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

Leslie Johns[†] and Francesca Parente[‡]

December 2023

Forthcoming in *International Studies Quarterly*

Online Appendix

Contents

[†]Professor, UCLA

[‡]Assistant Professor, Christopher Newport University

1 Theory

1.1 Baseline Model

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 power over the state. Then the generic utility function for player i is:

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

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

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

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

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

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

	Don't Join	Join
Dictator effort (e_D^*)	$\frac{W}{(1+c)^2}$	$\frac{W(1+\kappa_H)}{(1+c+\kappa_L+\kappa_H)^2}$
Political opposition effort (e_P^*)	$\frac{Wc}{(1+c)^2}$	$\frac{W(c+\kappa_L)}{(1+c+\kappa_L+\kappa_H)^2}$
Dictator survival probability (π^*)	$\frac{1}{1+c}$	$\frac{1+\kappa_H}{1+c+\kappa_L+\kappa_H}$

So the dictator's expected utility from his two choices is:

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

The dictator will thus be willing to join iff:

$$\Psi \equiv (1+\kappa_H)^2(1+c)^2 - (1+c+\kappa_L+\kappa_H)^2 \geq 0$$

Note that:

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

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. □

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

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

Total violence if the dictator joins the ICC is:

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

Note that the latter quantity is always less than the former quantity because $0 < \kappa_L + \kappa_H$ always.
□

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

$$\begin{aligned}
\frac{1}{1+c} \le \frac{1+\kappa_H}{1+c+\kappa_L+\kappa_H} \quad &\Leftrightarrow \quad 1+c+\kappa_L+\kappa_H \le (1+\kappa_H)(1+c) \\
&\Leftrightarrow \quad (1+c+\kappa_L+\kappa_H)^2 \le (1+\kappa_H)^2(1+c)^2
\end{aligned}$$

Note that this is equivalent to the selection constraint $\Psi \ge 0$. So for dictators that select into the ICC ($c > \bar{c}$), joining the ICC increases the probability of surviving in office. □

1.2 Extension 1: No ICC Costs for Dictators

Suppose $\kappa_L = 0$. Then the expected utility to players for outcomes in the theoretical model becomes:

Actor	Don't Join	Join
Dictator	$W\pi - ce_D$	$W\pi - ce_D$
Political Opponent	$W(1-\pi) - e_P$	$W(1-\pi) - (1+\kappa_H)\, e_P$

Proposition A1

The utility maximization calculations from the Proof of Proposition 1 continue to hold, yielding the same best response functions. For the two different subgames, this yields the following equilibrium behavior and outcomes:

	Don't Join	Join
Dictator effort (e_D^*)	$\frac{W}{(1+c)^2}$	$\frac{W(1+\kappa_H)}{(1+c+\kappa_H)^2}$
Political opposition effort (e_P^*)	$\frac{Wc}{(1+c)^2}$	$\frac{Wc}{(1+c+\kappa_H)^2}$
Dictator survival probability (π^*)	$\frac{1}{1+c}$	$\frac{1+\kappa_H}{1+c+\kappa_H}$

And the dictator's expected utility from his two choices is:

$$EU_D\,(\text{Don't Join}) = \frac{W}{(1+c)^2}$$
$$EU_D\,(\text{Join}) = \frac{W(1+\kappa_H)^2}{(1+c+\kappa_H)^2}$$

The dictator will thus be willing to join iff:

$$F \equiv (1+\kappa_H)^2\,(1+c)^2 - (1+c+\kappa_H)^2 \geq 0$$

Note that:

$$\begin{aligned} F(c=0) &= (1+\kappa_H)^2 - (1+\kappa_H)^2 = 0 \\ \frac{\partial F}{\partial c} &= 2(1+\kappa_H)^2(1+c) - 2(1+c+\kappa_H) \\ &= 2\kappa_H\left[1+\kappa_H + c(2+\kappa_H)\right] > 0 \end{aligned}$$

So the dictator is indifferent when $c = 0$ and strictly prefers to join whenever $c > 0$. QED.

Proposition A2

Total violence if the dictator does not join the ICC is:

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

Total violence if the dictator joins the ICC is:

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

Note that the latter quantity is always less than the former quantity because $0 < \kappa_H$ always. So joining the ICC lowers total violence. QED.

Proposition A3

The probability that the dictator survives in office if he joins (weakly) increases (relative to not joining) iff:

$$\begin{aligned} \frac{1}{1+c} \leq \frac{1+\kappa_H}{1+c+\kappa_H} &\Leftrightarrow 1+c+\kappa_H \leq (1+\kappa_H)(1+c) \\ &\Leftrightarrow (1+c+\kappa_H)^2 \leq (1+\kappa_H)^2(1+c)^2 \end{aligned}$$

Note that this is equivalent to the selection constraint $F \geq 0$. So for dictators that select into the ICC ($c > F$), joining the ICC increases the probability of surviving in office. QED.

1.3 Extension 2: ICC Benefits for Dictators

Assume the basic domestic and international costs as in the main model. Let $\beta > 0$ denote an added benefit to the dictator from joining the ICC. Then the expected utility to players for outcomes in the theoretical model becomes:

Actor	Don't Join	Join
Dictator	$W\pi - ce_D$	$W\pi - ce_D + \beta$
Political Opponent	$W(1-\pi) - e_P$	$W(1-\pi) - (1+\kappa_H)\, e_P$

Proposition B1

The value of β does not affect the best response functions for the choice of effort for either player. So it does not affect the equilibrium effort or survival probabilities.

So the dictator's expected utility from his two choices is:

$$\begin{aligned} EU_D\,(\text{Don't Join}) &= \frac{W}{(1+c)^2} \\ EU_D\,(\text{Join}) &= \frac{W(1+\kappa_H)^2}{(1+c+\kappa_L+\kappa_H)^2} + \beta \end{aligned}$$

The dictator will thus be willing to join iff:

$$G \equiv (1+\kappa_H)^2\,(1+c)^2 - (1+c+\kappa_L+\kappa_H)^2 + \beta \geq 0$$

Note that:

$$\begin{aligned} \frac{\partial G}{\partial c} &= 2\,(1+\kappa_H)^2\,(1+c) - 2\,(c+\kappa_L+\kappa_H+1) \\ &= 2\left[\kappa_H - \kappa_L + \kappa_H^2 + c\left(2\kappa_H + \kappa_H^2\right)\right] > 0 \\ \lim_{c\to\infty} G &= \lim_{c\to\infty}\left\{\left(1+2\kappa_H+\kappa_H^2\right)(1+c)^2 - \left[(1+c)^2 + 2\,(1+c)\,(\kappa_L+\kappa_H) + (\kappa_L+\kappa_H)^2\right] + \beta\right\} \\ &= \lim_{c\to\infty}\left\{\left(2\kappa_H+\kappa_H^2\right)(1+c)^2 - 2\,(1+c)\,(\kappa_L+\kappa_H) - (\kappa_L+\kappa_H)^2 + \beta\right\} \\ &= \lim_{c\to\infty}(1+c) \times \lim_{c\to\infty}\left[\left(2\kappa_H+\kappa_H^2\right)(1+c) - 2\,(\kappa_L+\kappa_H)\right] - (\kappa_L+\kappa_H)^2 + \beta = \infty \end{aligned}$$

And:

$$G(c=0) = (1+\kappa_H)^2 - (1+\kappa_L+\kappa_H)^2 + \beta$$

So for a small $\beta > 0$, there exists a unique $\bar{c} > 0$ such that $\Psi(c) < 0$ for all $c < \bar{c}$ and $\Psi(c) > 0$ for all $c > \bar{c}$. And for large $\beta > 0$, the dictator always joins the ICC, regardless of the level of political competition (c). QED.

For the remaining results, we assume a small $\beta > 0$.

Proposition B2
The Proof of Proposition 2 still holds. QED.

Proposition B3 (Survival logic does not necessarily hold)
The probability that the dictator survives in office if he joins (weakly) increases (relative to not joining) iff:

$$\frac{1}{1+c} \leq \frac{1+\kappa_H}{1+c+\kappa_L+\kappa_H} \quad \Leftrightarrow \quad 1+c+\kappa_L+\kappa_H \leq (1+\kappa_H)(1+c)$$
$$\Leftrightarrow \quad (1+c+\kappa_L+\kappa_H)^2 \leq (1+\kappa_H)^2 (1+c)^2$$

Because the added benefit β causes some types of the dictator to select into the ICC under constraint G even through this decision does not increase their survival in office suggests that Proposition 3 does not necessarily hold. That is, the dictator trades off a decrease in survival for an increase in rents.

1.4 Extension 3: Pay ICC Costs After Power Transfers

Continue to assume the same players, actions, and preferences if the dictator does not join the Rome Statute. However, assume that if the dictator joins the Rome Statute, then the ICC costs are imposed after the conflict ends. Namely, suppose that the actor that holds power pays a low unit cost, κ_L, while the group that does not hold power pays a high unit cost, κ_H, where $0 < \kappa_L < \kappa_H$.

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

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

Proposition C1

If the dictator does not join the ICC, then subgame behavior matches that in Proposition 1. However, if the dictator joins the ICC, then the following optimization process applies. We begin with the first- and second-order conditions:

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

Note that:

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

So the binding constraint to identify an optimal e_i is:

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

This translates to the system of constraints:

$$e_D \in \left(0,\ \frac{W}{1 + \kappa_H}\right)$$
$$e_P \in \left(0,\ \frac{W}{c + \kappa_H}\right)$$

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

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

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

This translates to the best response functions:

$$\begin{aligned} e_D(e_P) &= \left(\frac{We_P + (\kappa_L - \kappa_H)\, e_P^2}{c + \kappa_L}\right)^{\frac{1}{2}} - e_P \\ e_P(e_D) &= \left(\frac{We_D + (\kappa_L - \kappa_H)\, e_D^2}{1 + \kappa_L}\right)^{\frac{1}{2}} - e_D \end{aligned}$$

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

$$
\begin{aligned}
e_D &= \left(\frac{W\left[\left(\frac{We_D+(\kappa_L-\kappa_H)e_D^2}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D\right] + (\kappa_L-\kappa_H)\left[\left(\frac{We_D+(\kappa_L-\kappa_H)e_D^2}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D\right]^2}{c+\kappa_L} \right)^{\frac{1}{2}} \\
&\quad - \left[\left(\frac{We_D+(\kappa_L-\kappa_H)e_D^2}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D\right] \\
\Leftrightarrow &\quad \frac{W\left[\left(\frac{We_D+(\kappa_L-\kappa_H)e_D^2}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D\right] + (\kappa_L-\kappa_H)\left[\left(\frac{We_D+(\kappa_L-\kappa_H)e_D^2}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D\right]^2}{c+\kappa_L} \\
&\quad = \frac{We_D+(\kappa_L-\kappa_H)e_D^2}{1+\kappa_L} \\
\Leftrightarrow &\quad 0 = W\left[\left(\frac{W+(\kappa_L-\kappa_H)e_D}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D^{\frac{1}{2}}\right] + (\kappa_L-\kappa_H)e_D^{\frac{1}{2}}\left[\left(\frac{W+(\kappa_L-\kappa_H)e_D}{1+\kappa_L}\right)^{\frac{1}{2}} - e_D^{\frac{1}{2}}\right]^2 \\
&\quad - (c+\kappa_L)e_D^{\frac{1}{2}}\left(\frac{W+(\kappa_L-\kappa_H)e_D}{1+\kappa_L}\right)
\end{aligned}
$$

Define

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

Note the following property of the Δ function:

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

Also note that:

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

So:

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

Finally, note that:

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

Because this equation is continuous, it is possible that multiple values of $x \in \left(0, \frac{W}{1+\kappa_H}\right)$ solve $\Delta(x) = 0$. Each of these values would represent a possible equilibrium value of e_D for the subgame in which the dictator joins the ICC.

To ensure that we have a unique equilibrium, we according assume that the value of W is large. Note that $\lim_{W\to\infty}\Upsilon(x) = \infty$. So:

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

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

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 when the dictator joins the ICC is then:

$$e_D^* + e_P^* \quad = \quad \left[\frac{Wx + (\kappa_L - \kappa_H)\,x^2}{(1+\kappa_L)}\right]^{\frac{1}{2}}$$

The dictator probability of survival when the dictator joins the ICC is then:

$$\pi = \frac{e_D}{e_D + e_P} \quad = \quad \frac{e_D}{\left[\frac{We_D + (\kappa_L - \kappa_H) e_D^2}{(1+\kappa_L)}\right]^{\frac{1}{2}}} = \frac{e_D^{\frac{1}{2}}\,(1+\kappa_L)^{\frac{1}{2}}}{[W + (\kappa_L - \kappa_H)\,e_D]^{\frac{1}{2}}} = \left[\frac{x}{\Upsilon(x)}\right]^{\frac{1}{2}}$$

And the dictator's expected utility from joining the ICC is then:

$$\begin{aligned} U_G(\text{join}|x) &= W\pi_G - (c+\kappa_H)\,x + (\kappa_H - \kappa_L)\,x\pi_G \\ &= W\left[\frac{x}{\Upsilon(x)}\right]^{\frac{1}{2}} + (\kappa_H - \kappa_L)\left[\frac{x^{\frac{3}{2}}}{\Upsilon(x)^{\frac{1}{2}}}\right] - (c+\kappa_H)\,x \end{aligned}$$

Note that by definition of this value of x:

$$\begin{aligned} \Delta(x) = W\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right] + (\kappa_L + \kappa_H)\,x^{\frac{1}{2}}\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2 - (c+\kappa_L)\,x^{\frac{1}{2}}\Upsilon(x) = 0 \\ \Leftrightarrow (c+\kappa_L)\,x^{\frac{1}{2}}\Upsilon(x) = W\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right] + (\kappa_L - \kappa_H)\,x^{\frac{1}{2}}\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2 \\ \Leftrightarrow c = W\left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{x^{\frac{1}{2}}\Upsilon(x)}\right] + (\kappa_L - \kappa_H)\,\frac{\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2}{\Upsilon(x)} - \kappa_L \\ \Leftrightarrow (c+\kappa_H)\,x = Wx^{\frac{1}{2}}\left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{\Upsilon(x)}\right] + (\kappa_H - \kappa_L)\,x\left[\frac{\Upsilon(x) - \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2}{\Upsilon(x)}\right] \end{aligned}$$

So the dictator's expected utility from joining the ICC (given subsequent equilibrium behavior in the subgame) is:

$$\begin{aligned} \Delta(x) &= W\left[\frac{x^{\frac{1}{2}}}{\Upsilon(x)^{\frac{1}{2}}}\right] + (\kappa_H - \kappa_L)\left[\frac{x^{\frac{3}{2}}}{\Upsilon(x)^{\frac{1}{2}}}\right] \\ &\quad - \left\{Wx^{\frac{1}{2}}\left[\frac{\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}}{\Upsilon(x)}\right] + (\kappa_H - \kappa_L)\,x\left[\frac{\Upsilon(x) - \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2}{\Upsilon(x)}\right]\right\} \\ &= W\left[\frac{x}{\Upsilon(x)}\right] - (\kappa_H - \kappa_L)\left[\frac{x^{\frac{3}{2}}\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]}{\Upsilon(x)}\right] \end{aligned}$$

We can now say that the dictator wants to join the ICC iff:

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

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

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

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

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

Next note that:

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

Note that:

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

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 $\left(0, \frac{W}{1+\kappa_H}\right)$.

Consider the components of function M:

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

$$\lim_{W\to\infty}\left\{\frac{Wx^{\frac{1}{2}}}{(1+\kappa_L)\,\Delta_x\Upsilon(x)}\right\}=\frac{\hat{x}^{\frac{1}{2}}}{(1+\kappa_L)}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x}\right\}\lim_{W\to\infty}\left\{\frac{W}{\Upsilon(x)}\right\}$$
$$=\frac{\hat{x}^{\frac{1}{2}}}{(1+\kappa_L)}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x}\right\}\lim_{W\to\infty}\left\{\frac{1+\kappa_L}{1+\frac{(\kappa_L-\kappa_H)\hat{x}}{W}}\right\}=\hat{x}^{\frac{1}{2}}\lim_{W\to\infty}\left\{\frac{1}{\Delta_x}\right\}=0$$

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

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

So:

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

Finally, note that:

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

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

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

So:

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

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

Proposition C2

Step 1: Recall from the Proof of Proposition 1 that the best response function if the leader does not join the ICC in this model extension is:

$$e_i(e_j|\text{don't join})=\left(\frac{We_j}{c_i}\right)^{\frac{1}{2}}-e_j$$

Recall from the Proof of Proposition C1 that the best response function if the leader joins the ICC in this model extension is:

$$e_i(e_j|\text{join})=\left(\frac{We_j-(\kappa_H-\kappa_L)\,e_j^2}{c_i+\kappa_L}\right)^{\frac{1}{2}}-e_j$$

Note that as κ_L and κ_H become arbitrarily small, continuity of the $e_i(e_j|\text{join})$-function ensures that the equilibrium values of $e_D(\text{join})$ and $e_P(\text{join})$ will approach the equilibrium values of $e_D(\text{don't join})$ and $e_P(\text{don't join})$. So as κ_L and κ_H become arbitrarily small, total violence when the dictator joins that ICC will approach the level of total violence when the dictator does not join the ICC. We can therefore think of violence when the dictator has not joined the ICC as the limiting case of a model with ICC jurisdiction in which $\kappa_L=\kappa_H=0$.

Step 2: By the Proof of Proposition C1, the total violence when the dictator joins the ICC in this model extension is:

$$V\left(\text{join}|x\right) \equiv \left(\frac{Wx-\left(\kappa_H-\kappa_L\right)x^2}{1+\kappa_L}\right)^{\frac{1}{2}} = \left[x\Upsilon\left(x\right)\right]^{\frac{1}{2}} \quad \text{and} \quad \Upsilon\left(x\right) \equiv \frac{W-\left(\kappa_H-\kappa_L\right)x}{1+\kappa_L}$$

where x is the equilibrium value defined by:

$$\Delta\left(x\right) \equiv W\left[\Upsilon\left(x\right)^{\frac{1}{2}}-x^{\frac{1}{2}}\right]+\left(\kappa_L-\kappa_H\right)x^{\frac{1}{2}}\left[\Upsilon\left(x\right)^{\frac{1}{2}}-x^{\frac{1}{2}}\right]^2-\left(c+\kappa_L\right)\Upsilon\left(x\right)=0$$

To calculate the impact of changes in κ_L and κ_H on total violence we must consider the total derivatives:

$$\begin{aligned}\frac{dV\left(\text{join}|x\right)}{d\kappa_L} &= \frac{\partial V\left(\text{join}|x\right)}{\partial\kappa_L}+\frac{\partial V\left(\text{join}|x\right)}{\partial x}\left(\frac{\partial x}{\partial\kappa_L}\right)\\ \frac{dV\left(\text{join}|x\right)}{d\kappa_H} &= \frac{\partial V\left(\text{join}|x\right)}{\partial\kappa_H}+\frac{\partial V\left(\text{join}|x\right)}{\partial x}\left(\frac{\partial x}{\partial\kappa_H}\right)\end{aligned}$$

where:

$$\begin{aligned}\frac{\partial V\left(\text{join}|x\right)}{\partial\kappa_L} &= \frac{-1}{2\left[x\Upsilon\left(x\right)\right]^{\frac{1}{2}}}\left[\frac{Wx+\left(1+\kappa_H\right)x^2}{\left(1+\kappa_L\right)^2}\right]\\ \frac{\partial V\left(\text{join}|x\right)}{\partial\kappa_H} &= \frac{-1}{2\left[x\Upsilon\left(x\right)\right]^{\frac{1}{2}}}\left(\frac{x^2}{1+\kappa_L}\right)\\ \frac{\partial V\left(\text{join}|x\right)}{\partial x} &= \frac{1}{2\left[x\Upsilon\left(x\right)\right]^{\frac{1}{2}}}\left[\frac{W-2\left(\kappa_H-\kappa_L\right)x}{1+\kappa_L}\right]\end{aligned}$$

(a) Changes in κ_L

(i) Note that:

$$\frac{dV\left(\text{join}|x\right)}{d\kappa_L} = \frac{-1}{2\left[x\Upsilon\left(x\right)\right]^{\frac{1}{2}}}\left\{\frac{Wx+\left(1+\kappa_H\right)x^2}{\left(1+\kappa_L\right)^2}+\left[\frac{W-2\left(\kappa_H-\kappa_L\right)x}{1+\kappa_L}\right]\left(\frac{\Delta_{\kappa_L}}{\Delta_x}\right)\right\}$$

So:

$$\frac{dV\left(\text{join}|x\right)}{d\kappa_L}<0 \qquad \Leftrightarrow \qquad 0<Wx+\left(1+\kappa_H\right)x^2+\left(1+\kappa_L\right)\left[W-2\left(\kappa_H-\kappa_L\right)x\right]\left(\frac{\Delta_{\kappa_L}}{\Delta_x}\right)$$

Recall that the constraints on equilibrium solutions of x ensure that $\lim_{W\to\infty}x$ is finite and weakly positive. Also recall from the Proof of Proposition C1 that: $\lim_{W\to\infty}\Delta_x=-\infty$. So a sufficient condition for our result to hold when W is large is: $\lim_{W\to\infty}\Delta_{\kappa_L}-\infty$.

(ii) Now note that:

$$\begin{aligned}
\Delta_{\kappa_L} &= W\left[\frac{1}{2}\Upsilon(x)^{-\frac{1}{2}}\left(\frac{\partial\Upsilon}{\partial\kappa_L}\right)\right] + x^{\frac{1}{2}}\left\{(\kappa_L-\kappa_H)\,2\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]\frac{1}{2}\Upsilon(x)^{-\frac{1}{2}}\left(\frac{\partial\Upsilon}{\partial\kappa_L}\right) + \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2\right\} \\
&\quad - \left[(c+\kappa_L)\left(\frac{\partial\Upsilon}{\partial\kappa_L}\right) + \Upsilon(x)\right] \\
&= \left[\frac{W - 2x(\kappa_L-\kappa_H)}{2\Upsilon(x)^{\frac{1}{2}}} + x^{\frac{1}{2}}(\kappa_L-\kappa_H) - (c+\kappa_L)\right]\left(\frac{\partial\Upsilon}{\partial\kappa_L}\right) + x^{\frac{1}{2}}\left[\Upsilon(x) - 2\Upsilon(x)^{\frac{1}{2}}x^{\frac{1}{2}} + x\right] - \Upsilon(x) \\
&= \left[\frac{W - 2x(\kappa_L-\kappa_H)}{2\Upsilon(x)^{\frac{1}{2}}} + x^{\frac{1}{2}}(\kappa_L-\kappa_H) - (c+\kappa_L)\right]\left(\frac{\partial\Upsilon}{\partial\kappa_L}\right) + \left(x^{\frac{1}{2}}-1\right)\Upsilon(x) - 2x\Upsilon(x)^{\frac{1}{2}} + x^{\frac{3}{2}} \\
&= -\left[\frac{(1+\kappa_L)^{\frac{1}{2}}\left[W^{\frac{1}{2}} - \frac{2x(\kappa_L-\kappa_H)}{W^{\frac{1}{2}}}\right]}{2\left(1-\frac{(\kappa_H-\kappa_L)x}{W}\right)^{\frac{1}{2}}} + x^{\frac{1}{2}}(\kappa_L-\kappa_H) - (c+\kappa_L)\right]\left(\frac{W-(1+\kappa_H)x}{(1+\kappa_L)^2}\right) \\
&\quad + \left(x^{\frac{1}{2}}-1\right)\left(\frac{W-(\kappa_H-\kappa_L)x}{1+\kappa_L}\right) - 2x\left(\frac{W-(\kappa_H-\kappa_L)x}{1+\kappa_L}\right)^{\frac{1}{2}} + x^{\frac{3}{2}} \\
&= -W\left\{\left[\frac{(1+\kappa_L)^{\frac{1}{2}}\left[W^{\frac{1}{2}} - \frac{2x(\kappa_L-\kappa_H)}{W^{\frac{1}{2}}}\right]}{2\left(1-\frac{(\kappa_H-\kappa_L)x}{W}\right)^{\frac{1}{2}}} + x^{\frac{1}{2}}(\kappa_L-\kappa_H) - (c+\kappa_L)\right]\left(\frac{1-\frac{(1+\kappa_H)x}{W}}{(1+\kappa_L)^2}\right)\right. \\
&\quad \left. + \left(x^{\frac{1}{2}}-1\right)\left(\frac{1-\frac{(\kappa_H-\kappa_L)x}{W}}{1+\kappa_L}\right) - \frac{2x}{W^{\frac{1}{2}}}\left(\frac{1-\frac{(\kappa_H-\kappa_L)x}{W^{\frac{1}{2}}}}{1+\kappa_L}\right)^{\frac{1}{2}} + \frac{x^{\frac{3}{2}}}{W^{\frac{1}{2}}}\right\}
\end{aligned}$$

Recall that the constraints on equilibrium solutions of x ensure that $\lim_{W\to\infty} x$ is finite and weakly positive. So:

$$\begin{aligned}
\lim_{W\to\infty}\Delta_{\kappa_L} &= -\lim_{W\to\infty} W\times \\
&\left\{\lim_{W\to\infty}\frac{W^{\frac{1}{2}}}{2(1+\kappa_L)^{\frac{3}{2}}} + \lim_{W\to\infty}\left[\frac{x^{\frac{1}{2}}(\kappa_L-\kappa_H)-(c+\kappa_L)}{(1+\kappa_L)^2}\right] + \lim_{W\to\infty}\left(\frac{x^{\frac{1}{2}}-1}{1+\kappa_L}\right)\right\} = -\infty
\end{aligned}$$

So $V(\text{join}|x)$ is strictly decreasing in κ_L.

(b) Changes in κ_H

(i) Note that:

$$\frac{dV(\text{join}|x)}{d\kappa_H} = \frac{-1}{2[x\Upsilon(x)]^{\frac{1}{2}}}\left\{\left(\frac{x^2}{1+\kappa_L}\right) + \left[\frac{W-2(\kappa_H-\kappa_L)x}{1+\kappa_L}\right]\left(\frac{\Delta_{\kappa_H}}{\Delta_x}\right)\right\}$$

So:

$$\frac{dV(\text{join}|x)}{d\kappa_H} < 0 \quad \Leftrightarrow \quad 0 < \frac{x^2}{W} + \left[1 - \frac{2(\kappa_H - \kappa_L)x}{W}\right]\left(\frac{\Delta_{\kappa_H}}{\Delta_x}\right)$$

Recall that the constraints on equilibrium solutions of x ensure that $\lim_{W\to\infty} x$ is finite and weakly positive. Also recall from the Proof of Proposition C1 that: $\lim_{W\to\infty} \Delta_x = -\infty$. So our result holds when W is large if: $\lim_{W\to\infty} \Delta_{\kappa_H} - \infty$.

(ii) Now note that:

$$\begin{aligned}
\Delta_{\kappa_H} &= W\left[\frac{1}{2}\Upsilon(x)^{-\frac{1}{2}}\left(\frac{\partial\Upsilon}{\partial\kappa_H}\right)\right] + x^{\frac{1}{2}}\left\{(\kappa_L - \kappa_H)\,2\left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]\frac{1}{2}\Upsilon(x)^{-\frac{1}{2}}\left(\frac{\partial\Upsilon}{\partial\kappa_H}\right) - \left[\Upsilon(x)^{\frac{1}{2}} - x^{\frac{1}{2}}\right]^2\right\} \\
&\quad - (c+\kappa_L)\left(\frac{\partial\Upsilon}{\partial\kappa_H}\right) \\
&= \left[\frac{W - 2x(\kappa_L - \kappa_H)}{2\Upsilon(x)^{\frac{1}{2}}} + x^{\frac{1}{2}}(\kappa_L - \kappa_H) - (c+\kappa_L)\right]\left(\frac{\partial\Upsilon}{\partial\kappa_H}\right) - x^{\frac{1}{2}}\Upsilon(x) + 2x\Upsilon(x)^{\frac{1}{2}} - x^{\frac{3}{2}} \\
&= -\left[\frac{(1+\kappa_L)^{\frac{1}{2}}\left[W^{\frac{1}{2}} - \frac{2x(\kappa_L-\kappa_H)}{W^{\frac{1}{2}}}\right]}{2\left(1 - \frac{(\kappa_H-\kappa_L)x}{W}\right)^{\frac{1}{2}}} + x^{\frac{1}{2}}(\kappa_L - \kappa_H) - (c+\kappa_L)\right]\left(\frac{x}{1+\kappa_L}\right) \\
&\quad - x^{\frac{1}{2}}\left(\frac{W - (\kappa_H - \kappa_L)x}{1+\kappa_L}\right) + 2x\left(\frac{W - (\kappa_H - \kappa_L)x}{1+\kappa_L}\right)^{\frac{1}{2}} - x^{\frac{3}{2}} \\
&= -W\left\{\left[\frac{(1+\kappa_L)^{\frac{1}{2}}\left[1 - \frac{2x(\kappa_L-\kappa_H)}{W}\right]}{2W^{\frac{1}{2}}\left(1 - \frac{(\kappa_H-\kappa_L)x}{W}\right)^{\frac{1}{2}}} + \frac{x^{\frac{1}{2}}(\kappa_L - \kappa_H) - (c+\kappa_L)}{W}\right]\left[\frac{x}{(1+\kappa_L)}\right]\right. \\
&\quad \left. + x^{\frac{1}{2}}\left(\frac{1 - \frac{(\kappa_H-\kappa_L)x}{W}}{1+\kappa_L}\right) - \frac{2x}{W^{\frac{1}{2}}}\left(\frac{1 - \frac{(\kappa_H-\kappa_L)x}{W}}{1+\kappa_L}\right)^{\frac{1}{2}} + \frac{x^{\frac{3}{2}}}{W^{\frac{1}{2}}}\right\}
\end{aligned}$$

Recall that the constraints on equilibrium solutions of x ensure that $\lim_{W\to\infty} x$ is finite and weakly positive. So:

$$\lim_{W\to\infty} \Delta_{\kappa_H} = -\lim_{W\to\infty} W \times \lim_{W\to\infty}\left(\frac{x^{\frac{1}{2}}}{1+\kappa_L}\right) = -\infty$$

So $V(\text{join}|x)$ is strictly decreasing in κ_H.

This implies that for large W, joining the ICC lowers the total violence in the state. QED

Proposition C3

Note that the dictator's probabilities of survival in office, conditional on his joining decisions, are:

$$\pi\left(\text{don't join}\right)=\frac{1}{1+c}\qquad\text{and}\qquad\pi\left(\text{join}\right)=\left[\frac{x}{\Upsilon(x)}\right]^{\frac{1}{2}}$$

By the Proof of Proposition C1, a dictator only joins the ICC if:

$$\begin{aligned}\theta\left(c\right)\quad&\equiv\quad W\left[\frac{x}{\Upsilon\left(x\right)}\right]-\left(\kappa_H-\kappa_L\right)\left[\frac{x^{\frac{3}{2}}\left[\Upsilon\left(x\right)^{\frac{1}{2}}-x^{\frac{1}{2}}\right]}{\Upsilon\left(x\right)}\right]-\frac{W}{\left(1+c\right)^2}\geq 0\\&\Leftrightarrow\quad W\left[\frac{x}{\Upsilon\left(x\right)}-\frac{1}{\left(1+c\right)^2}\right]\geq\frac{\left(\kappa_H-\kappa_L\right)x^{\frac{3}{2}}\left[\Upsilon\left(x\right)^{\frac{1}{2}}-x^{\frac{1}{2}}\right]}{\Upsilon\left(x\right)}\end{aligned}$$

As shown in the Proof of Proposition C1, $x^{\frac{1}{2}}<\Upsilon\left(x\right)^{\frac{1}{2}}$ for all relevant values of x, meaning that the right-hand side of the equation above is always positive. This therefore ensures that the left-hand side of the equation is also always positive. So if the dictator joins the ICC, then:

$$\pi\left(\text{don't join}\right)=\frac{1}{1+c}<\pi\left(\text{join}\right)=\left[\frac{x}{\Upsilon(x)}\right]^{\frac{1}{2}}$$

QED.

2 Empirics

2.1 Descriptive Statistics

Table A1: Descriptive Statistics for Models in the Paper

	Political Competition				Violence				Leader Survival			
Variable	Mean	Min.	Max.	N	Mean	Min.	Max.	N	Mean	Min.	Max.	N
Year	2007	1998	2018	1,194								
ICC join	0.03	0	1	1,194								
Political competition	2.08	0	4	1,100								
Multiparty elections†	−0.50	−3.35	1.66	1,017								
Leader tenure†	11.01	0	49	1,180								
Post-join					0.75	0	1	672	0.75	0	1	783
Out of office									0.13	0	1	783
Tenure year									7.52	0	46	783
Total violence	0.81	0	7	1,194	0.68	0	7	672	0.67	0	7	783
Intra-state violence	0.72	0	7	1,194	0.67	0	7	651	0.66	0	7	783
Inter-state violence	0.09	0	6	1,194					0.01	0	1	783
PRIO violence†					0.34	0	6	672				
Log(GDP per capita)	7.90	5.23	11.15	1,107	6.98	5.35	9.39	646	6.78	5.35	9.39	755
Rule of law	−0.71	−2.61	1.84	1,194	−0.75	−2.13	0.46	672	−0.74	−2.13	0.46	783
Foreign aid*	19.31	−21.41	24.66	1,095	20.71	−17.05	23.89	672	20.74	−17.05	23.89	783
Polity					2.02	−7	8	638	2.21	−7	8	716
Africa	0.42	0	1	1,194	0.72	0	1	672	0.69	0	1	783
Middle East/North Africa	0.23	0	1	1,194	0.06	0	1	672	0.06	0	1	783
Asia	0.26	0	1	1,194					0.12	0	1	783
Central & South America	0.05	0	1	1,194								

†appendix only *hyperbolic sine transformation

2.2 Hypothesis 1: Political Competition is Associated with Joining the ICC

2.2.1 Robustness: Ratification

In these models, we exclude the two cases of joining the ICC by special declaration and only include ratification of the Rome Statute as the event. We thank an anonymous reviewer for suggesting the inclusion of these tests.

Table A2: Political Competition is Associated with Ratification

Event: Ratification

Explanatory Variable		
Political competition	0.66**	0.63**
	(0.28)	(0.28)
Control Variables		
GDP per capita, logged	-0.43	-0.47
	(0.30)	(0.31)
Foreign aid	0.02	0.02
	(0.11)	(0.11)
Rule of law	0.67	0.72
	(0.48)	(0.50)
Violence: total	-0.01	
	(0.16)	
Violence: intra-state		0.04
		(0.18)
Violence: inter-state		-0.29
		(0.61)
Region controls	Yes	Yes
Events	23	23
States	84	84
Observations (state-year)	940	940

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

2.2.2 Robustness: Regime Measures

To check the robustness of our definition of dictatorship, we consider three alternative measures. First, we use POLYARCHY scores from the Varieties of Democracy ("V-Dem") dataset (Coppedge et al., 2021). There is no standard cut-off on this variable to separate democracies from dictatorships (Boese, 2019). We thus employ two cut-offs for our state-year observations: a POLYARCHY score of 0.4 in one set of models ("V-DEM-0.4 dataset") and a score of 0.5 in a different set of models ("V-DEM-0.5 dataset").

There is, of course, significant overlap in these definitions, so many of the same states appear across each model. For example, 80 of the 83 states (96%) in the V-DEM-0.4 dataset also appear among the 93 dictatorships in the POLITY-5 dataset in the main paper. Likewise, 88 of the 98 states (90%) in the V-DEM-0.5 dataset also appear among the 93 dictatorships in the POLITY-5 dataset. Second, we use the dichotomous measure of regime-type provided in Boix, Miller, and Rosato (2013) at the request of an anonymous reviewer. Results are as follows.

Table A3: Political Competition Models - Alternative Measures of Dictatorship

Event: Joining the ICC

Explanatory Variable						
Political competition	0.74*	0.72*	0.54**	0.53**	0.68**	0.64**
	(0.41)	(0.42)	(0.27)	(0.27)	(0.30)	(0.30)
Control Variables						
GDP per capita, logged	-0.57	-0.59	-0.27	-0.28	-0.44	-0.54
	(0.38)	(0.39)	(0.27)	(0.28)	(0.33)	(0.35)
Foreign aid	-0.06	-0.06	-0.00	-0.00	0.01	0.01
	(0.08)	(0.08)	(0.09)	(0.09)	(0.11)	(0.10)
Rule of law	0.35	0.37	0.31	0.33	1.12**	1.25**
	(0.61)	(0.62)	(0.48)	(0.48)	(0.55)	(0.59)
Violence: total	-0.09		0.00		-0.01	
	(0.23)		(0.16)		(0.18)	
Violence: intra-state		-0.06		0.02		0.10
		(0.25)		(0.17)		(0.21)
Violence: inter-state		-0.21		-0.13		-0.34
		(0.60)		(0.60)		(0.62)
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5	BMR	BMR
Events	12	12	20	20	21	21
States	74	74	91	91	85	85
Observations (state-year)	809	809	1,017	1,017	945	945

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

2.2.3 Robustness: Political Competition Measures

Table A4: Multiparty Elections and Continuous Exposure

Event: Joining the ICC

Explanatory Variable						
Multiparty elections	0.56***	0.55***	0.66**	0.65**	0.48**	0.45**
	(0.21)	(0.21)	(0.28)	(0.29)	(0.21)	(0.22)
Control Variables						
GDP per capita, logged	-0.44	-0.45*	-0.76**	-0.76**	-0.34	-0.37
	(0.27)	(0.27)	(0.35)	(0.35)	(0.25)	(0.26)
Foreign aid	0.05	0.05	-0.03	-0.03	-0.00	-0.01
	(0.15)	(0.16)	(0.09)	(0.09)	(0.08)	(0.08)
Rule of law	0.64	0.65	0.58	0.59	0.29	0.35
	(0.45)	(0.46)	(0.59)	(0.60)	(0.45)	(0.47)
Violence: total	0.09		0.01		0.09	
	(0.12)		(0.16)		(0.12)	
Violence: intra-state		0.11		0.02		0.12
		(0.13)		(0.17)		(0.12)
Violence: inter-state		-0.08		-0.03		-0.04
		(0.53)		(0.54)		(0.51)
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	Polity-5	Polity-5	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5
Events	30	30	17	17	26	26
States	84	84	71	71	87	87
Observations (state-year)	892	892	747	747	971	971

$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table A5: Leader Tenure and Continuous Exposure

Event: Joining the ICC

Explanatory Variable						
Leader Tenure	-0.06**	-0.06**	-0.07*	-0.06*	-0.08**	-0.08**
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Control Variables						
GDP per capita, logged	-0.44	-0.48*	-0.64*	-0.68**	-0.35	-0.38
	(0.27)	(0.28)	(0.34)	(0.35)	(0.26)	(0.26)
Foreign aid	0.14	0.15	-0.02	-0.02	0.02	0.02
	(0.19)	(0.19)	(0.10)	(0.10)	(0.11)	(0.13)
Rule of law	0.47	0.57	0.05	0.16	0.15	0.19
	(0.46)	(0.48)	(0.56)	(0.58)	(0.47)	(0.48)
Violence: total	0.05		-0.02		0.06	
	(0.13)		(0.16)		(0.13)	
Violence: intra-state		0.12		0.03		0.10
		(0.13)		(0.17)		(0.14)
Violence: inter-state		-0.53		-0.46		-0.36
		(0.68)		(0.64)		(0.60)
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	Polity-5	Polity-5	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5
Events	31	31	18	18	27	27
States	90	90	79	79	94	94
Observations (state-year)	1,009	1,009	861	861	1,085	1,085

$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

2.2.4 Robustness: Risk Set

Table A6: Political Competition and New Exposure

Event: Joining the ICC

<u>*Explanatory Variable*</u>						
Political competition	0.64**	0.59**	0.80*	0.78*	0.52**	0.50*
	(0.28)	(0.29)	(0.41)	(0.42)	(0.26)	(0.27)
<u>*Control Variables*</u>						
GDP per capita, logged	-0.42	-0.50	-0.58	-0.60	-0.24	-0.26
	(0.31)	(0.33)	(0.38)	(0.39)	(0.28)	(0.28)
Foreign aid	0.07	0.06	-0.05	-0.05	0.03	0.03
	(0.15)	(0.14)	(0.08)	(0.08)	(0.09)	(0.09)
Rule of law	0.75	0.85	0.37	0.40	0.26	0.29
	(0.52)	(0.54)	(0.61)	(0.62)	(0.49)	(0.49)
Violence: total	0.04		-0.10		-0.05	
	(0.16)		(0.23)		(0.16)	
Violence: intra-state		0.10		-0.07		-0.02
		(0.17)		(0.25)		(0.17)
Violence: inter-state		-0.33		-0.19		-0.24
		(0.62)		(0.60)		(0.59)
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	Polity-5	Polity-5	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5
Events	23	23	12	12	20	20
States	84	84	74	74	91	91
Observations (state-year)	933	933	809	809	1,017	1,017

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

Table A7: Multiparty Elections and New Exposure

Event: Joining the ICC

Explanatory Variable						
Multiparty elections	0.55***	0.53**	0.75**	0.75**	0.48**	0.46**
	(0.21)	(0.22)	(0.29)	(0.30)	(0.21)	(0.22)
Control Variables						
GDP per capita, logged	-0.48*	-0.49*	-0.83**	-0.83**	-0.36	-0.37
	(0.27)	(0.28)	(0.35)	(0.36)	(0.25)	(0.25)
Foreign aid	0.07	0.08	-0.04	-0.04	0.02	0.03
	(0.15)	(0.16)	(0.08)	(0.08)	(0.09)	(0.09)
Rule of law	0.72	0.73	0.65	0.65	0.32	0.33
	(0.47)	(0.47)	(0.59)	(0.60)	(0.45)	(0.46)
Violence: total	0.09		-0.02		0.05	
	(0.13)		(0.17)		(0.12)	
Violence: intra-state		0.10		-0.02		0.06
		(0.13)		(0.18)		(0.13)
Violence: inter-state		-0.08		0.00		-0.10
		(0.54)		(0.54)		(0.52)
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	Polity-5	Polity-5	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5
Events	30	30	17	17	26	26
States	84	84	71	71	87	87
Observations (state-year)	892	892	747	747	971	971

$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table A8: Leader Tenure and New Exposure

Event: Joining the ICC

Explanatory Variable						
Leader Tenure	-0.06**	-0.06*	-0.08*	-0.07*	-0.08**	-0.08**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Control Variables						
GDP per capita, logged	-0.46*	-0.52*	-0.70**	-0.74**	-0.38	-0.42
	(0.28)	(0.29)	(0.34)	(0.34)	(0.26)	(0.27)
Foreign aid	0.17	0.18	-0.03	-0.03	0.06	0.06
	(0.18)	(0.19)	(0.08)	(0.09)	(0.14)	(0.15)
Rule of law	0.54	0.65	0.06	0.16	0.19	0.27
	(0.47)	(0.49)	(0.56)	(0.59)	(0.47)	(0.49)
Violence: total	0.05		-0.08		0.01	
	(0.13)		(0.16)		(0.13)	
Violence: intra-state		0.12		-0.03		0.06
		(0.13)		(0.17)		(0.14)
Violence: inter-state		-0.56		-0.47		-0.44
		(0.69)		(0.64)		(0.61)
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	Polity-5	Polity-5	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5
Events	31	31	18	18	27	27
States	90	90	79	79	94	94
Observations (state-year)	1,009	1,009	861	861	1,085	1,085

$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

2.2.5 Ancillary Test: Rebel Group Competition

As an ancillary test of our theory, we examined rebel group competition as an alternative measure of our explanatory variable. Unfortunately, existing data for this concept do not provide enough coverage to run a time-series, cross-national model. However, they do provide some suggestive, albeit limited, evidence in favor of our theory. In Figure A1, we use a measure of the relative strength of rebel groups from Cunningham, Gleditsch and Salehyan (2013). These data include 42 of the dictatorships in our larger dataset, which are coded for a maximum of 14 years.[1] Each observation corresponds to one state-year in which the state was fighting against at least one rebel group; thus, the number of observations per state is not consistent.[2] This generates 208 state-year observations.[3] For each observation, we plot the frequency of states facing rebel groups of various strengths, disaggregated by whether that state eventually joined the ICC. Our coding of rebel group competition corresponds to Cunningham, Gleditsch and Salehyan (2013)'s classification of rebel group strength as much weaker (low competition), weaker (moderate competition), and parity/stronger (high competition). A state-year unit is colored black in Figure A1 if the state joined the ICC at some point in time and grey if the state never joined the ICC.

Figure A1 provides additional evidence for our theory. The data include 92 state-year observations in which the rebel group is "much weaker" than the government, meaning that the government faces low levels of political competition. In 9 (or 9.8%) of these observations, the government joined the ICC at some point in time.[4] Overall, the data show that governments facing low rebel group competition are unlikely to join the ICC.

Next, the data include 99 state-year observations in which the rebel group is "weaker" than the government, meaning that the government faces moderate rebel group competition. In 38 (or 38.4%) of these observations, the state joined the ICC at some point in time. This statistic suggests that a government facing moderate rebel group competition receives more benefit from being an ICC member than a government facing low competition.

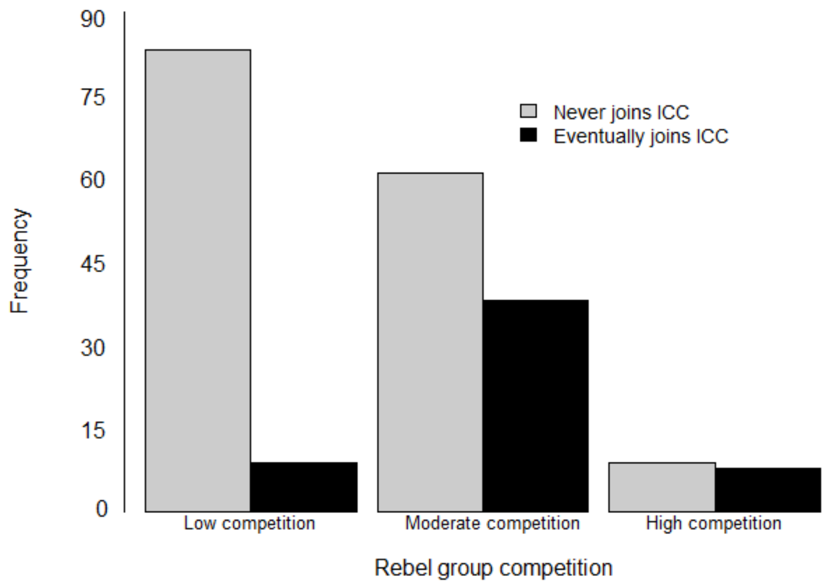
[1] The dataset covers 1998-2011 only.

[2] Additionally, states that join the ICC are only included for years prior to joining. For example, Afghanistan generates five observations (1998-2001 and 2003), Guinea generates two (1998-1999), and Myanmar generates 14 (1998-2011).

[3] We use state-years as the unit because rebel groups change strength over time and aggregating to the level of the state would obscure these differences.

[4] There is only one observation in which a state facing a much weaker rebel group joined the ICC in the same year: Afghanistan in 2003. Obviously, the extensive US government support for the Afghan government in 2003 makes this an idiosyncratic case. Had the US military not been deployed on Afghan soil, the true level of political competition within Afghanistan would have been very high.

Figure A1: Rebel Group Competition and Joining the ICC



Note: Level of competition corresponds to rebel group strength relative to government as coded by Cunningham, Gleditsch and Salehyan (2013): low competition ("much weaker"), moderate competition ("weaker"), or high competition ("parity" or "stronger"). The unit of analysis is a state-year.

Finally, it is very rare for a rebel group to be at "parity or stronger" than the government (meaning that the rebel group is of equal or greater strength). Only 17 of the 208 observations fall into this category. In such circumstances, a government faces high rebel group competition. In 8 (or 47%) of the observations in this category, the government joined the ICC at some point in time. So governments that face high rebel group competition appear to receive a higher benefit from being an ICC member than governments that face moderate or low rebel group competition. Overall, the evidence from Figure A1 supports the first observable implication of our theory: political competition increases the likelihood of joining the ICC.

2.3 Hypothesis 2: Joining is Associated with a Decrease in Violence

The inclusion of a state into H2 is defined based on the inclusion rule for H1. Any state that joins the ICC as a dictatorship (defined using either the Polity, V-Dem 0.4, or V-Dem 0.5 cutoff) is included in the sample for H2. Thus the numbers are slightly different across each set of models.

2.3.1 Robustness: Ratification

In these models, we exclude the two cases of joining the ICC by special declaration and only include ratification of the Rome Statute as the event. We thank an anonymous reviewer for suggesting the inclusion of these tests.

Table A9: Ratification is Associated with Decreased Violence

Dependent Variable: Violence

	Total	Intra-state
Explanatory Variable		
Post-ratification	-0.52*	-0.55*
	(0.30)	(0.30)
Control Variables		
GDP per capita, logged	-0.35*	-0.36*
	(0.21)	(0.21)
Foreign aid	1.25***	1.26***
	(0.16)	(0.16)
Rule of law	-4.46***	-4.58***
	(0.37)	(0.38)
Polity	-0.05	-0.05
	(0.04)	(0.04)
Region controls	Yes	Yes
States	30	30
Observations (state-year)	592	592

$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

2.3.2 Robustness: Regime Measures

To check the robustness of our definition of dictatorship, we consider three alternative measures. First, we use POLYARCHY scores from the Varieties of Democracy ("V-Dem") dataset (Coppedge et al., 2021). There is no standard cut-off on this variable to separate democracies from dictatorships (Boese, 2019). We thus employ two cut-offs for our state-year observations: a POLYARCHY score of 0.4 in one set of models ("V-DEM-0.4 dataset") and a score of 0.5 in a different set of models ("V-DEM-0.5 dataset").

There is, of course, significant overlap in these definitions, so many of the same states appear across each model. For example, 80 of the 83 states (96%) in the V-DEM-0.4 dataset also appear among the 93 dictatorships in the POLITY-5 dataset in the main paper. Likewise, 88 of the 98 states (90%) in the V-DEM-0.5 dataset also appear among the 93 dictatorships in the POLITY-5 dataset. Second, we use the dichotomous measure of regime-type provided in Boix, Miller, and Rosato (2013) at the request of an anonymous reviewer. Results are as follows.

Table A10: Violence Models – Alternative Measures of Dictatorship

Dependent Variable: Violence

	Total	Intra-state	Total	Intra-state	Total	Intra-state
Explanatory Variable						
Post-join	-1.20***	-1.17***	-0.20	-0.22	-1.56***	-1.63***
	(0.38)	(0.38)	(0.29)	(0.29)	(0.37)	(0.37)
Control Variables						
GDP per capita, logged	0.06**	0.06**	0.02	0.02	-0.20	-0.19
	(0.03)	(0.03)	(0.02)	(0.02).	(0.21)	(0.21)
Foreign aid	1.71***	1.70***	1.23***	1.24***	1.50***	1.52***
	(0.20)	(0.20)	(0.15)	(0.15)	(0.19)	(0.19)
Rule of law	-4.19***	-4.42***	-4.26***	-4.38***	-4.58***	-4.75***
	(0.45)	(0.47)	(0.38)	(0.39).	(0.43)	(0.45)
Polity					0.00	0.01
					(0.05)	(0.05)
Polyarchy	-5.76***	-6.15***	-3.43***	-3.42***		
	(1.91)	(2.00)	(1.30)	(1.31)		
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5	BMR	BMR
States	19	19	28	28	29	29
Observations (state-year)	373	373	554	554	571	571

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

2.3.3 Robustness: Violence Measures

Table A11: Joining the ICC Decreases PRIO Violence

Dependent Variable: Violence

Explanatory Variable			
Post-join	-0.42*	-1.43***	-0.50**
	(0.25)	(0.33)	(0.25)
Control Variables			
GDP per capita, logged	-0.23	-0.04*	-0.06***
	(0.17)	(0.02)	(0.02)
Foreign aid	0.73***	1.06***	0.92***
	(0.13)	(0.14)	(0.13)
Rule of law	-1.71***	-1.52***	-1.92***
	(0.24)	(0.27)	(0.26)
Polity	-0.06**		
	(0.03)		
Polyarchy (V-Dem)		-2.64**	-4.53***
		(1.29)	(1.11)
Region controls	Yes	Yes	Yes
Dataset	Polity-5	V-Dem-0.4	V-Dem-0.5
States	30	19	28
Observations (state-year)	592	373	554

$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

2.3.4 Robustness: Post-2002 Only

Table A12: Joining the ICC Decreases Violence: Post-2002 Data Only

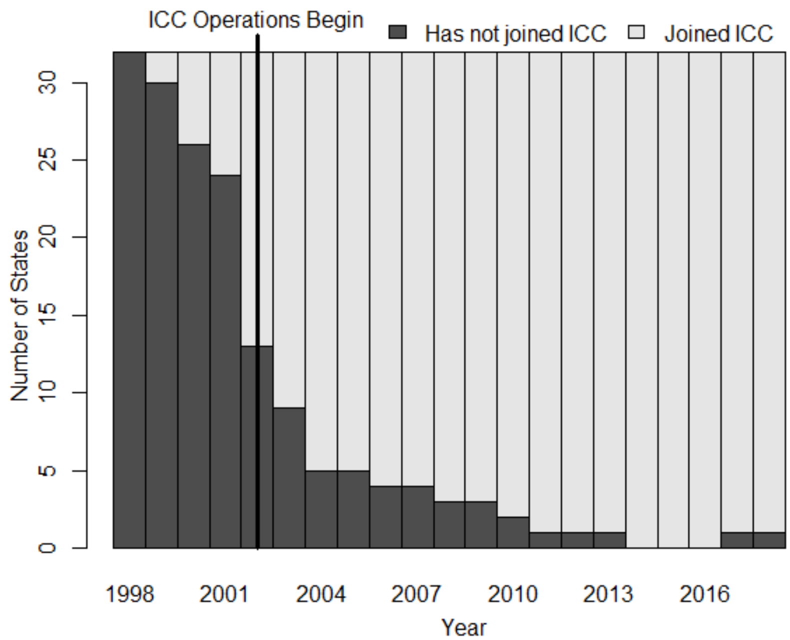
Dependent Variable: Violence

Type of Violence	Total	Intra-state
Explanatory Variable		
Post-join	-0.04	-0.12
	(0.48)	(0.49)
Control Variables		
GDP per capita, logged	0.25	0.25
	(0.22)	(0.22)
Foreign aid	1.37***	1.37***
	(0.18)	(0.18)
Rule of law	-5.03***	-5.09***
	(0.47)	(0.47)
Polity	0.03	0.03
	(0.05)	(0.05)
Region controls	Yes	Yes
States	31	31
Observations (state-year)	499	499

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

The coefficient on Post-join in the post-2002 sample is not statistically significant, although it is signed consistently negative. The issue is that most dictators who would join the ICC had done so by 2002. Thus, most of the "Has not joined ICC" observations come from 1998–2001. Excluding those years excludes the vast majority of the "Has not joined ICC" observations, making comparisons of behavior before and after joining nearly impossible. Figure A2 illustrates this issue, using the Polity sample of the 31 dictatorships that join. When subsetting the sample to 2002 and later, one is losing everything to the left of the solid black line, leaving only the sample to the right. The light grey bars indicate the states that are under the ICC's jurisdiction in that year, while the darker grey bars indicate the states that have not yet joined. About 90% of the observations in the post-2002 period are from states that already joined the ICC, leaving an insufficient number of "has not joined ICC" states to make adequate comparisons.

Figure A2: Composition of the Polity Sample



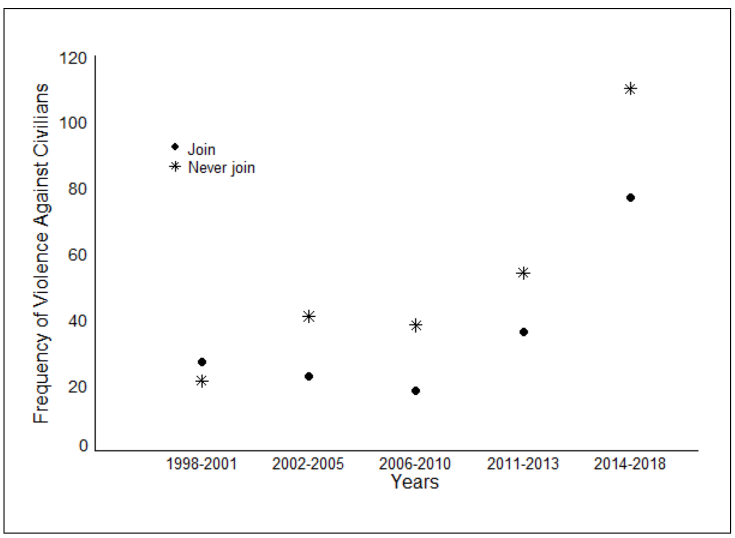
Note: The 31 states included here are the 31 dictatorships that join the ICC as dictatorships based on the Polity sample.

2.3.5 Ancillary Test: Violence Against Civilians

Our main modes use ordinal data on violence. While serious international crimes cannot occur with violence, we acknowledge that more precise measures of our theoretical concept may exist.[5] However, events-based measures of violence against civilians (from, for example, the Uppsala Conflict Data Program or the Armed Conflict Location and Event Data) will be subject to recency bias due to data availability, making them ill-suited for a pre/post test like the ones in the main text of the paper that goes back to 1998. However, as an ancillary test, we demonstrate the effect of ICC ratification on ACLED violence against civilians in African dictatorships that did and did not ratify the Rome Statute.[6]

The results appear below as Figure A3. While violence against civilians has increased universally over time, it has increased significantly less in African states that joined the ICC, versus those that never join.

Figure A3: ACLED Data



Note: Plot includes African dictatorships only due to data availability. Those that ever join the ICC are coded as "Join" and those that never join the ICC are coded as "Never join".

[5] We thank an anonymous reviewer for this observation.

[6] ACLED temporal coverage going back to 1998 is only available for Africa. However, since African states comprise the majority of our dictatorships that ratify, we have good coverage for most of our dataset.

2.4 Hypothesis 3: Joining is Associated with an Increase in Leader Survival

The inclusion of a state into H3 is defined based on the inclusion rule for H1. Any state that joins the ICC as a dictatorship (defined using either the Polity, V-Dem 0.4, or V-Dem 0.5 cutoff) is included in the sample for H3. Thus the numbers are slightly different across each set of models.

2.4.1 Robustness: Ratification

In these models, we exclude the two cases of joining the ICC by special declaration and only include ratification of the Rome Statute as the event. We thank an anonymous reviewer for suggesting the inclusion of these tests.

Table A13: Ratification is Associated with Increased Leader Survival

Event: Leader Removal

Explanatory Variable		
Post-ratification	-0.65*	-0.67*
	(0.38)	(0.39)
Control Variables		
GDP per capita, logged	0.05	0.05
	(0.21)	(0.21)
Foreign aid	-0.21	-0.21
	(0.20)	(0.20)
Violence: total	0.27**	
	(0.11)	
Violence: intra-state		0.27**
		(0.11)
Violence: inter-state		2.26
		(1.68)
Polity	0.25***	0.25***
	(0.06)	(0.06)
Region controls	Yes	Yes
Events	64	64
States	30	30
Observations (leader-year)	665	665

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

2.4.2 Robustness: Regime Measures

To check the robustness of our definition of dictatorship, we consider three alternative measures. First, we use POLYARCHY scores from the Varieties of Democracy ("V-Dem") dataset (Coppedge et al., 2021).[7] There is no standard cut-off on this variable to separate democracies from dictatorships (Boese, 2019). We thus employ two cut-offs for our state-year observations: a POLYARCHY score of 0.4 in one set of models ("V-DEM-0.4 dataset") and a score of 0.5 in a different set of models ("V-DEM-0.5 dataset").

There is, of course, significant overlap in these definitions, so many of the same states appear across each model. For example, 80 of the 83 states (96%) in the V-DEM-0.4 dataset also appear among the 93 dictatorships in the POLITY-5 dataset in the main paper. Likewise, 88 of the 98 states (90%) in the V-DEM-0.5 dataset also appear among the 93 dictatorships in the POLITY-5 dataset. Second, we use the dichotomous measure of regime-type provided in Boix, Miller and Rosato (2013) at the request of an anonymous reviewer. Results are as follows.

[7] Robustness checks using other measures of dictatorship are available from the authors on request.

Table A14: Leader Survival Models – Alternative Measures of Dictatorship
Event: Leader Removal

Explanatory Variable						
Post-join	-0.38	-0.33	-1.02***	-1.01***	-0.69	-0.69
	(0.79)	(0.80)	(0.36)	(0.36)	(0.44)	(0.44)
Control Variables						
GDP per capita, logged	-0.08	-0.07	-0.03	-0.03	0.05	0.05
	(0.06)	(0.06)	(0.03)	(0.03)	(0.21)	(0.21)
Foreign aid	-0.85**	-0.88**	-0.27*	-0.28*	-0.12	-0.13
	(0.37)	(0.37)	(0.16)	(0.17)	(0.17)	(0.17)
Violence: total	0.74***		0.30**		0.22*	
	(0.25)		(0.13)		(0.12)	
Violence: intra-state		0.73***		0.29**		0.22*
		(0.26)		(0.13)		(0.12)
Violence: inter-state		2.12		1.44		1.58
		(1.66)		(1.39)		(1.31)
Polity					0.25***	0.25***
					(0.06)	(0.06)
Polyarchy	7.09***	7.05***	4.62***	4.58***		
	(2.52)	(2.55)	(1.52)	(1.53)		
Region controls	Yes	Yes	Yes	Yes	Yes	Yes
Dataset	V-Dem-0.4	V-Dem-0.4	V-Dem-0.5	V-Dem-0.5	BMR	BMR
Events	25	25	55	55	64	64
States	19	19	28	28	29	29
Observations (leader-year)	382	373	598	598	643	643

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

2.4.3 Robustness: Post-2002 Only

Table A15: Joining the ICC Increases Leader Survival in Office: Post-2002 Data Only

Dependent Variable: Years to Losing Office
(Event: Removal from Office)

Explanatory Variable		
Post-join	-1.01^{*}	-1.02^{*}
	(0.53)	(0.53)
Control Variables		
GDP per capita, logged	0.02	0.02
	(0.22)	(0.22)
Foreign aid	-0.03	-0.03
	(0.19)	(0.19)
Violence: total	0.22^{*}	
	(0.11)	
Violence: intra-state		0.21^{*}
		(0.11)
Violence: inter-state		1.89
		(1.66)
Polity	0.18^{***}	0.18^{***}
	(0.06)	(0.06)
Region controls	Yes	Yes
Events	56	56
States	31	31
Observations (leader-year)	562	562

$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

References

Boese, Vanessa A. 2019. "How (Not) to Measure Democracy." *International Area Studies Review* 22(2):95–127.

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. "A Complete Data Set of Political Regimes, 1800–2007." *Comparative Political Studies* 46(12):1523–1554.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, A. Cornell, M.S. Fish, L. Gastaldi et al. 2021. "V-Dem codebook v11. Varieties of Democracy (V-dem) Project.".

Cunningham, David E., Kristian Skrede Gleditsch and Idean Salehyan. 2013. "Non-State Actors in Civil Wars: A New Dataset." *Conflict Management and Peace Science* 30(5):516–531.