Feigning Weakness

Branislav L. Slantchev

Department of Political Science
UCSD

2008

The Chinese Intervention in the Korean War

> *"It is not easy to explain why the Chinese entered North Korea so secretly and so suddenly. [. . .] They chose instead to launch a surprise attack, with stunning tactical advantages but no prospect of deterrence."*
>
> Thomas Schelling

Costly Signaling in Crisis Bargaining

A signal is credible if a weak actor is unwilling/unable to mimic it.

(a) Such signals must usually:

- involve high risk of war
- be very costly (immediately or later)

(b) Some mechanisms for credible signaling:

- sinking costs (Fearon 1997)
- tying hands (audience costs, Fearon 1994)
- autonomous risk of war (Schelling 1960)
- military mobilization (Slantchev 2005)
- domestic political actors (Schultz 1998)
- foreign political actors (Sartori 2002)

In other words. . .

- a strong actor never wants to pretend to be weak
- absence of costly signal is *prima facie* evidence of weakness

Sun Tzu's Principle

> *"If your opponent is of choleric temper, seek to irritate him. Pretend to be weak, that he may grow arrogant."*
>
> Sun Tzu

Was Sun Tzu wrong?
Does not look like it when it comes to **fighting**:

- Warfare is costly, so always conserve effort

  $\Rightarrow$ less effort if A believes B is weak

  $\Rightarrow$ a strong B can take advantage of A's belief

. . . but what about bargaining **before** fighting?

Implications for Crisis Bargaining

Contradictory incentives for a strong actor:

- during crisis: wants opponent to believe he's strong (so she agrees to larger concessions)
- if negotiations break down: wants opponent to believe he's weak (so she expends lower effort fighting)

Seems that the strong actor must somehow simultaneously signal strength and weakness.

How Can We Study This?

A minimalist model should have:

- bargaining in the shadow of power
- endogenous distribution of power in war-fighting
- fighting decisions depend on information gleaned from crisis

Model structure:

- Fearon's original take-it-or-leave-it (TILI) crisis game
- war-fighting as costly probabilistic contest in efforts
- effort depends on beliefs that may be based on crisis behavior

What Will We See?

Show that:

a) the strong actor benefits from opponent thinking him weak when war begins (not surprising, but nice)

b) this causes the strong actor to pretend to be weak during the crisis with positive probability

The Model: Structure & Payoffs

(a) Bargaining (Ultimatum) Phase:

- two risk-neutral players, 1 and 2
- bargain over division of benefit $[0, 1]$
- player 1 makes TILI offer $(x, 1 - x)$, $x \in [0, 1]$
- if player 2 accepts, game ends (payoffs from shares)
- if player 2 rejects, war begins

(b) Contest (War) Phase:

- players simultaneously spend effort, $m_i \geq 0$
- victory determined by technology of war:

$$\pi_i(m_1, m_2) = \begin{cases} \frac{m_i}{m_1 + m_2} & \text{if } m_1 + m_2 > 0 \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

- payoff: $\pi_i(m_1, m_2) - \frac{m_i}{c_i}$

The Model: Information

Two-sided incomplete information about costs of effort:

- each player knows own costs;
- player 1 believes player 2 is strong, $\overline{c}_2 > 0$, with probability p, and weak, $\underline{c}_2 < \overline{c}_2$, with probability $1 - p$;
- player 2 believes player 1 is strong, $\overline{c}_1 > 0$, with probability q, and weak, $\underline{c}_1 < \overline{c}_1$, with probability $1 - q$;
- beliefs are common knowledge

Assume strong type's costs are at least somewhat lower than the costs of his weak opponent: $\overline{c}_j > \sqrt{\underline{c}_i \overline{c}_i}$.

Common Contest Nash Equilibrium

Players optimize:

$$\max_{m_i} \left\{ \frac{m_i}{m_1 + m_2} - \frac{m_i}{c_i} \right\},$$

for which the interior solution is:

$$m_1^* = c_2 \left(\frac{c_1}{c_1 + c_2} \right)^2 \quad \text{and} \quad m_2^* = c_1 \left(\frac{c_2}{c_1 + c_2} \right)^2.$$

Equilibrium expected payoffs:

$$W_1 = \left(\frac{c_1}{c_1 + c_2} \right)^2 \quad \text{and} \quad W_2 = \left(\frac{c_2}{c_1 + c_2} \right)^2.$$

War Still Inefficient

Basic setup from Fearon's model the same:

$$W_1 + W_2 < 1,$$

so war is inefficient, so mutually acceptable peaceful division still exists under complete information.

The Informed Player

The informed player (1) optimizes:

$$\max_{m_1} \left\{ \frac{m_1}{m_1 + m_2} - \frac{m_1}{c_1} \right\},$$

for which we already know the solution:

$$m_1(m_2; c_1) = \max\left(\sqrt{c_1 m_2} - m_2, 0\right).$$

This is enough for the following Lemma 1:

Lemma

In equilibrium, either both types of the informed player participate in the contest (skirmish), or only the strong type does (war).

The Uninformed Player: Skirmish Equilibrium

Let $m_1(\overline{c}_1) > 0$ and $m_1(\underline{c}_1) > 0$ be player 1's type-contingent effort levels. Player 2 has a (posterior) belief $\hat{q}$ and optimizes:

$$\max_{m_2} \left\{ \frac{\hat{q} m_2}{m_1(\overline{c}_1) + m_2} + \frac{(1-\hat{q}) m_2}{m_1(\underline{c}_1) + m_2} - \frac{m_2}{c_2} \right\},$$

for which the solution is:

$$m_2^* = \underline{c}_1 \overline{c}_1 \left[\frac{f(\hat{q})}{g(\hat{q}; c_2)} \right]^2,$$

where $f(\hat{q}) = \hat{q}\sqrt{\underline{c}_1} + (1-\hat{q})\sqrt{\overline{c}_1} > 0$ and $g(\hat{q}; c_2) = \frac{\underline{c}_1 \overline{c}_1}{c_2} + \hat{q}\underline{c}_1 + (1-\hat{q})\overline{c}_1 > 0$.

The Uninformed Player: War Equilibrium

Since $m_1(\underline{c}_1) = 0$, player 2 optimizes:

$$\max_{m_2} \left\{ \frac{\hat{q} m_2}{m_1(\overline{c}_1) + m_2} + (1 - \hat{q}) - \frac{m_2}{c_2} \right\},$$

for which the solution is:

$$m_2^* = \overline{c}_1 \left(\frac{\hat{q} c_2}{\overline{c}_1 + \hat{q} c_2} \right)^2 .$$

Sun Tzu's Principle of Feigning Weakness

Lemma

The more confident player 2 gets that player 1 is strong, the more effort will she spend fighting him.

Lemma (Sun Tzu)

Player 1's expected payoff from fighting decreases as player 2 gets more confident that he is strong.

(If player 2 thinks player 1 is likely to be weak, she expends less effort than she would have if she knew player 1 was strong. The strong player 1 benefits.)

Feint Equilibrium Structure

I construct a class of equilibria with following structure:

- player 1 makes either a low-value, low-risk demand, $\underline{x}$, or a high-value high-risk demand $\overline{x} > \underline{x}$ as follows:
  - weak type always demands $\underline{x}$
  - strong type mixes between $\underline{x}$ and $\overline{x}$
- weak player 2 accepts both demands
- strong player 2 accepts $\underline{x}$ with positive probability and rejects $\overline{x}$ with certainty

In this equilibrium, the strong player 1 pretends to be weak with positive probability during the crisis.

The Equilibrium Demands

Since only the strong player 2 rejects an equilibrium offer with positive probability, in the contest player 1 knows he faces the strong opponent and:

- after $\underline{x}$: player 2 is unsure if player 1 is strong (one-sided asymmetric info)
  - strong player 2's contest payoff: $W_2(\hat{q}; \overline{c}_2)$
  - since she's willing to mix: $1 - \underline{x} = W_2(\hat{q}; \overline{c}_2)$, or

$$\underline{x} = 1 - W_2(\hat{q}(\underline{x}); \overline{c}_2)$$

- after $\overline{x}$: player 2 knows player 1 is strong (complete info)
  - weak player 2 accepts but strong player 2 rejects
  - if weak player 2 deviates and fights, she gets $W_2' < W_2(\overline{c}_1, \underline{c}_2)$
  - player 1 offers enough to get weak to accept: $1 - \overline{x} = W_2'$, or

$$\overline{x} = 1 - W_2'$$

Range of Possible Low-Value Demands

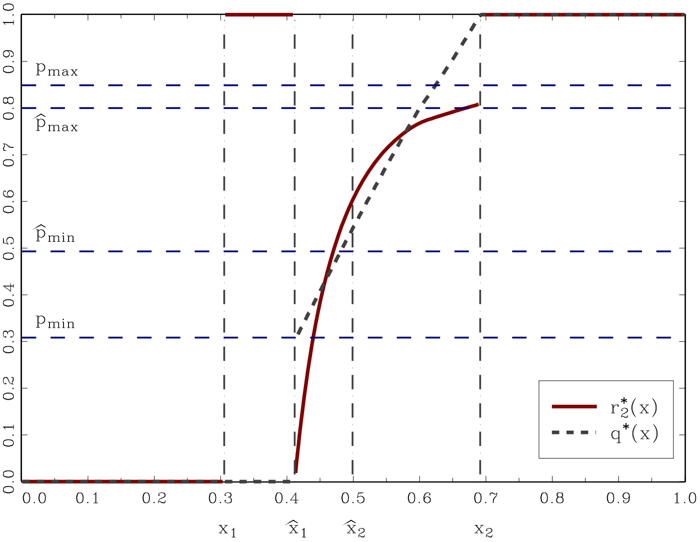
Since $\overline{W}_2 = W_2(\underline{c}_1, \overline{c}_2)$ is the *best* possible contest payoff for the strong player 2 and $\underline{W}_2 = W_2(\overline{c}_1, \overline{c}_2)$ is the *worst*, she rejects any $1 - x < \underline{W}_2$ and accepts any $1 - x > \overline{W}_2$ regardless of her beliefs (Lemma 4).

The only belief-contingent responses are to $x \in [x_1, x_2]$ where

$$x_1 = 1 - \overline{W}_2 < 1 - \underline{W}_2 = x_2$$

By Lemma 5, for any such x, there exists a unique $\hat{q} \in [0, 1]$ that satisfies $x = 1 - W_2(\hat{q}(x); \overline{c}_2)$, the requirement for $\underline{x}$. Moreover, this $\hat{q}$ is strictly increasing in x.

Posterior Beliefs for Player 2



Strong Player 2's Probability of Rejecting Demands

Let $r_2(x)$ be probability of strong player 2 rejecting $x \in [x_1, x_2]$.
The strong player 1's payoff from the low-value demand:

$$U_1(x; \overline{c}_1) = pr_2(x) W_1(\hat{q}(x); \overline{c}_1) + (1 - pr_2(x))x,$$

The strong player 1's payoff from the high-value demand $\overline{x}$:

$$U_1(\overline{x}; \overline{c}_1) = pW_1(\overline{c}_1, \overline{c}_2) + (1 - p)\overline{x} \equiv \hat{x}_1$$

(recall that $\overline{x}$ is fixed by the exogenous parameters)

Strong Player 2's Probability of Rejecting Demands

- If strong player 2 were to accept $x \in [x_1, \hat{x}_1]$ for sure, the strong player 1 will never make such low demands.

  $\Rightarrow$ posterior belief $\hat{q}(x) = 0$

  $\Rightarrow$ strong player 2 certain to reject such x

  $\Rightarrow$ strong player 1 tempted to deviate to x (player 2 erroneously believes he's weak)

  $\Rightarrow$ posterior belief cannot be $\hat{q}(x) = 0$!

Only if the strong player 1 does not want to deviate despite temptation, can this belief be rationalized. (Restrictions on player 1's prior.)

The Feint Probability

- Let $r_1(\underline{x})$ the probability with which the strong player 1 makes the low-value demand (feigns weakness).
- Since it must induce $q^*(\underline{x})$, Bayes rule requires:

$$r_1^*(\underline{x}) = \frac{q^*(\underline{x})(1-q)}{q(1-q^*(\underline{x}))},$$

which is a valid probability if $q^*(\underline{x}) < q$ ($p > \hat{p}_{\min}$).

- To ensure that the weak player 1 cannot profit by deviating to x_2, we find $\hat{x}_2 \in [\hat{x}_1, x_2]$ such that $U_1(x; \underline{c}_1) \geq U_1(x_2; \underline{c}_1)$.
- The set $[\hat{x}_1, \hat{x}_2]$ exists (Lemma 8).

The Feint Equilibria

Theorem

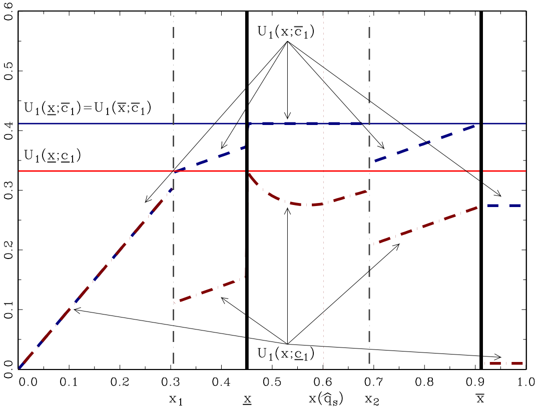
When the necessary conditions are satisfied, then any $\underline{x} \in [\hat{x}_1, \hat{x}_2]$ *can be supported in a perfect Bayesian equilibrium of the crisis bargaining game using the following strategies and beliefs:*

- *The weak player 1 demands* $\underline{x}$*. The strong player 1, demands* $\underline{x}$ *with probability* $r_1^*(\underline{x})$ *and* $\overline{x}$ *with probability* $1 - r_1^*(\underline{x})$*.*
- *The weak player 2 accepts* $x \leq \overline{x}$*, and rejects every other demand. The strong player 2 accepts* $x \leq x_1$*, rejects* $x \in (x_1, \underline{x})$*, accepts* $x \in [\underline{x}, x_2]$ *with probability* $1 - r_2^*(x)$*, and rejects* $x > x_2$*.*

On and off the path, beliefs are given by $q^*(x)$*. In the contest, players use the belief-contingent equilibrium strategies.*

A Numerical Example

Let $\underline{c}_1 = 1$, $\overline{c}_1 = 4$, $\underline{c}_2 = 2$, $\overline{c}_2 = 5$, and $p = q = 0.7$, so $\overline{x} = 0.91$, and $\underline{x} \in [0.41, 0.50]$. Take $\underline{x} = 0.45$, so $r_1^* = 0.30$, $r_2^* = 0.38$, and $q^* = 0.41$.



Feint benefit: $\underline{x}$ ($\overline{x}$) is rejected with probability 0.27 (0.70), but the strong player 1's war payoff is 0.31 (0.20). Pr(feint) $= qr_1^* \approx 0.21$.

Costly Signaling and Feints

The costly signaling logic *remains* in the feint equilibria:

- the strong player 1 can only get $\overline{x}$ by running a *larger* risk of a *costlier* war.
- this discourages the weak from demanding $\overline{x}$ as well.

$\Rightarrow$ Credible revelation of information requires costly signal.

However, with endogenous belief-contingent war payoffs:

- opponent's beliefs matter in war
- incentives to manipulate these beliefs when one is strong
- strong reason not to reveal strength

$\Rightarrow$ strong player may foster *false optimism*

Overcoming Mutual Optimism

- disagreement about how war will "play out"
- credible signaling: imperfect cure for mutual optimism
- but. . . costly signaling not the only problem
- A's optimism may be deliberately induced by B

  $\Rightarrow$ B cannot use lack of credible signal as evidence that A is necessarily weak

  $\Rightarrow$ information remains private because the only type of A with incentives to reveal it, does not want to

  $\Rightarrow$ A cannot use B's costly signal to correct his own optimism because B's signal may be a product of B's false optimism that he cultivated

  $\Rightarrow$ war provides the "stinging ice of reality" (Blainey)