The Spotlight’s Harsh Glare: Rethinking Publicity and International Order

Allison Carnegie and Austin Carson

November 13, 2014
Question: How does publicity of rule violations affect international political order?

- Conventional wisdom: publicity reinforces rules and social order
- Rationale for “naming-and-shaming” and transparency function of IOs
- Yet studies in sociology and criminology suggest a different relationship
- Revealing information about drug use, crime, and other deviant behaviors can induce these behaviors

Puzzle: When does publicity of rule violations strengthen norms/regimes, as suggested by the conventional wisdom, and when might it undermine them?
A theory of “ignorance as bliss”

Exposure of rule violations can:

- Correct beneficial misperceptions about the overall rate of compliance (pessimism mechanism)
- Sharpen the threat posed by acts of deviance (threat mechanism)

Both mechanisms encourage additional violations by other states
Greater publicity for rule violations can endanger norms/regimes under certain conditions

- Threat environment
- Initial rule violation unlikely to be reversed

Observable implication: When these conditions are present, regime advocates should intentionally conceal and overlook rule violations ("strategically obfuscate")
We formalize this and generate hypotheses about when regime advocates should strategically obfuscate.

We test these hypotheses empirically by analyzing all relevant cases of violations of the nuclear non-proliferation norm.

We find that enforcers strategically obfuscate when conditions favor the order-degrading mechanisms we identify.

We find that primary evidence concerns about pessimism and threat dynamics drive such decisions.

We also find that enforcers embrace publicity consistent with conventional wisdom under reversed conditions.
Why Nuclear Non-Proliferation?

- Ideal area to test our claims:
  - Features a powerful state—the US (primarily)—with strong incentives to enforce the regime
  - Prominent and substantively important norm
  - Variation in our key independent variables over time and across regions
  - Features unusually candid archival material

- However, logic applies to many international areas: human rights, environment, trade, etc.
Model: Set-up and Utility

- Three states
  - State $A$, initial rule violator
  - State $B$, reacting states
  - State $E$, regime "enforcer" with knowledge of $A$'s behavior

- $E$ and $B$ are the strategic actors

- $B$’s utility: rate of compliance with the norm (more utility when more other states comply) and self defense benefits from its own nuclear weapons capability (more utility if $A$ has weapons) plus other domestic and international factors

- $E$’s utility: rate of compliance, other domestic and international factors
Timing

- A chooses whether to violate the norm with probability $q$
  - $B$ does not observe this decision but $E$ does
- $E$ decides whether to publicize the violation
- $B$ forms beliefs about whether $A$ violated the norm, determined by Bayes’ Rule
- If $E$ publicizes, $A$ may or may not come into compliance. The proportion of $A$s that comply is given by $r$
- $B$ decides whether to violate the norm
Results: \( E \) publicizes \( A \)'s rule violation if...

Publicity is rational if \( A \) is expected to comply and/or \( A \)'s rule violation creates few security risks for \( B \):

- If \( A \) will comply, \( E \) publicizes expecting such pressure to reverse \( A \)'s violating behavior.
- \( B \) is made certain of \( A \)'s violation but reversal avoids pessimism, eliminates danger.
- If \( A \) won’t comply but \( A \)'s rule violation poses few dangers to \( B \), \( E \) still publicizes.
Strategic obfuscation is rational if $A$ is not expected to reverse its activity and it poses security risks to $B$

- $E$ anticipates publicity will not help reverse $A$’s behavior
- $B$ remains ignorant or uncertain about $A$’s behavior
- $E$ strategically obfuscates to reduce the risk of ”$Bs$” violating due to pessimism and threat
Table: Choice of Strategic Obfuscation versus Publicity

<table>
<thead>
<tr>
<th></th>
<th>A Will Comply</th>
<th>A Will Not Comply</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Risk of Reactive Violations</td>
<td>Publicize</td>
<td>Strategic Obfuscation</td>
</tr>
<tr>
<td>Low Risk of Reactive Violations</td>
<td>Publicize</td>
<td>Publicize</td>
</tr>
</tbody>
</table>
Empirical Strategy: Case Selection

Examine all cases (13) in which:

- States had credible nuclear weapons program after the emergence of the non-proliferation norm
- American intelligence concluded such a program existed
- The proliferating state had not overtly demonstrated a capability

Take advantage of recently declassified U.S. archival material
Variables Coded

- Dependent variable: U.S. decision to hide or publicize a given state’s nuclear weapons program
- Independent variable 1: Likelihood the proliferating state (A) would reverse its program following publicity
- Independent variable 2: Risks of second-order proliferation
  - Likelihood other states will be threatened by nuclear program and build weapons in self-defense
  - Vulnerability of overall regime to pessimism (first, second, or third decade of NPT)
Table: Choice of Strategic Obfuscation versus Publicity

<table>
<thead>
<tr>
<th></th>
<th>A is Expected to Comply</th>
<th>A is Not Expected to Comply</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Risk of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reactive Violations</td>
<td>$E$ publicizes</td>
<td>$E$ obfuscates</td>
</tr>
<tr>
<td>Low Risk of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reactive Violations</td>
<td>$E$ publicizes</td>
<td>$E$ publicizes</td>
</tr>
</tbody>
</table>
### Table: Choice of Strategic Obfuscation versus Publicity

<table>
<thead>
<tr>
<th>High Risk of Reactive Violations</th>
<th>A is Expected to Comply</th>
<th>A is Not Expected to Comply</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Korea</td>
<td>Israel</td>
<td></td>
</tr>
<tr>
<td>Iraq After 1990</td>
<td>India</td>
<td></td>
</tr>
<tr>
<td>Taiwan</td>
<td>Pakistan</td>
<td></td>
</tr>
<tr>
<td>South Korea</td>
<td>Iraq Before 1990</td>
<td>South Africa</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low Risk of Reactive Violations</th>
<th>Algeria</th>
<th>Brazil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Libya</td>
<td>Argentina</td>
<td></td>
</tr>
</tbody>
</table>
US intel concludes nuclear weapon capability in late 1960s
US leaders skeptical publicity and pressure lead Israel to reverse course
US leaders acutely concerned about second-order proliferation
  - Early stages of NPT: signature, ratification
  - Arab states’ likely to react in self-defense
  - Kissinger: “Nuclear capable countries would be more likely to opt in favor of nuclear weapons for themselves and...would be less likely to sign the NPT”
As a result, US chose strategic obfuscation

Kissinger: “Public knowledge is almost as dangerous as possession itself”
   “Our primary concern was that the Israelis make no visible introduction of nuclear weapons”

Strategic obfuscation worked: no major leaks until 1980s, minimal damage to NPT
Conclusion

- Assumption: publicizing rule violations strengthens social and political order ("sunlight as disinfectant")
- Our theory: publicizing rule violations can *undermine* order
  - Prompts pessimism about overall compliance; exacerbates threat
  - Makes second-order rule violations more likely
- Prediction: regime enforcers *strategically obfuscate* if conditions favor these mechanisms
- Payoffs
  - Revisit core assumption in theories of norms, international institutions
  - Identifies and sheds light on puzzling cases of obfuscation
  - Consequence of a new era of leaks?