\theta = H) = Pr(\tilde{L}|L) = \tau \in (\frac{1}{2}, 1)$

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A (or A') and B: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type
   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
   - if replace: draw new A', from same distribution as A
4. A (or A') and B: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A (or A') and B: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Figure: Conflict payoffs $W_i(d_A, d_B; z)$

	$d_B = 0$	$d_B = 1$
$d_A = 0$	$0, 0$	$p_B - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_B) - c_B^t + z\delta$
$d_A = 1$	$p_A - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_A) - c_B^t + z\delta$	$p_{AB} - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_{AB}) - c_B^t + z\delta$

Figure: Conflict payoffs $W_i(d_A, d_B; z)$

	$d_B = 0$	$d_B = 1$
$d_A = 0$	$0, 0$	$p_B - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_B) - c_B^t + z\delta$
$d_A = 1$	$p_A - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_A) - c_B^t + z\delta$	$p_{AB} - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_{AB}) - c_B^t + z\delta$

Figure: Conflict payoffs $W_i(d_A, d_B; z)$

	$d_B = 0$	$d_B = 1$
$d_A = 0$	$0, 0$	$p_B - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_B) - c_B^t + z\delta$
$d_A = 1$	$p_A - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_A) - c_B^t + z\delta$	$p_{AB} - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_{AB}) - c_B^t + z\delta$

- $p_A > p_{AB} > p_B$
  - first-strike advantage, or enhanced bargaining leverage
- if i expects j to defect, i's BR is defect

Figure: Conflict payoffs $W_i(d_A, d_B; z)$

	$d_B = 0$	$d_B = 1$
$d_A = 0$	$0, 0$	$p_B - \alpha_i c_A^\theta - z\lambda$, $(1 - p_B) - c_B^t + z\delta$
$d_A = 1$	$p_A - \alpha_i c_A^\theta - z\lambda$, $(1 - p_A) - c_B^t + z\delta$	$p_{AB} - \alpha_i c_A^\theta - z\lambda$, $(1 - p_{AB}) - c_B^t + z\delta$

- mobilizing improves B's conflict payoffs $(+z\delta)$
  - worsens A's conflict payoffs $(-z\lambda)$
- $\delta > p_{AB} - p_B$
  - preparing in advance $>$ catching A off-guard

Figure: Conflict payoffs $W_i(d_A, d_B; z)$

	$d_B = 0$	$d_B = 1$
$d_A = 0$	$0, 0$	$p_B - \alpha_i c_A^{\theta} - z\lambda$, $(1 - p_B) - c_B^t + z\delta$
$d_A = 1$	$p_A - \alpha_i c_A^{\theta} - z\lambda$, $(1 - p_A) - c_B^t + z\delta$	$p_{AB} - \alpha_i c_A^{\theta} - z\lambda$, $(1 - p_{AB}) - c_B^t + z\delta$

- A types, $\theta \in \{L, H\}$, with $c_A^L < c_A^H$
  - prior $Pr(A_H) = \pi_A$
- B types, $t \in \{\ell, m, h\}$, with $c_B^\ell < c_B^m < c_B^h$
  - prior $\pi_B^\ell + \pi_B^m + \pi_B^h = 1$
- low types = more "hawkish": **strictly prefer defecting**
- high/moderate types = more "dovish", conflict-averse
  - pref. for coop/defect **conditional on other side's action**

Figure: Conflict payoffs $W_i(d_A, d_B; z)$

	$d_B = 0$	$d_B = 1$
$d_A = 0$	$0, 0$	$p_B - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_B) - c_B^t + z\delta$
$d_A = 1$	$p_A - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_A) - c_B^t + z\delta$	$p_{AB} - \alpha_i c_A^\theta - z\lambda,$ $(1 - p_{AB}) - c_B^t + z\delta$

- $\alpha_A = 1$, and $\alpha_D >> 1$
- D shares A_H's preference for mutual cooperation
- but, conditional on conflict happening, D prefers A_L in office

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A and B simultaneously: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A and B simultaneously: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Total payoffs:

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A and B simultaneously: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Total payoffs:

- $U_A = r(\psi + W_A)$
  - officeholding value ψ large

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A and B simultaneously: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Total payoffs:

- $U_A = r(\psi + W_A)$
  - officeholding value ψ large
- $U_D = W_D$

Game setup, without communication

Sequence:

1. A and B types drawn by nature, observed privately
2. B: mobilize for conflict $(z = 1)$ or not $(z = 0)$
3. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader $(r = 1)$ or replace her $(r = 0)$
4. A and B simultaneously: cooperate $(d_i = 0)$ or defect $(d_i = 1)$

Total payoffs:

- $U_A = r(\psi + W_A)$
  - officeholding value ψ large
- $U_D = W_D$
- $U_B = W_B - z\kappa$
  - direct cost of mobilizing $\kappa \geqslant 0$

Equilibrium, without communication

Non-communication equilibrium

- B's mobilization strategy:
  - B^{ℓ} always mobilizes ($z = 1$)
  - B^{h} never mobilizes ($z = 0$)
  - B^{m} mobilizes iff $\pi_A < \overline{\pi}_A$ ("low-trust" environment)
- D retains if signal of A's type matches B's action, i.e.:
  - $r = 1$ if $(x = \tilde{H}, z = 0)$ or $(x = \tilde{L}, z = 1)$
  - $r = 0$ otherwise
- Conflict strategies:
  - B defect iff mobilized
  - A_L always defect; A_H defect iff B mobilized

Non-communication equilibrium – intuition

A_L and B^ℓ, always defect:

- trivial (assumed preference for defecting)

Non-communication equilibrium – intuition

A_L and B^ℓ, always defect:

- trivial (assumed preference for defecting)

A_H and B^m, defect iff B mobilized:

- $z = 1 \implies$ high $Pr(B^\ell)$
- in reality, if $t \neq \ell$, both sides would prefer mutual cooperation
  - but no way for them to know this!
- defensive mobilization $\implies$ mutual mistrust, mutual defection

Non-communication equilibrium – intuition

Effects of B's mobilization:

1. improves B's conflict payoffs

Non-communication equilibrium – intuition

Effects of B's mobilization:

1. improves B's conflict payoffs
2. signals B's hostility $\Longrightarrow$ influences A_H behavior

Non-communication equilibrium – intuition

Effects of B's mobilization:

1. improves B's conflict payoffs
2. signals B's hostility $\implies$ influences A_H behavior
3. signals B's hostility $\implies$ makes D demand A_L over A_H

Non-communication equilibrium – intuition

Effects of B's mobilization:

1. improves B's conflict payoffs
2. signals B's hostility $\implies$ influences A_H behavior
3. signals B's hostility $\implies$ makes D demand A_L over A_H

B^ℓ never mobilizes:

- reducing Pr(conflict) $>>$ being prepared for conflict

Non-communication equilibrium – intuition

Effects of B's mobilization:

1. improves B's conflict payoffs
2. signals B's hostility $\implies$ influences A_H behavior
3. signals B's hostility $\implies$ makes D demand A_L over A_H

B^ℓ never mobilizes:

- reducing Pr(conflict) $>>$ being prepared for conflict

B^m, mobilize iff prior trust is low ($\pi_A < \overline{\pi}_A$):

- balancing value of reducing Pr(conflict) vs. being prepared

Inefficiency of the non-communication equilibrium

In a "high-trust" environment:

- B^m would want to mobilize if he knew A was A_L
  - doesn't know A's type, so defaults to $z = 0$

Inefficiency of the non-communication equilibrium

In a "high-trust" environment:

- B^m would want to mobilize if he knew A was A_L
  - doesn't know A's type, so defaults to $z = 0$
- A_L wants her audience to believe B is a threat
  - argument is undermined by B^m not mobilizing

Inefficiency of the non-communication equilibrium

In a "high-trust" environment:

- B^m would want to mobilize if he knew A was A_L
  - doesn't know A's type, so defaults to $z = 0$
- A_L wants her audience to believe B is a threat
  - argument is undermined by B^m not mobilizing

Symmetrical problems in the "low-trust" environment

$\rightarrow$ mitigated by communication between A and B

Game setup, with communication

Sequence:

1. A and B types drawn by nature, observed privately
2. A: send private, costless message to B
   - conciliatory ($s = 0$) or hostile ($s = 1$)
3. B: mobilize for conflict ($z = 1$) or not ($z = 0$)
4. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader ($r = 1$) or replace her ($r = 0$)
5. A and B simultaneously: cooperate ($d_i = 0$) or defect ($d_i = 1$)

All other game features same as before

Game setup, with communication

Sequence:

1. A and B types drawn by nature, observed privately
2. A: send private, costless message to B
   - conciliatory ($s = 0$) or hostile ($s = 1$)
3. B: mobilize for conflict ($z = 1$) or not ($z = 0$)
4. D observes: B's action z, and signal x of A's type

   D: retain the incumbent leader ($r = 1$) or replace her ($r = 0$)
5. A and B simultaneously: cooperate ($d_i = 0$) or defect ($d_i = 1$)

All other game features same as before

Equilibrium with informative communication

Communication equilibrium

- A strategy:
  - A_L sends hostile message ($s = 1$)
  - A_H sends conciliatory message ($s = 0$)
- B strategy:
  - B^ℓ always mobilizes ($z = 1$)
  - B^h never mobilizes ($z = 0$)
  - B^m mobilizes iff receives hostile message ($z = s$)
- D strategy (same as before):
  - retain if signal of A's type matches B's action, i.e.:
    - $r = 1$ if $(x = \tilde{H}, z = 1)$ or $(x = \tilde{L}, z = 0)$
    - $r = 0$ otherwise
- Conflict strategies:
  - A_L and B^ℓ always defect
  - other types: cooperate only if
    - (i) B did not mobilize, and
    - (ii) A sent $s = 0$, or new A' was selected

Equilibrium with vs. without communication

Figure: Path-of-play conflict behavior, (d_A, d_B)

Non-communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,z$	$1,0$
	A_H	$1,1$	z,z	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ 1, & t = m, \text{ "low trust"} \\ 0, & t = m, \text{ "high trust"} \\ 0, & t = h \end{cases}$$

Private communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,1$	$1,1$
	A_H	$1,1$	$0,0$	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ s, & t = m \\ 0, & t = h \end{cases}, \quad s = \mathbb{1}[\theta = L]$$

Equilibrium with vs. without communication

Figure: Path-of-play conflict behavior, (d_A, d_B)

Non-communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,z$	$1,0$
	A_H	$1,1$	z,z	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ 1, & t = m, \text{ "low trust"} \\ 0, & t = m, \text{ "high trust"} \\ 0, & t = h \end{cases}$$

Private communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,1$	$1,1$
	A_H	$1,1$	$0,0$	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ s, & t = m \\ 0, & t = h \end{cases}, \quad s = \mathbb{1}[\theta = L]$$

Equilibrium with vs. without communication

Figure: Path-of-play conflict behavior, (d_A, d_B)

Non-communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,z$	$1,0$
	A_H	$1,1$	z,z	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ 1, & t = m, \text{ "low trust"} \\ 0, & t = m, \text{ "high trust"} \\ 0, & t = h \end{cases}$$

Private communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,1$	$1,1$
	A_H	$1,1$	$0,0$	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ s, & t = m \\ 0, & t = h \end{cases}, \quad s = \mathbb{1}[\theta = L]$$

Equilibrium with vs. without communication

Figure: Path-of-play conflict behavior, (d_A, d_B)

Non-communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,z$	$1,0$
	A_H	$1,1$	z,z	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ 1, & t = m, \text{ "low trust"} \\ 0, & t = m, \text{ "high trust"} \\ 0, & t = h \end{cases}$$

Private communication eqm

		B^ℓ	B^m	B^h
(r=1)	A_L	$1,1$	$1,1$	$1,1$
	A_H	$1,1$	$0,0$	$0,0$
(r=0)	A'_L	$1,1$	$1,z$	$1,0$
	A'_H	$1,1$	z,z	$0,0$

$$z = \begin{cases} 1, & t = \ell \\ s, & t = m \\ 0, & t = h \end{cases}, \quad s = \mathbb{1}[\theta = L]$$

Informative communication – incentive-compatibility

A_H incentive for non-provocation ($s = 0$):

Informative communication – incentive-compatibility

A_H incentive for non-provocation ($s = 0$):

- politically advantageous, and improves conflict payoffs

Informative communication – incentive-compatibility

A_H incentive for non-provocation ($s = 0$):

- politically advantageous, and improves conflict payoffs

A_L incentive for provocation ($s = 1$):

Informative communication – incentive-compatibility

A_H incentive for non-provocation ($s = 0$):

- politically advantageous, and improves conflict payoffs

A_L incentive for provocation ($s = 1$):

- B mobilizing is strictly harmful, for A_L's conflict payoffs
- but, demonstrates to audience that B poses a threat
  - so they need a leader like A, to manage the threat
- beneficial for A_L if office-holding value ψ large

Informative communication – incentive-compatibility

B^m following A's message:

Informative communication – incentive-compatibility

B^m following A's message:

- if he learns that A is type A_H (given message $s = 0$):
  - can help A_H survive in office, by showing D he is not a threat
  - if conflict can be avoided, mobilizing is unnecessary

Informative communication – incentive-compatibility

B^m following A's message:

- if he learns that A is type A_H (given message $s = 0$):
  - can help A_H survive in office, by showing D he is not a threat
  - if conflict can be avoided, mobilizing is unnecessary
- if he learns that A is type A_L (given message $s = 1$):
  - possible that he could undermine A_L and get A'_H, but unlikely
  - conflict is likely, so better to be prepared

Informative communication – incentive-compatibility

B^m following A's message:

- if he learns that A is type A_H (given message $s = 0$):
  - can help A_H survive in office, by showing D he is not a threat
  - if conflict can be avoided, mobilizing is unnecessary
- if he learns that A is type A_L (given message $s = 1$):
  - possible that he could undermine A_L and get A'_H, but unlikely
  - conflict is likely, so better to be prepared
- both conditions satisfied if c_B^m in intermediate range
  - $\implies$ message is influential, $z = s$

Extension: Public communication

What if A's message to B is public (observable by D)?

Extension: Public communication

What if A's message to B is public (observable by D)?

Equilibrium with public communication
Suppose office-holding incentives are dominant ($\psi \to \infty$). Suppose $\pi^\ell > \pi^h$. Then there exists an equilibrium similar to the private communication equilibrium, with the following exceptions:

- A_H mixes her messages, with $Pr(s = 1) = \frac{1-\tau}{\tau}$
- If D observes ($s = 1$, $x = \tilde{H}$), then D retains A with

$$Pr(r = 1) = \frac{1}{\tau}\left(\pi^h + \tau\pi^m - (1-\tau)\pi^l\right)$$

If $\pi^\ell < \pi^h$, a symmetrical equilibrium exists.

Extension: Public communication

D prefers mutual cooperation over conflict.

Why do they let A get away with provocative behavior?

Extension: Public communication

D prefers mutual cooperation over conflict.

Why do they let A get away with provocative behavior?

- Because D is *prospective*
  - doesn't matter who "started" the conflict, or who is "responsible" for the onset of tensions
- Provocation + mobilization $\Longrightarrow$ cooperation impossible
  - Given that conflict is inevitable, D wants leader who can manage it most effectively

Extension: Public communication

D prefers mutual cooperation over conflict.

Why do they let A get away with provocative behavior?

- Because D is *prospective*
  - doesn't matter who "started" the conflict, or who is "responsible" for the onset of tensions
- Provocation + mobilization $\Longrightarrow$ cooperation impossible
  - Given that conflict is inevitable, D wants leader who can manage it most effectively

Implication: A_L does not have to "deceive" her audience

- can provoke openly, and still be politically rewarded for it

Influential communication

To summarize:

- A and B collude against A's domestic opposition, to extract a divisible benefit

Influential communication

To summarize:

- A and B collude against A's domestic opposition, to extract a divisible benefit
  - A gives B a security benefit:
    - lets B know whether he should or shouldn't mobilize

Influential communication

To summarize:

- A and B collude against A's domestic opposition, to extract a divisible benefit
  - A gives B a security benefit:
    - lets B know whether he should or shouldn't mobilize
  - B gives A a political benefit:
    - helps convince A's audience that A is the right type of leader for the moment

Influential communication

To summarize:

- A and B collude against A's domestic opposition, to extract a divisible benefit
  - A gives B a security benefit:
    - lets B know whether he should or shouldn't mobilize
  - B gives A a political benefit:
    - helps convince A's audience that A is the right type of leader for the moment
- communication allows them to coordinate their actions to realize these benefits

Cases

Franco-Prussian War, 1870:

- Bismarck (A) wanted smaller German states (D) to support unification under Prussia ($r = 1$)
  - needed to demonstrate that French Emperor Napoleon III (B) had hostile intent

Cases

Franco-Prussian War, 1870:

- Bismarck (A) wanted smaller German states (D) to support unification under Prussia ($r = 1$)
  - needed to demonstrate that French Emperor Napoleon III (B) had hostile intent
- Ems Dispatch ($s = 1$):
  - costless message that communicated Bismarck's hostile intent

Cases

Franco-Prussian War, 1870:

- Bismarck (A) wanted smaller German states (D) to support unification under Prussia ($r = 1$)
  - needed to demonstrate that French Emperor Napoleon III (B) had hostile intent
- Ems Dispatch ($s = 1$):
  - costless message that communicated Bismarck's hostile intent
- France "mobilizing" ($z = 1$):
  - initiating war on its own timeline, rather than waiting
  - advantageous for France—power shifting towards Prussia

Cases

Franco-Prussian War, 1870:

- Bismarck (A) wanted smaller German states (D) to support unification under Prussia ($r = 1$)
  - needed to demonstrate that French Emperor Napoleon III (B) had hostile intent
- Ems Dispatch ($s = 1$):
  - costless message that communicated Bismarck's hostile intent
- France "mobilizing" ($z = 1$):
  - initiating war on its own timeline, rather than waiting
  - advantageous for France—power shifting towards Prussia

(Complication: Napoleon had his own domestic politics, and wanted Bismarck to initiate the war...)

Cases

South Korean martial law episode, 2024:

- Yoon government (A) wanted to overcome domestic gridlock
  - needed to demonstrate to public that NK (B) was a threat

Cases

South Korean martial law episode, 2024:

- Yoon government (A) wanted to overcome domestic gridlock
  - needed to demonstrate to public that NK (B) was a threat
- "Provoke the North's attack at the NLL" ($s = 1$)
  - drone flights dropping propaganda leaflets
  - shooting down trash balloons

Cases

South Korean martial law episode, 2024:

- Yoon government (A) wanted to overcome domestic gridlock
  - needed to demonstrate to public that NK (B) was a threat
- "Provoke the North's attack at the NLL" ($s = 1$)
  - drone flights dropping propaganda leaflets
  - shooting down trash balloons
- Provocation unsuccessful; Kim Jong Un did not mobilize
  - Yoon's martial law attempt failed, because no external threat
  - Kim Jong Un not actually a hostile type ($t \neq \ell$)?
    - (relative to prior expectations)

Contribution

Distinction from **audience costs**:

- message is intended to *provoke*, rather than *deter* as in AC
- message can be private or public

Contribution

Distinction from **audience costs**:

- message is intended to *provoke*, rather than *deter* as in AC
- message can be private or public

Distinction from **diversionary war**:

- insufficient for A to show hawkishness / competence in conflict
- must also show that the international environment (i.e. B's type) makes those attributes valuable

Contribution

Distinction from **audience costs**:

- message is intended to *provoke*, rather than *deter* as in AC
- message can be private or public

Distinction from **diversionary war**:

- insufficient for A to show hawkishness / competence in conflict
- must also show that the international environment (i.e. B's type) makes those attributes valuable

Novel mechanism of cheap-talk diplomacy between adversaries:

- coordinating action to collude against a third party

Thank you!

- Matt Malis, Texas A&M University
- Comments welcome and appreciated: `malis@tamu.edu`