The Politics of Punishment:
Why Dictators Join the International Criminal Court

Leslie Johns[†] and Francesca Parente[‡]

September 2022

Online Appendix: Theory

Contents

[†]Professor, UCLA

[‡]Assistant Professor, Christopher Newport University

1 Theory

1.1 Main Model

1.1.1 Assumptions

We focus on strategic interactions between two actors: the government, G, and a rebel group, R. We assume that these two actors are competing for power over the state, which yields a payoff of $W > 0$. To isolate our causal mechanism, we assume that both players have complete information all aspects of the game.

The game begins when the government decides whether to join the ICC. We assume that the only impact of this publicly-observed decision is to raise the cost of violence to both players. After the government makes its decision, each player simultaneously chooses a level of effort to invest in violence, $e_i \geq 0$. These effort levels in turn determine the likelihood that each player wins power over the state. Namely, we assume that the probability that the government survives in power is:

$$\pi\left(e_G, e_R\right) = \frac{e_G}{e_G + e_R}$$

The probability that the rebel group gains power over the state is therefore: $1 - \pi$.

The payoffs of the players from the various outcomes depend on whether the government previously joined the ICC. If the government has not previously joined the ICC, we assume that the government's unit cost of violence is $c > 0$. We interpret this parameter as the strength of the government, relative to the rebel group. We assume that when the government is relatively strong, it can easily deploy violence at little cost. The corresponding government cost of violence, c, is therefore small. In contrast, we assume that when the government is relatively weak, violence is more costly in economic and political terms. The corresponding government cost of violence, c, is therefore large. We normalize the cost of violence for the rebel group to be equal to 1.[1]

In contrast, if the government has previously joined the ICC, we assume that the costs of violence are higher for both the government and the rebel group. Namely, we let parameter $\rho_H > 0$ denote the added unit cost of violence to the rebel group, and $\rho_L > 0$ denote the added unit cost of violence to the government. Higher values of ρ therefore correspond to situations in which players expect for the ICC to have more ability and willingness to prosecute international crimes. We additionally assume that $\rho_L < \rho_H$.

[1] This normalizing assumption does not affect any of our results. It simply reduces the number of parameters in the model.

The expected payoffs for both players from the various outcomes of the game are shown in Table A1.

Table A1: Expected Utility from Outcomes in the Theoretical Model

Actor	Don't Join	Join
Government	$W\pi - ce_G$	$W(1-\pi) - (c+\rho_L)\, e_G$
Rebel group	$W\pi - e_R$	$W(1-\pi) - (1+\rho_H)\, e_R$

1.1.2 Results

<u>**Selection**</u>

Proposition 1. *In equilibrium, weaker governments (with high values of c) join the ICC, while stronger governments (with lower values of c) do not join.*

Proof of Proposition 1. To solve the model generally, let c_i denote the cost of effort to actor i, and $\pi_i = \frac{e_i}{e_i+e_j}$ denote the probability that actor i gains the payoff from winning political power. Then the generic utility function for player i is:

$$U_i\left(e_i, e_j\right) = W\pi_i - c_i e_i$$

Maximizing the utility function with respect to the choice variable e_i yields the best response function:

$$e_i\left(e_j\right) = \left(\frac{e_j W}{c_i}\right)^{\frac{1}{2}} - e_j$$

Taking the intersection of the best response functions yields the equilibrium effort level:

$$e_i^* = \frac{W c_j}{\left(c_i + c_j\right)^2}$$

For the two different subgames, this yields the following behavior and outcomes:

	Don't Join	Join
Government effort (e_G)	$\frac{W}{(1+c)^2}$	$\frac{W(1+\rho_H)}{(1+c+\rho_L+\rho_H)^2}$
Rebel effort (e_R)	$\frac{Wc}{(1+c)^2}$	$\frac{W(c+\rho_L)}{(1+c+\rho_L+\rho_H)^2}$
Government survival probability (π)	$\frac{1}{1+c}$	$\frac{1+\rho_H}{1+c+\rho_L+\rho_H}$

So the government's expected utility from its two choices is:

$$\begin{aligned} EU_G\left(\text{Don't Join}\right) &= W\left[\frac{1}{1+c}\right] - c\left[\frac{W}{(1+c)^2}\right] = \frac{W}{(1+c)^2} \\ EU_G\left(\text{Join}\right) &= W\left[\frac{1+\rho_H}{1+c+\rho_L+\rho_H}\right] - (c+\rho_L)\left[\frac{W\left(1+\rho_H\right)}{\left(1+c+\rho_L+\rho_H\right)^2}\right] = \frac{W(1+\rho_H)^2}{\left(1+c+\rho_L+\rho_H\right)^2} \end{aligned}$$

The government will thus be willing to join iff:

$$\Psi \equiv \left(1+\rho_H\right)^2\left(1+c\right)^2 - \left(1+c+\rho_L+\rho_H\right)^2 \geq 0$$

Note that:

$$\begin{aligned} \Psi\left(c=0\right) &= \left(1+\rho_H\right)^2 - \left(1+\rho_H+\rho_L\right)^2 < 0 \\ \frac{\partial \Psi}{\partial c} &= 2\left(1+\rho_H\right)^2\left(1+c\right) - 2\left(c+\rho_L+\rho_H+1\right) \\ &= 2\left[\rho_H - \rho_L + \rho_H^2 + c\left(2\rho_H+\rho_H^2\right)\right] > 0 \\ \lim_{c\to\infty}\Psi &= \lim_{c\to\infty}\left\{\left(1+2\rho_H+\rho_H^2\right)\left(1+c\right)^2 - \left[\left(1+c\right)^2 + 2\left(1+c\right)\left(\rho_L+\rho_H\right) + \left(\rho_L+\rho_H\right)^2\right]\right\} \\ &= \lim_{c\to\infty}\left\{\left(2\rho_H+\rho_H^2\right)\left(1+c\right)^2 - 2\left(1+c\right)\left(\rho_L+\rho_H\right) - \left(\rho_L+\rho_H\right)^2\right\} \\ &= \lim_{c\to\infty}\left(1+c\right)\times\lim_{c\to\infty}\left[\left(2\rho_H+\rho_H^2\right)\left(1+c\right) - 2\left(\rho_L+\rho_H\right)\right] - \left(\rho_L+\rho_H\right)^2 = \infty \end{aligned}$$

So by the intermediate value theorem, there exists a unique $\bar{c} > 0$ such that $\Psi(c) < 0$ for all $c < \bar{c}$ and $\Psi(c) > 0$ for all $c > \bar{c}$. So higher values of c are (weakly) more likely to join than lower values of c. □

Violence

Proposition 2. *Joining the ICC always lowers the total violence in the state.*

Proof of Proposition 2. Total violence if the government does not join the ICC is:

$$\frac{W}{(1+c)^2} + \frac{Wc}{(1+c)^2} = \frac{W}{1+c}$$

Total violence if the government joins the ICC is:

$$\frac{W(1+\rho_H)}{(1+c+\rho_L+\rho_H)^2} + \frac{W(c+\rho_L)}{(1+c+\rho_L+\rho_H)^2} = \frac{W}{1+c+\rho_L+\rho_H}$$

Note that the latter quantity is always less than the former quantity. □

<u>**Survival**</u>

Proposition 3. *For governments that select into the ICC, joining the ICC increases the probability of surviving in office.*

Proof of Proposition 3. The probability that the government survives in office if it joins increases (relative to not joining) iff:

$$\frac{1}{1+c} < \frac{1+\rho_H}{1+c+\rho_L+\rho_H} \quad \Leftrightarrow \quad 1+c+\rho_L+\rho_H < (1+\rho_H)(1+c) \quad \Leftrightarrow \quad \frac{\rho_L}{\rho_H} < c$$

Note that:

$$\begin{aligned} \Psi\left(c = \frac{\rho_L}{\rho_H}\right) &= (1+\rho_H)^2\left(1+\frac{\rho_L}{\rho_H}\right)^2 - \left(1+\frac{\rho_L}{\rho_H}+\rho_L+\rho_H\right)^2 \\ &= \frac{(1+\rho_H)^2(\rho_L+\rho_H)^2}{\rho_H^2} - \left[\frac{(1+\rho_H)(\rho_L+\rho_H)}{\rho_H}\right]^2 = 0 \end{aligned}$$

So for governments that select into the ICC ($c > \bar{c}$), joining the ICC increases the probability of surviving in office. □

1.2 Model Extension: Punish the Winner Less than the Loser

1.2.1 Assumptions

Continue to assume the same players, actions, and preferences if the government does not join the Rome Statute

However, assume that if the government joins the Rome Statute, then the ICC costs are imposed after the conflict ends. Namely, suppose that the group that hold power pays a low unit cost, ρ_L, while the group that does not hold power pays a high unit cost, ρ_H, where $0 < \rho_L < \rho_H$.

Then expected utility for the subgame in which the government signs the Rome Statute is:

$$\begin{aligned} U_i(e) &= W\pi_i - c_i e_i - \rho_L e_i \pi_i - \rho_H e_i (1-\pi_i) \\ &= W\left(\frac{e_i}{e_i+e_j}\right) - c_i e_i - \rho_L \left(\frac{e_i^2}{e_i+e_j}\right) - \rho_H \left(\frac{e_i e_j}{e_i+e_j}\right) \end{aligned}$$

1.2.2 Results

Selection

Proposition 4. *In equilibrium, weaker governments (with high values of c) join the ICC, while stronger governments (with lower values of c) do not join.*

Proof of Proposition 4. If the government does not join the Rome Statute, then subgame behavior matches that in Proposition 1.

However, if the government joins the Rome Statute, then the following optimization process applies. We begin with the first- and second-order conditions:

$$\begin{aligned} \frac{\partial U_i(e)}{\partial e_i} &= W\left[\frac{e_j}{(e_i+e_j)^2}\right] - c_i - \rho_L \left[\frac{e_i^2 + 2e_i e_j}{(e_i+e_j)^2}\right] - \rho_H \left[\frac{e_j^2}{(e_i+e_j)^2}\right] \\ \frac{\partial^2 U_i(e)}{\partial e_i^2} &= \frac{-2We_j}{(e_i+e_j)^3} - \rho_L \left[\frac{(e_i+e_j)^2 (2e_i + 2e_j) - \left(e_i^2 + 2e_i e_j\right) 2(e_i+e_j)}{(e_i+e_j)^4}\right] + \frac{2\rho_H e_j^2}{(e_i+e_j)^3} \\ &= \frac{2e_j\left[(\rho_H - \rho_L) e_j - W\right]}{(e_i+e_j)^3} \end{aligned}$$

Note that:

$$\begin{aligned} \frac{\partial U_i(e)}{\partial e_i}(e_i = 0) &= W\left(\frac{1}{e_j}\right) - c_i - \rho_H \\ &\Rightarrow \quad \frac{\partial U_i(e)}{\partial e_i}(e_i = 0 | e_j > 0) > 0 \quad \Leftrightarrow \quad \frac{W}{c_i + \rho_H} > e_j \\ \frac{\partial^2 U_i(e)}{\partial e_i^2} < 0 &\Leftrightarrow e_j \in \left(0, \ \frac{W}{\rho_H - \rho_L}\right) \end{aligned}$$

So the binding constraint to identify an optimal e_i is:

$$e_j \in \left(0, \ \frac{W}{c_i + \rho_H}\right)$$

This translates to the system of constraints:

$$e_G \in \left(0,\ \frac{W}{1+\rho_H}\right)$$
$$e_R \in \left(0,\ \frac{W}{c+\rho_H}\right)$$

We can now identify the best response function for each player:

$$\frac{\partial U_i\left(e\right)}{\partial e_i} = 0 \Leftrightarrow We_j - c_i\left(e_i + e_j\right)^2 - \rho_L\left(e_i^2 + 2e_ie_j\right) - \rho_H e_j^2 = 0$$
$$\Leftrightarrow \left(c_i + \rho_L\right)e_i^2 + 2\left(c_i + \rho_L\right)e_ie_j + \left[\left(c_i - \rho_H\right)e_j^2 - We_j\right] = 0$$

$$\begin{aligned} e_i\left(e_j\right) &= \frac{-2\left(c_i+\rho_L\right)e_j \pm \sqrt{4\left(c_i+\rho_L\right)^2 e_j^2 - 4\left(c_i+\rho_L\right)\left[\left(c_i+\rho_H\right)e_j^2 - We_j\right]}}{2\left(c_i+\rho_L\right)} \\ &= \frac{\left(c_i+\rho_L\right)^{\frac{1}{2}}\left[We_j + \left(\rho_L - \rho_H\right)e_j^2\right]^{\frac{1}{2}} - \left(c_i+\rho_L\right)e_j}{\left(c_i+\rho_L\right)} \\ &= \left[\frac{We_j + \left(\rho_L-\rho_H\right)e_j^2}{\left(c_i+\rho_L\right)}\right]^{\frac{1}{2}} - e_j \end{aligned}$$

This translates to the best response functions:

$$\begin{aligned} e_G\left(e_R\right) &= \left(\frac{We_R + \left(\rho_L-\rho_H\right)e_R^2}{c+\rho_L}\right)^{\frac{1}{2}} - e_R \\ e_R\left(e_G\right) &= \left(\frac{We_G + \left(\rho_L-\rho_H\right)e_G^2}{1+\rho_L}\right)^{\frac{1}{2}} - e_G \end{aligned}$$

One possible solution to this system is: $e_G = e_R = 0$. However, these values do not maximize the utility function because of the second-order condition. Alternatively, substitution yields:

$$
\begin{aligned}
e_G &= \left(\frac{W\left[\left(\frac{We_G+(\rho_L-\rho_H)e_G^2}{1+\rho_L}\right)^{\frac{1}{2}}-e_G\right]+(\rho_L-\rho_H)\left[\left(\frac{We_G+(\rho_L-\rho_H)e_G^2}{1+\rho_L}\right)^{\frac{1}{2}}-e_G\right]^2}{c+\rho_L}\right)^{\frac{1}{2}} \\
&\quad -\left[\left(\frac{We_G+(\rho_L-\rho_H)\,e_G^2}{1+\rho_L}\right)^{\frac{1}{2}}-e_G\right] \\
\Leftrightarrow &\quad \frac{W\left[\left(\frac{We_G+(\rho_L-\rho_H)e_G^2}{1+\rho_L}\right)^{\frac{1}{2}}-e_G\right]+(\rho_L-\rho_H)\left[\left(\frac{We_G+(\rho_L-\rho_H)e_G^2}{1+\rho_L}\right)^{\frac{1}{2}}-e_G\right]^2}{c+\rho_L} \\
&\quad =\frac{We_G+(\rho_L-\rho_H)\,e_G^2}{1+\rho_L} \\
\Leftrightarrow &\quad 0=W\left[\left(\frac{W+(\rho_L-\rho_H)\,e_G}{1+\rho_L}\right)^{\frac{1}{2}}-e_G^{\frac{1}{2}}\right]+(\rho_L-\rho_H)\,e_G^{\frac{1}{2}}\left[\left(\frac{W+(\rho_L-\rho_H)\,e_G}{1+\rho_L}\right)^{\frac{1}{2}}-e_G^{\frac{1}{2}}\right]^2 \\
&\quad -(c+\rho_L)\,e_G^{\frac{1}{2}}\left(\frac{W+(\rho_L-\rho_H)\,e_G}{1+\rho_L}\right)
\end{aligned}
$$

Define

$$
\begin{aligned}
\Delta(x) &\equiv W\left[\Upsilon(x)^{\frac{1}{2}}-x^{\frac{1}{2}}\right]+(\rho_L-\rho_H)\,x^{\frac{1}{2}}\left[\Upsilon(x)^{\frac{1}{2}}-x^{\frac{1}{2}}\right]^2-(c+\rho_L)\,x^{\frac{1}{2}}\Upsilon(x) \\
&\text{where}\quad \Upsilon(x)\equiv\frac{W+(\rho_L-\rho_H)\,x}{1+\rho_L}
\end{aligned}
$$

Note the following property of the Δ function:

$$
\Delta(0)=W\Upsilon(0)^{\frac{1}{2}}=\frac{W^{\frac{3}{2}}}{(1+\rho_L)^{\frac{1}{2}}}>0
$$

Also note that:

$$
\Upsilon\left(\frac{W}{1+\rho_H}\right)^{\frac{1}{2}}-\left(\frac{W}{1+\rho_H}\right)^{\frac{1}{2}}=\left(\frac{W+(\rho_L-\rho_H)\left(\frac{W}{1+\rho_H}\right)}{1+\rho_L}\right)^{\frac{1}{2}}-\left(\frac{W}{1+\rho_H}\right)^{\frac{1}{2}}=0
$$

So:

$$
\Delta\left(\frac{W}{1+\rho_H}\right)=-(c+\rho_L)\left(\frac{W}{1+\rho_H}\right)^{\frac{1}{2}}\Upsilon\left(\frac{W}{1+\rho_H}\right)=-(c+\rho_L)\left(\frac{W}{1+\rho_H}\right)^{\frac{3}{2}}<0
$$

Finally, note that:

$$\Delta'(x) = W\left[\frac{\Upsilon'(x)}{2\Upsilon(x)^{\frac{1}{2}}} - \frac{1}{2x^{\frac{1}{2}}}\right] + (\rho_L - \rho_H)\left\{x^{\frac{1}{2}}2\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]\left[\frac{\Upsilon'(x)}{2\Upsilon(x)^{\frac{1}{2}}} - \frac{1}{2x^{\frac{1}{2}}}\right] + \frac{1}{2x^{\frac{1}{2}}}\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2\right\}$$
$$- (c+\rho_L)\left[x^{\frac{1}{2}}\Upsilon'(x) + \frac{\Upsilon(x)}{2x^{\frac{1}{2}}}\right]$$
$$= \frac{W}{2}\left[\frac{\left(\frac{\rho_L-\rho_H}{1+\rho_L}\right)}{\left(\frac{W+(\rho_L-\rho_H)x}{1+\rho_L}\right)^{\frac{1}{2}}} - \frac{1}{x^{\frac{1}{2}}}\right]$$
$$+ (\rho_L - \rho_H)\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]\left[\frac{\left(\frac{\rho_L-\rho_H}{1+\rho_L}\right)x^{\frac{1}{2}}}{\left(\frac{W+(\rho_L-\rho_H)x}{1+\rho_L}\right)^{\frac{1}{2}}} + \frac{\Upsilon(x)^{\frac{1}{2}}}{2x^{\frac{1}{2}}} - \frac{3}{2}\right]$$
$$- \left(\frac{c+\rho_L}{2x^{\frac{1}{2}}}\right)\left[2x\left(\frac{\rho_L-\rho_H}{1+\rho_L}\right) + \left(\frac{W+(\rho_L-\rho_H)x}{1+\rho_L}\right)\right]$$
$$= \frac{W}{2}\left[\frac{\rho_L-\rho_H}{[W+(\rho_L-\rho_H)x]^{\frac{1}{2}}(1+\rho_L)^{\frac{1}{2}}} - \frac{1}{x^{\frac{1}{2}}}\right]$$
$$+ (\rho_L - \rho_H)\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]\left[\frac{(\rho_L-\rho_H)x^{\frac{1}{2}}}{[W+(\rho_L-\rho_H)x]^{\frac{1}{2}}(1+\rho_L)^{\frac{1}{2}}} + \frac{\Upsilon(x)^{\frac{1}{2}}}{2x^{\frac{1}{2}}} - \frac{3}{2}\right]$$
$$- \left(\frac{c+\rho_L}{2x^{\frac{1}{2}}}\right)\left(\frac{W-3(\rho_H-\rho_L)x}{1+\rho_L}\right)$$

Note that $\lim_{W\to\infty}\Upsilon(x) = \infty$. So:

$$\lim_{W\to\infty}\Delta'(x) = -\frac{1}{2x^{\frac{1}{2}}}\lim_{W\to\infty}[W] - (\rho_H-\rho_L)\left[\lim_{W\to\infty}\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]\left[\lim_{W\to\infty}\frac{\Upsilon(x)^{\frac{1}{2}}}{2x^{\frac{1}{2}}} - \frac{3}{2}\right]$$
$$- \left(\frac{c+\rho_L}{2x^{\frac{1}{2}}}\right)\lim_{W\to\infty}\left(\frac{W-3(\rho_H-\rho_L)x}{1+\rho_L}\right) = -\infty < 0$$

By the IVT, for large W there exists a unique $x \in \left(0, \frac{W}{1+\rho_H}\right)$ that solves $\Delta(x) = 0$. This is the government's equilibrium level of effort.

Note that:

$$\frac{\partial x}{\partial c} = \frac{-\Delta_c}{\Delta_x} = \frac{x^{\frac{1}{2}}\Upsilon(x)}{\Delta_x} < 0 \quad \text{for large } W$$

The total equilibrium violence for this subgame is:

$$e_G^* + e_R^* = \left[\frac{Wx + (\rho_L+\rho_H)x^2}{(1+\rho_L)}\right]^{\frac{1}{2}}$$

The government probability of survival for this subgame is:

$$\pi_G = \frac{e_G}{e_G + e_R} = \frac{e_G}{\left[\frac{We_G + (\rho_L + \rho_H) e_G^2}{(1+\rho_L)}\right]^{\frac{1}{2}}} = \frac{e_G^{\frac{1}{2}} (1 + \rho_L)^{\frac{1}{2}}}{\left[W + (\rho_L + \rho_H) e_G\right]^{\frac{1}{2}}} = \left[\frac{x}{\Upsilon(x)}\right]^{\frac{1}{2}}$$

The government's expected utility from this subgame is:

$$\begin{aligned} U_G(x) &= W\pi_G - (c + \rho_H) x + (\rho_H - \rho_L) x \pi_G \\ &= W \left[\frac{x}{\Upsilon(x)}\right]^{\frac{1}{2}} + (\rho_H - \rho_L) \left[\frac{x^{\frac{3}{2}}}{\Upsilon(x)^{\frac{1}{2}}}\right] - (c + \rho_H) x \end{aligned}$$

Note that:

$$\begin{aligned} \Delta(x) &= W \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right] + (\rho_L + \rho_H) x^{\frac{1}{2}} \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2 - (c + \rho_L) x^{\frac{1}{2}} \Upsilon(x) = 0 \\ &\Leftrightarrow (c + \rho_L) x^{\frac{1}{2}} \Upsilon(x) = W \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right] + (\rho_L - \rho_H) x^{\frac{1}{2}} \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2 \\ &\Leftrightarrow c = W \left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{x^{\frac{1}{2}} \Upsilon(x)}\right] + (\rho_L - \rho_H) \frac{\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2}{\Upsilon(x)} - \rho_L \\ &\Leftrightarrow (c + \rho_H) x = W x^{\frac{1}{2}} \left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{\Upsilon(x)}\right] + (\rho_H - \rho_L) x \left[\frac{\Upsilon(x) - \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2}{\Upsilon(x)}\right] \end{aligned}$$

So the government's expected utility from this subgame–given equilibrium behavior–is equivalent to:

$$\begin{aligned} U_G(x) &= W \left[\frac{x^{\frac{1}{2}}}{\Upsilon(x)^{\frac{1}{2}}}\right] + (\rho_H - \rho_L) \left[\frac{x^{\frac{3}{2}}}{\Upsilon(x)^{\frac{1}{2}}}\right] \\ &\quad - \left\{ W x^{\frac{1}{2}} \left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{\Upsilon(x)}\right] + (\rho_H - \rho_L) x \left[\frac{\Upsilon(x) - \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2}{\Upsilon(x)}\right] \right\} \\ &= W \left[\frac{x}{\Upsilon(x)}\right] - (\rho_H - \rho_L) \left[\frac{x^{\frac{3}{2}} \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]}{\Upsilon(x)}\right] \end{aligned}$$

The government wants to join the ICC iff:

$$\theta(c) \equiv W \left[\frac{x}{\Upsilon(x)}\right] - (\rho_H - \rho_L) \left[\frac{x^{\frac{3}{2}} \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]}{\Upsilon(x)}\right] - \frac{W}{(1 + c)^2} > 0$$

Define $x_0 \equiv \lim_{c \to 0} x$. Recall that x_0 is finite because it must fall within the interval $\left(0, \frac{W}{1+\rho_H}\right)$.

So:

$$\theta(0) = W\left[\frac{x_0}{\Upsilon(x_0)}\right] - (\rho_H - \rho_L)\left[\frac{x_0^{\frac{3}{2}}\left[\Upsilon(x_0)^{\frac{1}{2}} - x_0^{\frac{1}{2}}\right]}{\Upsilon(x_0)}\right] - W < 0$$

$$\Leftrightarrow -(\rho_H - \rho_L)\, x_0^{\frac{3}{2}}\left[\Upsilon(x_0)^{\frac{1}{2}} - x_0^{\frac{1}{2}}\right] < W\left[\Upsilon(x_0) - x_0\right]$$

This always holds because $x < \Upsilon(x)$ and $x^{\frac{1}{2}} < \Upsilon(x)^{\frac{1}{2}}$ for all $x \in \left(0, \frac{W}{1+\rho_H}\right)$.

Next note that:

$$\begin{aligned}
\frac{d\theta(c)}{dc} &= \frac{\partial \theta(c)}{\partial c} + \frac{\partial \theta(c)}{\partial x}\left(\frac{\partial x}{\partial c}\right) \\
&= \frac{2W}{(1+c)^3} + \left(\frac{\partial x}{\partial c}\right)\frac{\partial}{\partial x}\left\{W\left[\frac{x}{\Upsilon(x)}\right] - (\rho_H - \rho_L)\left[\frac{x^{\frac{3}{2}}\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]}{\Upsilon(x)}\right]\right\} \\
&= \frac{2W}{(1+c)^3} + \frac{W^2 x^{\frac{1}{2}}}{(1+\rho_L)\,\Delta_x \Upsilon(x)} \\
&\quad - \left[\frac{(\rho_H - \rho_L)\,x}{2\Delta_x}\right]\left\{\frac{x\Upsilon'(x)}{\Upsilon(x)^{\frac{1}{2}}} - x^{\frac{1}{2}} + 3\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right] - 2x\left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{\Upsilon(x)}\right]\Upsilon'(x)\right\}
\end{aligned}$$

Note that:

$$\begin{aligned}
0 < \frac{d\theta(c)}{dc} \Leftrightarrow 0 &< \frac{2}{(1+c)^3} + \frac{W x^{\frac{1}{2}}}{(1+\rho_L)\,\Delta_x \Upsilon(x)} \\
&\quad - \left[\frac{(\rho_H - \rho_L)\,x^{\frac{3}{2}}}{2\Delta_x \Upsilon(x)}\right]\left\{\frac{x\Upsilon'(x)}{W\Upsilon(x)^{\frac{1}{2}}} - \frac{x^{\frac{1}{2}}}{W} + \frac{3\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]}{W} - 2x\left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{W\Upsilon(x)}\right]\Upsilon'(x)\right\} \equiv M
\end{aligned}$$

Recall that $\lim_{W \to \infty} \Delta_x = -\infty$. Define $\hat{x} \equiv \lim_{W \to \infty} x$. Recall that $\hat{x}$ is finite because it must fall within the interval $\left(0, \frac{W}{1+\rho_H}\right)$.

Consider the components of function M:

$$\lim_{W \to \infty}\left\{\frac{2}{(1+c)^3}\right\} = \frac{2}{(1+c)^3} > 0$$

$$\lim_{W\to\infty}\left\{\frac{Wx^{\frac{1}{2}}}{(1+\rho_L)\,\Delta_x\Upsilon(x)}\right\} = \frac{\hat{x}^{\frac{1}{2}}}{(1+\rho_L)}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x}\right\}\lim_{W\to\infty}\left\{\frac{W}{\Upsilon(x)}\right\}$$

$$= \frac{\hat{x}^{\frac{1}{2}}}{(1+\rho_L)}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x}\right\}\lim_{W\to\infty}\left\{\frac{1+\rho_L}{1+\frac{(\rho_L-\rho_H)\hat{x}}{W}}\right\} = \hat{x}^{\frac{1}{2}}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x}\right\} = 0$$

$$\lim_{W\to\infty}\left\{\frac{(\rho_H-\rho_L)\,x^{\frac{3}{2}}}{2\Delta_x\Upsilon(x)}\right\} = \frac{(\rho_H-\rho_L)\,\hat{x}^{\frac{3}{2}}}{2}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x\Upsilon(x)}\right\} = 0$$

$$\lim_{W\to\infty}\left\{\frac{x\Upsilon'(x)}{W\Upsilon(x)^{\frac{1}{2}}} - \frac{x^{\frac{1}{2}}}{W} + \frac{3\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]}{W} - 2x\left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{W\Upsilon(x)}\right]\Upsilon'(x)\right\}$$

$$= \hat{x}\Upsilon'(\hat{x})\lim_{W\to\infty}\left\{\frac{1}{W\Upsilon(\hat{x})^{\frac{1}{2}}}\right\} - \hat{x}^{\frac{1}{2}}\lim_{W\to\infty}\left\{\frac{1}{W}\right\} + 3\lim_{W\to\infty}\left\{\frac{\Upsilon(\hat{x})^{\frac{1}{2}} - \hat{x}^{\frac{1}{2}}}{W}\right\} - 2\hat{x}\Upsilon'(\hat{x})\lim_{W\to\infty}\left\{\frac{\Upsilon(\hat{x})^{\frac{1}{2}} - \hat{x}^{\frac{1}{2}}}{W\Upsilon(\hat{x})}\right\}$$

$$= 3\left[\lim_{W\to\infty}\left\{\frac{\Upsilon(\hat{x})^{\frac{1}{2}}}{W}\right\} - \hat{x}^{\frac{1}{2}}\lim_{W\to\infty}\left\{\frac{1}{W}\right\}\right] - 2\hat{x}\Upsilon'(\hat{x})\left[\lim_{W\to\infty}\left\{\frac{1}{W\Upsilon(\hat{x})^{\frac{1}{2}}}\right\} - \hat{x}^{\frac{1}{2}}\lim_{W\to\infty}\left\{\frac{1}{W\Upsilon(\hat{x})}\right\}\right]$$

$$= 3\lim_{W\to\infty}\left\{\frac{\left(\frac{W+(\rho_L-\rho_H)\hat{x}}{1+\rho_L}\right)^{\frac{1}{2}}}{W}\right\} = \frac{3}{(1+\rho_L)^{\frac{1}{2}}}\lim_{W\to\infty}\left\{\frac{\left[W+(\rho_L-\rho_H)\,\hat{x}\right]^{\frac{1}{2}}}{W}\right\}$$

$$= \frac{3}{(1+\rho_L)^{\frac{1}{2}}}\lim_{W\to\infty}\left\{\frac{1}{2\left[W+(\rho_L-\rho_H)\,\hat{x}\right]^{\frac{1}{2}}}\right\} = 0$$

So:

$$\lim_{W\to\infty} M = \frac{2}{(1+c)^3} > 0$$

Finally, note that:

$$0 < \theta(c) \quad \Leftrightarrow \quad 0 < Wx - (\rho_H-\rho_L)\,x^{\frac{3}{2}}\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right] - \frac{\Upsilon(x)\,W}{(1+c)^2} \equiv N(c)$$

Define $\bar{x} \equiv \lim_{c\to\infty} x(c)$. Note that $0 \le \bar{x}$ and $\bar{x}$ is finite. Then:

$$\lim_{c\to\infty} N(c) = W\bar{x} - (\rho_H-\rho_L)\,\bar{x}^{\frac{3}{2}}\left[\Upsilon(\bar{x})^{\frac{1}{2}} - \bar{x}^{\frac{1}{2}}\right]$$

So:

$$0 < \lim_{c \to \infty} N(c) \quad \Leftrightarrow \quad (\rho_H - \rho_L)\, \bar{x}^{\frac{1}{2}} \left[\Upsilon(\bar{x})^{\frac{1}{2}} - \bar{x}^{\frac{1}{2}} \right] < W$$

This holds for large W. So by the IVT, there is a unique crossing for large W at which $\theta(c) = 0$. The government joins iff c is sufficiently large. □

Violence

Proposition 5. *Joining the ICC always lowers the total violence in the state.*

Proof of Proposition 5. Total violence if the government does not join the ICC is:

$$\frac{W}{1+c}$$

Total violence if the government joins the ICC is:

$$\left[\frac{Wx + (\rho_L + \rho_H)\, x^2}{(1+\rho_L)} \right]^{\frac{1}{2}}$$

where x is implicitly defined by $\Delta(x) = 0$ from the Proof of Proposition 4. □