Boundedly Rational Sources of State Stability and Failure

Michael Gill, mzgill@fas.harvard.edu

Motivation

- Huge literature on causes of wars/ends of wars:
- Formal theories abound:
  - Costly lottery models (e.g., Fearon, Powell)
  - Repeated bargaining with uncertainty (Powell)
  - Commitment issues due to power shifts or resources
  - Role of ethnic or economic divisions (Padro i Miquel)
  - Global games (Yanagizawa)

Basic Model

- Evolutionary model
- Population of mass 1
- Citizens can be from four possible strategy combinations
- Each period t is marked by a single attack on a target
  - Only *two individuals* interact each period
- Attacks are either "won" (by the attacker) or "defended" (by the target)
- After each period, one with lower payoff has chance to change strategy.

Table: Possible Citizen Strategy Combinations

	$G = 0$	$G = 1$
$C = 0$	$\pi_{1,t}(C = 0, G = 0)$	$\pi_{3,t}(C = 0, G = 1)$
$C = 1$	$\pi_{2,t}(C = 1, G = 0)$	$\pi_{4,t}(C = 1, G = 1)$

Outlaws Loyalists

Table: Key Symbols Used in Paper

Item	Description
t	the period (or round) indicator
π_i	citizen strategy type for group i; also, proportion of π_i relative to total population
α_i	attacker type in period t
τ_j	target type in period t
$\omega = \Omega$	attack outcome in period t. 1, if attack succeeds; 0, if attack defended
$\frac{u^{\tau_i,t}(\omega)}{u^{\alpha_i,t}(\omega,\tau_j) + u^{\tau_i,t}(\omega)}$	ratio of target's payoff over the attacker and target's summed payoffs
$\frac{u^{\alpha_i,t}(\omega,\tau_j)}{u^{\alpha_i,t}(\omega,\tau_j) + u^{\tau_i,t}(\omega)}$	ratio of attacker's payoff over the attacker and target's summed payoffs
μ	per period base utility for each citizen
β	amount removed from target (and given to attacker) if attack is successful
δ	additional penalty to a loyalist following a successful attack
κ	punishment to attacker if attack is unsuccessful, and if target was a loyalist

How it works...

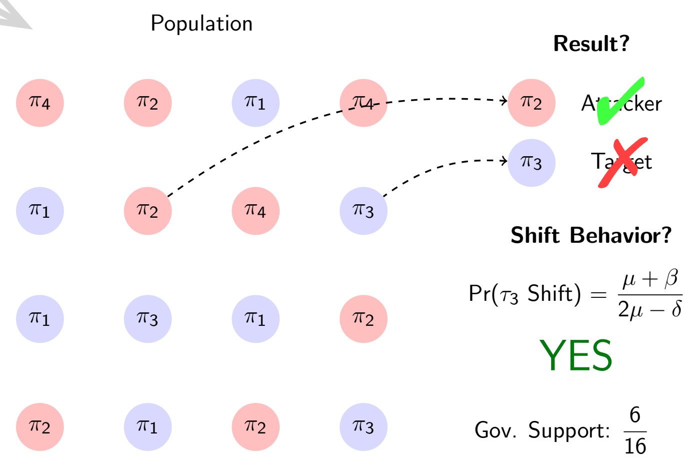
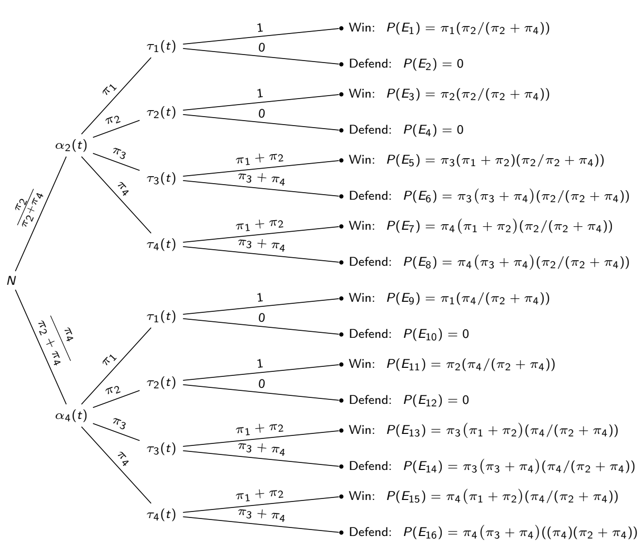


Table: Effects of Attacks on Various Targets

	Non-Loyalists		Loyalists			
	τ_1	τ_2	τ_3		τ_4	
	Wins	Wins	Wins	Defended	Wins	Defended
Pr(Attacker Wins):	1	1	$\pi_3 + \pi_4$		$\pi_3 + \pi_4$	
Pr(Attack Defended):	0	0	$\pi_1 + \pi_2$		$\pi_1 + \pi_2$	
$u_{\alpha_i}(t\vert\tau_j,\omega)$:	$\mu + \beta$	$\mu + \beta$	$\mu + \beta$	$\mu - \kappa$	$\mu + \beta$	$\mu - \kappa$
$u_{\tau_j}(t\vert\alpha_i,\omega)$:	$\mu - \beta$	$\mu - \beta$	$\mu - \beta - \delta$	μ	$\mu - \beta - \delta$	μ
$\frac{u_{\alpha_i}(t\vert\tau_j,\omega)}{u_{\alpha_i}(t\vert\tau_j,\omega) + u_{\tau_j}(t\vert\alpha_i,\omega)}$:	$\frac{\mu+\beta}{2\mu}$	$\frac{\mu+\beta}{2\mu}$	$\frac{\mu+\beta}{2\mu-\delta}$	$\frac{\mu-\kappa}{2\mu-\kappa}$	$\frac{\mu+\beta}{2\mu-\delta}$	$\frac{\mu-\kappa}{2\mu-\kappa}$
$\frac{u_{\tau_j}(t\vert\alpha_i,\omega)}{u_{\alpha_i}(t\vert\tau_j,\omega) + u_{\tau_j}(t\vert\alpha_i,\omega)}$:	$\frac{\mu-\beta}{2\mu}$	$\frac{\mu-\beta}{2\mu}$	$\frac{\mu-\beta-\delta}{2\mu-\delta}$	$\frac{\mu}{2\mu-\kappa}$	$\frac{\mu-\beta-\delta}{2\mu-\delta}$	$\frac{\mu}{2\mu-\kappa}$



Some results...

$$\mathbf{f_1} \equiv \frac{d\pi_1}{dt} = (\pi_3 + \pi_4)\left(\pi_2\frac{\mu - \beta}{2\mu} - \pi_1\frac{\mu + \beta}{2\mu}\right) \tag{1a}$$

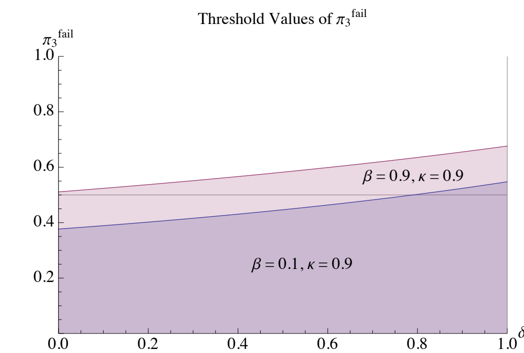
$$\mathbf{f_2} \equiv \frac{d\pi_2}{dt} = \pi_2\Bigg((\pi_3 + \pi_4)(\pi_1 + \pi_2)\frac{\mu + \beta}{2\mu - \delta} + (\pi_1 - \pi_4)\frac{\mu + \beta}{2\mu} - (\pi_3 + \pi_4)^2\frac{\mu}{2\mu - \kappa} - (\pi_2 + \pi_4)\frac{\mu - \beta}{2\mu}\Bigg) \tag{1b}$$

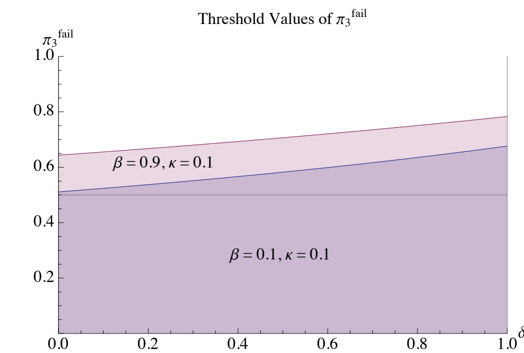
$$\mathbf{f_3} \equiv \frac{d\pi_3}{dt} = (\pi_3 + \pi_4)\Bigg((\pi_3 + \pi_4)^2\frac{\mu}{2\mu - \kappa} + \pi_4(\pi_1 + \pi_2)\frac{\mu - \beta - \delta}{2\mu - \delta} - \pi_3(\pi_1 + \pi_2)\frac{\mu + \beta}{2\mu - \delta}\Bigg) \tag{1c}$$

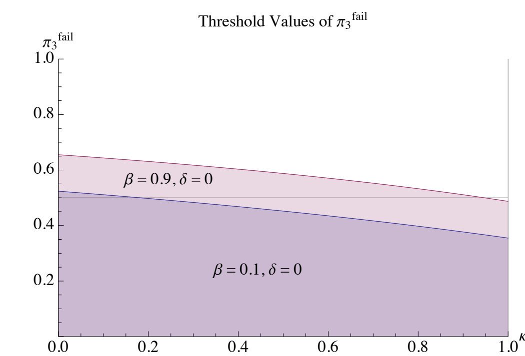
$$\mathbf{f_4} \equiv \frac{d\pi_4}{dt} = \pi_4\Bigg((\pi_1 + \pi_2)\frac{\mu + \beta}{2\mu} + \pi_3(\pi_1 + \pi_2)\frac{\mu + \beta}{2\mu - \delta} - (\pi_3 + \pi_4)^2\frac{\mu}{2\mu - \kappa} - \pi_4(\pi_1 + \pi_2)\frac{\mu - \beta - \delta}{2\mu - \delta} - \pi_2(\pi_1 + \pi_2)\Bigg) \tag{1d}$$

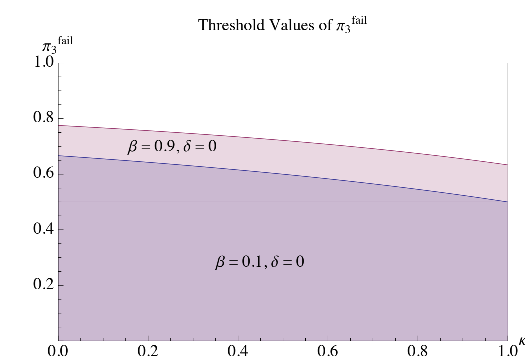
$$\pi_3^{fail} = \frac{(\beta + 1)(\kappa - 2)}{\delta + \kappa - \beta(2 - \kappa) - 4} \tag{2}$$

This equation provides our first primary result—namely, if the initial conditions of $\pi_{3,t=0} < \pi_3^{fail}$, then society will decay over time to a point where no citizens support the government. We know this because if $\pi_{3,t=0} < \pi_3^{fail}$ it implies $\mathbf{f_3} < 0$.









Extensions

- Include government as (rational?) actor
- Make population discrete
  - Permits stochastic interpretation of model
- Explore μ_t, etc. (Example: $\mu_t = \int_{\underline{u}}^{\overline{u}} u\, dH$).
- Explore simulations/relative rates of convergence
- Different replicator dynamics
- Non-loyalists have chance for defense
- Change targeting structure