Choosing How to Cooperate:
A Repeated Public-Goods Model of International Relations
(Technical Supplement)

Randall W. Stone
Department of Political Science, University of Rochester

Branislav L. Slantchev
Department of Political Science, University of California – San Diego

Tamar R. London
Department of Political Science, Pennsylvania State University

February 11, 2008

Abstract. This technical supplement includes the full proofs for the formal results in the paper.

Proof of Lemma 1. Take some $C \in S$. To simplify notation, let

$$f(a) = a^{c-b} > 0 \quad \text{and} \quad f'(a) = (c-b)a^{c-b-1} > 0.$$

We can now rewrite the constraints as follows:

$$\begin{aligned} \mathrm{MC}(a, C) &= (1-\delta)^{-1}\left[(1-\delta^{T+1})f(a) - (\delta - \delta^{T+1})(C + w_h)\right] \\ \mathrm{CR}(a, C) &= (1-\delta^T)\left[C + w_h - f(a)\right]. \end{aligned}$$

To see how the two constraints behave as a function of demand, take the derivatives:

$$\frac{\partial \mathrm{MC}}{\partial a} = \left(\frac{1-\delta^{T+1}}{1-\delta}\right) f'(a) > 0 \quad \text{and} \quad \frac{\partial \mathrm{CR}}{\partial a} = -(1-\delta^T)f'(a) < 0.$$

These now imply that the constraints either intersect precisely once or not at all. The necessary and sufficient condition for an intersection is $\mathrm{MC}(1, C) \leq \mathrm{CR}(1, C) \Leftrightarrow (1-\delta)/(1-\delta^T) \leq C + w_h - 1$, which yields

$$\mathrm{X}(C) = \frac{(1-\delta^T)(C + w_h)}{2 - \delta - \delta^T} \geq 1, \tag{1}$$

as the condition for the intersection. If $\mathrm{X}(C)$ satisfies (1), then the intersection is at

$$f(\tilde{a}) = \mathrm{X}(C), \tag{2}$$

where $\tilde{a} = \sqrt[c-b]{\mathrm{X}(C)}$ by the definition of $f(\cdot)$. We can now use this fact to substitute (2) into (MC) and (CR) to obtain:

$$\mathrm{MC}(\tilde{a}, C) = \mathrm{CR}(\tilde{a}, C) = (1-\delta)\,\mathrm{X}(C)$$

at the intersection, which yields the value of $\widetilde{w}(C)$ for the case of intersection stated in the lemma. Note in particular that $(1-\delta)\,\mathrm{X}(C)$ does not depend on a itself, only on the coalition size. This now gives us the smallest contributor size for C that admits a solution which would satisfy both constraints. If $\underline{w}(C) \geq (1-\delta)\,\mathrm{X}(C)$, then there exist values of a such that (a, C) is admissible. Otherwise, there is no solution and this coalition does not permit any admissible profiles. If the constraints do not intersect, then $\mathrm{MC}(a, C) > \mathrm{CR}(a, C)$ for all values of a. Therefore, it suffices to check whether $\mathrm{MC}(1, C)$ admits a solution. That is, if $\underline{w}(C) \geq \mathrm{MC}(1, C)$, then there exist values of a such that (a, C) is admissible. □

Proof of Proposition 1. Given that there exist admissible profiles, the hegemon picks the one that maximizes its payoff. Doing so is optimal for the hegemon because any alternative admissible profile will yield a worse payoff and any inadmissible profile fails at least one of the constraints, which implies that cooperation will unravel and the hegemon's payoff will be zero. Since the profile the hegemon chooses is admissible, both (MC) and (CR) are satisfied, and all members prefer to contribute while non-members free-ride on their efforts. Since the constraints are derived from the strategies specified in this section, these strategies form a stationary SPE of the game. In fact, these strategies will support any admissible profile in a stationary SPE of the continuation game after the initial choice by the hegemon, so the hegemon is effectively choosing which stationary SPE to play. □

Proof of Proposition 2. Define the optimal choices as follows:

(i) if $c/b < K_1$, then $\underline{w}^* = \text{MC}_1(a^*) > \text{CR}_1(a^*)$, where $a^* = \sqrt[c-b]{\frac{b(\mathcal{W}+w_h)}{c(1+\epsilon)}}$;

(ii) if $c/b \in [K_1, K_2]$, then $\underline{w}^* = \text{MC}_1(a^*) = \text{CR}_1(a^*)$, where: $a^* = \sqrt[c-b]{\frac{\mathcal{W}+w_h}{K_2}}$;

(iii) if $c/b > K_2$, then $\underline{w}^* = \text{CR}_1(a^*) > \text{MC}_1(a^*)$ where: $a^* = \sqrt[c-b]{\frac{b(\mathcal{W}+w_h)}{c}}$,

where $K_2 = 1 + (1-\delta)(1+\epsilon(1-\delta^T))/(1-\delta^T) > K_2/(1+\epsilon) = K_1 > 1$. The hegemon's optimization problem is:

$$\max_{a,\underline{w}} \pi_h(a,\underline{w}) \text{ subject to } \underline{w} \geq \max\{\text{MC}_1(a), \text{CR}_1(a)\} \;\&\; \underline{w} \geq 0 \;\&\; a \geq 1.$$

The Lagrangian is:

$$\mathcal{L} = \pi_h(a,\underline{w}) - \lambda_1(\text{MC}_1(a) - \underline{w}) - \lambda_2(\text{CR}_1(a) - \underline{w}) - \lambda_3(1-a) - \lambda_4(-\underline{w}),$$

with the Kuhn-Tucker conditions:

$$\frac{\partial \mathcal{L}}{\partial a} = \frac{\partial \pi_h(a,\underline{w})}{\partial a} - \lambda_1 \frac{\partial \text{MC}_1(a)}{\partial a} - \lambda_2 \frac{\partial \text{CR}_1(a)}{\partial a} + \lambda_3 = 0; \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial \underline{w}} = -\epsilon a^b + \lambda_1 + \lambda_2 + \lambda_4 = 0; \tag{4}$$

$$\lambda_1 \geq 0, \quad \text{MC}_1(a) \leq \underline{w} \quad \& \quad \lambda_1(\text{MC}_1(a) - \underline{w}) = 0; \tag{5}$$

$$\lambda_2 \geq 0, \quad \text{CR}_1(a) \leq \underline{w} \quad \& \quad \lambda_2(\text{CR}_1(a) - \underline{w}) = 0; \tag{6}$$

$$\lambda_3 \geq 0, \quad -a \leq -1 \quad \& \quad \lambda_3(1-a) = 0;$$

$$\lambda_4 \geq 0, \quad -\underline{w} \leq 0 \quad \& \quad \lambda_4(-\underline{w}) = 0. \tag{7}$$

Since we are looking for an interior solution, we set $\lambda_3 = \lambda_4 = 0$. Observe now that at least one of the constraints must be binding at a solution. To see that this must be the case, suppose that it is not. Since both constraints are slack, $\text{MC}_1(a) - \underline{w} > 0$ and $\text{CR}_1(a) - \underline{w} > 0$ at the solution, and so (5) and (6) imply that $\lambda_1 = \lambda_2 = 0$. But then (4) requires that $\epsilon a^b + \lambda_4 = 0$. Since $\epsilon a^b > 0$, this inequality cannot be satisfied for any $a \geq 1$ because $\lambda_4 \geq 0$ must hold by (7). Therefore, at least one of the constraints must be binding at a solution.

To simplify the algebra, define:

$$\Delta = \left[1 - \delta(1+\epsilon(1-\delta^T))\right]^{-1} > 0 \qquad R = (1-\delta^{T+1})\Delta > 0$$

$$P = [1+\epsilon(1-\delta^T)](1-\delta^T)^{-1} > 0 \qquad S = \delta(1-\delta^T)\Delta > 0.$$

and note that we can rewrite $K_2 = (1 + PR)/(1 + PS) > K_2/(1 + \epsilon) = K_1 > 1$.

Suppose first that $\mathrm{CR}_1(a)$ is slack, and so $\lambda_2 = 0$. In this case, (4) implies that $\lambda_1 = \epsilon a^b > 0$, which means that (5) requires that $\underline{w} = \mathrm{MC}_1(a)$; that is, $\mathrm{MC}_1(a)$ must bind. Substituting this in (3), we obtain:

$$b\left[\mathcal{W} + w_h - \epsilon \mathrm{MC}_1(a)\right] a^{b-1} - c a^{c-1} - \epsilon a^b \frac{\partial \mathrm{MC}_1(a)}{\partial a} = 0,$$

which simplifies to:

$$\frac{(1-\delta)\left[b(\mathcal{W} + w_h)a^{b-1} - c(1+\epsilon)a^{c-1}\right]}{1 - \delta(1 + \epsilon(1-\delta^T))} = 0,$$

whose unique solution is:

$$a_1 = \sqrt[c-b]{\frac{b(\mathcal{W} + w_h)}{c(1+\epsilon)}}.$$

Since we require that $\underline{w}_1 = \mathrm{MC}_1(a_1) > 0$, observe that this will hold whenever

$$\frac{c}{b} \leq \frac{1 - \delta^{T+1}}{(1+\epsilon)(\delta - \delta^{T+1})}.$$

We now need to ensure that $\mathrm{CR}_1(a_1) - \underline{w}_1 < 0$. This inequality reduces to:[1]

$$\frac{(1-\delta^T)(c(1+\epsilon) - b)}{1 + \epsilon(1-\delta^T)} - \frac{(1-\delta^{T+1})b - (\delta - \delta^{T+1})(1+\epsilon)c}{1 - \delta(1 + \epsilon(1-\delta^T))} < 0,$$

which we can compactly rewrite as $(c(1+\epsilon) - b)/P - Rb - S(1+\epsilon)c < 0$, or:

$$\frac{c}{b} < \frac{1 + PR}{(1+\epsilon)(1+PS)} \equiv K_1.$$

This inequality can be satisfied because $K_1 > 1 \Leftrightarrow \epsilon < \overline{\epsilon}$, which holds by Assumption 1, or:

$$\begin{aligned}
\frac{1+PR}{1+PS} &> 1 + \epsilon \\
1 + \frac{(1-\delta)[1 + \epsilon(1-\delta^T)]}{1-\delta^T} &> 1 + \epsilon \\
1 + \epsilon(1-\delta^T) &> \frac{(1-\delta^T)\epsilon}{1-\delta} \\
1 &> \epsilon\left(\frac{\delta(1-\delta^T)}{1-\delta}\right) \\
\frac{1-\delta}{\delta(1-\delta^T)} &> \epsilon,
\end{aligned}$$

[1]Note that, letting $x = 1 + \epsilon(1-\delta^T)$, we obtain:

$$\frac{1+PR}{1+PS} = 1 + \frac{(1-\delta)x}{1-\delta^T} = 1 + (1-\delta)P$$

where the last line follows from Assumption 1. Now observe that:

$$\frac{1-\delta^{T+1}}{(1+\epsilon)(\delta-\delta^{T+1})} > \frac{1+PR}{(1+\epsilon)(1+PS)} \Leftrightarrow \epsilon < \overline{\epsilon}$$

becase:

$$\begin{aligned}\frac{1-\delta^{T+1}}{\delta(1-\delta^T)} &> 1 + \frac{(1-\delta)[1+\epsilon(1-\delta^T)]}{1-\delta^T}\\ 1-\delta^{T+1} &> \delta(1-\delta^T) + \delta(1-\delta)[1+\epsilon(1-\delta^T)]\\ 1 &> \delta[1+\epsilon(1-\delta^T)]\\ \frac{1-\delta}{\delta(1-\delta^T)} &> \epsilon,\end{aligned}$$

In other words, whenever $c/b < K_1$, then $\underline{w}_1 = \mathrm{MC}_1(a_1) > 0$ will certainly be satisfied as well. Therefore, $(a_1, \underline{w}_1)$ is valid solution if, and only if, $b/c < K_1$, which yields the first case in the proposition.

Suppose now that $\mathrm{MC}_1(a)$ is slack, and so $\lambda_1 = 0$. In this case, (4) implies that $\lambda_2 = \epsilon a^b > 0$, which means that (6) requires that $\underline{w} = \mathrm{CR}_1(a)$; that is, $\mathrm{CR}_1(a)$ must bind. Substituting this in (3), we obtain:

$$b\left[\mathcal{W} + w_h - \epsilon \mathrm{CR}_1(a)\right] a^{b-1} - ca^{c-1} - \epsilon a^b \frac{\partial \mathrm{CR}_1(a)}{\partial a} = 0,$$

which simplifies to:

$$\frac{b(\mathcal{W}+w_h)a^{b-1} - ca^{c-1}}{1+\epsilon(1-\delta^T)} = 0,$$

whose unique solution is:

$$a_2 = \sqrt[c-b]{\frac{b(\mathcal{W}+w_h)}{c}}.$$

Observe now that $\underline{w}_2 = \mathrm{CR}_1(a_2) > 0$ is always satisfied, so we only need to check that $\mathrm{MC}_1(a_2) - \underline{w}_2 < 0$ to ensure that this is a solution. The inequality reduces to:

$$\frac{(1-\delta^{T+1})b - (\delta-\delta^{T+1})c}{1-\delta(1+\epsilon(1-\delta^T))} - \frac{(1-\delta^T)(c-b)}{1+\epsilon(1-\delta^T)} < 0,$$

which we can simplify to:

$$\frac{c}{b} < 1 + \left(\frac{1-\delta}{1-\delta^T}\right)\left[1+\epsilon(1-\delta^T)\right]$$

which can be satisfied. We can then rewrite the condition as $Rb - Sc - (c-b)/P < 0$, or:

$$\frac{c}{b} > \frac{1+PR}{1+PS} \equiv K_2 = (1+\epsilon)K_1 > K_1.$$

We have already established that $K_1 > 1$, so this inequality can be satisfied for some values of b and c. Hence, $(a_2, \underline{w}_2)$ is a valid solution if, and only if, $c/b > K_2$. This yields the third case in the proposition.

Finally, suppose that both constraints are binding. Solving $\mathrm{MC}_1(a) = \mathrm{CR}_1(a)$ yields:

$$\begin{aligned}
a_3^{c-b} &= \frac{(1-\delta^T)(\mathcal{W} + w_h)}{1 - \delta^T + (1-\delta)(1 + \epsilon(1-\delta^T))} \\
a_3^{c-b} &= \frac{(\mathcal{W} + w_h)(1 + PS)}{1 + PR}, \\
a_3 &= \sqrt[c-b]{\frac{\mathcal{W} + w_h}{K_2}},
\end{aligned}$$

and so $\underline{w}_3 = \mathrm{MC}_1(a_3) = \mathrm{CR}_1(a_3)$. The derivation ensures that $\mathrm{MC}_1(a_3) - \underline{w}_3 = 0$ and $\mathrm{CR}_1(a_3) - \underline{w}_3 = 0$, as required. To satisfy (4), we need $\lambda_1 > 0$, $\lambda_2 > 0$, and $\lambda_1 + \lambda_2 = \epsilon a_3^b > 0$, which can certainly be satisfied. Noting now that:

$$\begin{aligned}
\frac{\partial \pi_h(a)}{\partial a} &= b(\mathcal{W} + w_h - \epsilon \underline{w})a^{b-1} - ca^{c-1} \\
\frac{\partial \mathrm{MC}_1(a)}{\partial a} &= \frac{(1-\delta^{T+1})(c-b)a^{c-b-1}}{1 - \delta x} = R(c-b)a^{c-b-1} \\
\frac{\partial \mathrm{CR}_1(a)}{\partial a} &= \frac{-(1-\delta^T)(c-b)a^{c-b-1}}{x} = \frac{-(c-b)a^{c-b-1}}{P},
\end{aligned}$$

where we let $x = 1 + \epsilon(1 - \delta^T)$. Using our definitions, we can also rewrite:

$$\begin{aligned}
\mathrm{MC}_1(a) &= Ra^{c-b} - S(\mathcal{W} + w_h) \\
\mathrm{CR}_1(a) &= \frac{\mathcal{W} + w_h - a^{c-b}}{P}.
\end{aligned}$$

We can rewrite the Lagrangian as follows after dividing through by a_3^{b-1}:

$$b(\mathcal{W} + w_h - \epsilon \underline{w}_3) - ca_3^{c-b} - \lambda_1 R(c-b)a_3^c + \frac{(\epsilon a_3^b - \lambda_1)(c-b)a_3^c}{P} = 0.$$

Since $\underline{w}_3 = \mathrm{CR}_1(a_3) = (\mathcal{W} + w_h)(K_2 - 1)/(PK_2) = (\mathcal{W} + w_h)(R - S)/(1 + PR) = (\mathcal{W} + w_h)(1-\delta)/K_2$, we can simplify this to:

$$\frac{b(\mathcal{W} + w_h)}{K_2}\left[1 + \frac{1-\delta}{1-\delta^T}\right] - ca_3^{c-b} - \lambda_1 R(c-b)a_3^c + \frac{(\epsilon a_3^b - \lambda_1)(c-b)a_3^c}{P} = 0.$$

Substituting for a_3^{c-b} and rearranging terms gives us:

$$\frac{\mathcal{W} + w_h}{K_2}\left[c - b\left(1 + \frac{1-\delta}{1-\delta^T}\right)\right] = (c-b)a_3^c\left[\frac{\epsilon a_3^b - \lambda_1(1 + PR)}{P}\right].$$

Dividing through by $c - b$ then yields:

$$\frac{\mathcal{W} + w_h}{K_2}\left[1 - \frac{b(1-\delta)}{(c-b)(1-\delta^T)}\right] = a_3^c\left[\frac{\epsilon a_3^b - \lambda_1(1 + PR)}{P}\right].$$

Solving for λ_1 and noting that $(\mathcal{W} + w_h)/K_2 = a_3^{c-b}$ results in:

$$\lambda_1 = \left(\frac{1}{1+PR}\right)\left[\epsilon a_3^b - P a_3^{-b}\left[1 - \frac{b(1-\delta)}{(c-b)(1-\delta^T)}\right]\right].$$

We need to ensure that $0 < \lambda_1 < \epsilon a_3^b$. Let $\zeta = 1 - b(1-\delta)/[(c-b)(1-\delta^T)]$ and note that $\zeta < 0 \Leftrightarrow c/b < (2-\delta)/(1-\delta^T)$. Since $1 + PR > 0$, $\lambda_1 > 0$ requires that $a_3^{2b} > \zeta P/\epsilon$. Analogously, $\lambda_1 < \epsilon a_3^b$ requires that $a_3^{2b} > -\zeta/(\epsilon R)$. We now have two general cases to consider.

Suppose first that $\zeta < 0$. This immediately satisfies $\lambda_1 > 0$ because $a_3^{2b} > 0$ and $P/\epsilon > 0$. Hence, we only need to ensure that $a_3^{2b} > -\zeta/(\epsilon R)$ holds. This inequality simplifies to $c/b > 1 + (1+\delta)/[(1-\delta^T)(1+\epsilon R a_3^{2b})] \equiv \underline{K}$. Since $a_3^{2b} \geq 1$, it follows that $\underline{K} \leq 1 + (1-\delta)/[(1-\delta^T)(1+\epsilon R)] = K_1$. [2]

Therefore, $c/b > K_1 \Rightarrow c/b > \underline{K}$ is sufficient to guarantee λ_1 is valid whenever $\zeta < 0$.

Suppose now that $\zeta > 0$. This immediately satisfies $\lambda_1 < \epsilon a_3^{2b}$ because $\epsilon R > 0$. Hence, we only need to ensure that $a_3^{2b} > \zeta P/ve$ holds. This inequality simplifies to $c/b < 1 + (1-\delta)P/[(1-\delta^T)(P - \epsilon a_3^{2b})] \equiv \overline{K}$. Since $a_3^{2b} \geq 1$, it follows that $\overline{K} \geq 1 + (1-\delta)P/[(1-\delta^T)(P-\epsilon)] = K_2$.[3] Therefore, $c/b < K_2 \Rightarrow c/b < \overline{K}_2$ is sufficient to guarantee that λ_1 is valid whenever $\zeta > 0$.

Putting all of these results together yields the range $c/b \in [K_1, K_2]$ that ensures that there exists a solution that satisfies the Kuhn-Tucker conditions when both constraints are binding. This yields the second case in the proposition.

To establish the claim about the comparative statics of the optimal demand, we need to show that $\frac{\partial a^*}{\partial w_h} > 0$, which is clearly true from inspection. To see that $\frac{\partial \underline{w}^*}{\partial w_h} > 0$ as well, we examine each case separately:

(i) $\underline{w}^* = \text{MC}_1(a^*) = \frac{Rb(\mathcal{W}+w_h)}{c(1+\epsilon)} - S(\mathcal{W} + w_h)$, and the derivative is:

$$\frac{\partial \underline{w}^*}{\partial w_h} = \frac{Rb}{c(1+\epsilon)} - S > 0 \Leftrightarrow \frac{c}{b} < \frac{1-\delta^{T+1}}{\delta(1-\delta^T)(1+\epsilon)} \equiv K.$$

[2] To see that this equality holds:

$$1 + \frac{1-\delta}{(1-\delta^T)(1+\epsilon R)} = 1 + \frac{(1-\delta)(1-\delta x)}{(1-\delta^T)(1-\delta x + \epsilon(1-\delta^{T+1}))}$$

where $x = 1 + \epsilon(1-\delta^T)$ and $1 + \epsilon R = 1 + \epsilon(1-\delta^{T+1})/(1-\delta x)$

$$= 1 + \frac{1-\delta x}{(1+\epsilon)(1-\delta^T)}$$

because $1 - \delta x + \epsilon(1-\delta^{T+1}) = (1+\epsilon)(1-\delta)$

$$= \left(\frac{1}{1+\epsilon}\right)\left[1 + \frac{(1-\delta)x}{1-\delta^T}\right] = \frac{K_2}{1+\epsilon} = K_1.$$

[3] Since $P - \epsilon = 1/(1-\delta^T)$, we have $1 + (1-\delta)P/[(1-\delta^T)(P-\epsilon)] = 1 + (1-\delta)P = K_2$.

Since $c/b < K_1$ in this case, it will be sufficient to show that $K_1 < K$. Using $K_1 = K_2/(1+\epsilon)$, this last inequality reduces to $\epsilon < \overline{\epsilon}$, which holds.

(ii) $\underline{w}^* = \text{MC}_1(a^*) = \text{CR}_1(a^*) = (\mathcal{W} + w_h)(K_2 - 1)/(PK_2)$, and the derivative is:

$$\frac{\partial \underline{w}^*}{\partial w_h} = \frac{K_2 - 1}{PK_2} > 0,$$

where the inequality follows from $K_2 > 1$ and $P > 0$.

(iii) $\underline{w}^* = \text{CR}_1(a^*) = (\mathcal{W} + w_h)(c - b)/(Pc)$, and the derivative is:

$$\frac{\partial \underline{w}^*}{\partial w_h} = \frac{c - b}{Pc} > 0,$$

where the inequality follows from $P > 0$.

Therefore, the optimal minimum contributor size is strictly increasing in w_h as well. This establishes the comparative statics results. □

Proof of Proposition 3. First we show that if the exclusion cost is sufficiently low, the discriminatory regime is always preferred to the public goods regime. Let the subscripts D and P denote the discriminatory and public goods regimes, respectively. Since a^*_D is exactly the unconstrained optimum when $\underline{w} = 0$, it follows that the hegemon's payoff under the discriminatory regime is precisely the maximum unconstrained payoff net the discrimination cost: $\pi_h(a^*_D, 0) - m$. But $\pi_h(a^*_D, 0) > \pi_h(a^*, \underline{w}^*)$ implies that there exists $\overline{m} > 0$ such that $\pi_h(a^*_D, 0) - \overline{m} = \pi_h(a^*, \underline{w}^*)$. Hence, for any $m < \overline{m}$, the hegemon will strictly prefer the discriminatory regime to the public goods regime.

We now show that for any exclusion cost there exists a size threshold above which hegemons prefer to discriminate. The payoff from the public goods regime is:

$$\pi_h(a, \underline{w}) = (\mathcal{W} + w_h - \epsilon \underline{w})a^b - a^c = a^b[\mathcal{W} + w_h - \epsilon \underline{w} - a^{c-b}]$$

which, using $(a^*, \underline{w}^*)$ for the case where both constraints bind, is:

$$= (\mathcal{W} + w_h)\left(\frac{\mathcal{W} + w_h}{K_2}\right)^{\frac{b}{c-b}}\left[1 - \frac{\epsilon(K_2 - 1)}{PK_2} - \frac{1}{K_2}\right] \tag{8}$$

$$= (\mathcal{W} + w_h)\left(\frac{\mathcal{W} + w_h}{K_2}\right)^{\frac{b}{c-b}}\left[\frac{(P - \epsilon)(K_2 - 1)}{PK_2}\right] \tag{9}$$

and since $K_2 - 1 = (1 - \delta)P$, we get:

$$= (\mathcal{W} + w_h)^{\frac{c}{c-b}}(1 - \delta)(P - \epsilon)\left(\frac{1}{K_2}\right)^{\frac{c}{c-b}} \tag{10}$$

The utility obtained from the discriminatory regime is:

$$\pi_h(a_D^*,0) - m = \left(\frac{b(W+w_h)}{c}\right)^{\frac{b}{c-b}} \left[\mathcal{W} + w_h - \frac{b(\mathcal{W}+w_h)}{c}\right] - m$$
$$= (\mathcal{W}+w_h)^{\frac{c}{c-b}} \left(\frac{c-b}{c}\right)\left(\frac{b}{c}\right)^{\frac{b}{c-b}} - m.$$

The hegemon will prefer to pay the exclusion cost and discriminate rather than produce public goods whenever $\pi_h(a_D^*,0) - \pi_h(a^*,\underline{w}^*) > m$, or whenever

$$\Pi \cdot (\mathcal{W}+w_h)^{\frac{c}{c-b}} > m, \tag{11}$$

with

$$\Pi = \left(\frac{c-b}{c}\right)\left(\frac{b}{c}\right)^{\frac{b}{c-b}} - (1-\delta)(P-\epsilon)\left(\frac{1}{K_2}\right)^{\frac{c}{c-b}} > 0, \tag{12}$$

where the $\Pi > 0$ follows from the fact that $\pi_h(a_D^*,0) > \pi_h(a^*,\underline{w}^*)$, which implies that the left-hand side of (11) must be positive. Define:

$$\underline{w}_h(m) = \left(\frac{m}{\Pi}\right)^{\frac{c-b}{c}} - \mathcal{W},$$

and observe that for a given m, taking $w_h > \underline{w}_h(m)$ will satisfy (11). Therefore, for any $m > 0$, every hegemon with $w_h > \underline{w}_h(m)$ will strictly prefer to discriminate. □

Proof of Proposition 4. Let the subscript S denote slow-time (without an institution). We first derive the optimal profile without institution where players interact every other period only using the usual trigger strategies. To simplify notation, define the following short-hand expressions:

$$\hat{\Delta} = \left[1-\delta^2-\epsilon\delta^2\left(1-\delta^T\right)\right]^{-1} \quad \text{and} \quad P = \epsilon + \left(1-\delta^T\right)^{-1}$$
$$\hat{R} = \left(1-\delta^{T+2}\right)\hat{\Delta} \quad \text{and} \quad \hat{S} = \delta^2\left(1-\delta^T\right)\hat{\Delta}$$

Note that $\hat{\Delta} > \Delta > 0$, $\hat{R} > \hat{S} > 0$, $R > \hat{R}$, and $S > \hat{S}$. For the contributor to be willing to contribute it must be the case that:

$$\frac{\pi_i(a,C)}{1-\delta} \geq (1+\delta)\left[(C+w_h)a^b - w_i a^b\right] + \left(\frac{\delta^{T+2}}{1-\delta}\right)\pi_i(a,C)$$

Notice that the only change from $\text{MC}_1(a)$ is that when the potential contributor defects, it benefits from the defection for two periods instead of one. Making the necessary substitutions and simplifying yields:

$$\underline{w} \geq \hat{R}a^{c-b} - \hat{S}(\mathcal{W}+w_h) \equiv \text{MC}_S(a)$$

We now find the optimal level of contribution and optimal size of the marginal contributor in the case where both constraints are binding by setting the contributor's constraint equal to the credibility constraint. The credibility constraint is the same as before, $\mathrm{CR}_1(a)$, because the hegemon decides whether to punish by looking at present and future payoffs, which have not changed. (The loss from an additional period of free riding in the past is irrelevant.) The constraints are equal when:

$$P^{-1}\left(\mathcal{W} + w_h - a^{c-b}\right) = \hat{R}a^{c-b} - \hat{S}(\mathcal{W} + w_h).$$

Solving this for a yields: $(a_S^*)^{c-b} = (\mathcal{W} + w_h)(1 + P\hat{S})/(1 + P\hat{R})$. Note now that $(1 + P\hat{S}) = \Delta$ and $(1 + P\hat{R}) = \Delta\hat{K}_2$, where

$$\hat{K}_2 = 1 + \frac{(1-\delta^2)[1+\epsilon(1-\delta^T)]}{1-\delta^T}.$$

We can rewrite this result as follows:

$$a_S^* = \sqrt[c-b]{\frac{\mathcal{W} + w_h}{\hat{K}_2}}.$$

Since both constraints are binding, we obtain $\underline{w}_S^* = \mathrm{CR}_1(a_S^*) = \mathrm{MC}_S(a_S^*)$.

Turning now to the proof of the claim, let $(a_S^*, \underline{w}_S^*)$ be the optimal profile we just derived and $(a^*, \underline{w}^*)$ be the optimal (fast-time, with institution present) profile from Proposition 2 when both constraints are binding. Note now that:

$$a^* > a_S^* \Leftrightarrow \hat{K}_2 > K_2 \Leftrightarrow 1 > \delta,$$

and so the optimal demand with the institution is strictly larger than without it. Observe now that:

$$\underline{w}^* < \underline{w}_S^* \Leftrightarrow \hat{K}_2 > K_2,$$

which we already know to hold. That is, the size of the smallest contributor is strictly smaller with the institution than without it. □

Proof of Proposition 5. We need to compare the expected payoffs with an institution, $\pi_h(a^*, \underline{w}^*)$ from Proposition 2 with both constraints binding, and without an institution, $\pi_h(a_S^*, \underline{w}_S^*)$ from the proof of Proposition 4. Letting $k > 0$ represent the cost of building the institution, the payoff from doing so exceeds the payoff of having no institution whenever $\pi_h(a^*, \underline{w}^*) - k > \pi_h(a_S^*, \underline{w}_S^*)$. The first payoff is simplified in (10). Noting that $\hat{K}_2 = 1 + (1-\delta^2)P$, we can simplify the latter payoff as:

$$\pi_h(a_S^*, \underline{w}_S^*) = (\mathcal{W} + w_h)^{\frac{c}{c-b}}(1-\delta^2)(P-\epsilon)\left(\frac{1}{\hat{K}_2}\right)^{\frac{c}{c-b}}.$$

We can now rewrite the condition as follows:

$$\hat{\Pi} \cdot (\mathcal{W} + w_h)^{\frac{c}{c-b}} > k, \tag{13}$$

where

$$\widehat{\Pi} = \left[(1-\delta) \left(\frac{1}{K_2} \right)^{\frac{c}{c-b}} - (1-\delta^2) \left(\frac{1}{\widehat{K}_2} \right)^{\frac{c}{c-b}} \right] (P-\epsilon) > 0, \tag{14}$$

where the inequality follows from the fact that we can show that the expression in the brackets is positive:

$$\frac{1-\delta}{1-\delta^2} > \left(\frac{K_2}{\widehat{K}_2} \right)^{\frac{c}{c-b}} \Leftrightarrow \widehat{K}_2 > K_2.$$

This means that we can now define the following function:

$$\underline{\widehat{w}}_h(k) = \left(\frac{k}{\widehat{\Pi}} \right)^{\frac{c-b}{c}} - \mathcal{W}.$$

For any $k > 0$, every hegemon with $w_h > \underline{\widehat{w}}_h(k)$ strictly prefers to build an institution because its size satisfies (13). This establishes the claim. □

Proof of Proposition 6. The hegemon will prefer the private-goods regime to "slow time" whenever $\pi_h(a_D^*, 0) - m > \pi_h(a_S^*, \underline{w}_S^*)$, which reduces to:

$$\Pi_S \cdot (\mathcal{W} + w_h)^{\frac{c}{c-b}} > m,$$

where

$$\Pi_S = \left(\frac{c-b}{c} \right)^{\frac{b}{c-b}} - (1-\delta^2)(P-\epsilon) \left(\frac{1}{\widehat{K}_2} \right)^{\frac{c}{c-b}}.$$

Hence, any hegemon with size

$$w_h > \left(\frac{m}{\Pi_S} \right)^{\frac{c-b}{c}} - \mathcal{W} \equiv \underline{h}$$

will strictly prefer paying the exclusion cost to continuing without an institution.

We now find the size threshold beyond which a hegemon would prefer the private-goods regime to the public-goods regime with an institution. The threshold will be similar to the one above, but actions take place in "fast time," and the hegemon bears an additional cost to make this change to "fast time." The hegemon will prefer to pay the exclusion cost whenever $\pi_h(a_D^*, 0) - m > \pi_h(a^*, \underline{w}^*) - k$, or:

$$\Pi \cdot (\mathcal{W} + w_h)^{\frac{c}{c-b}} > m - k,$$

where Π is defined in (12). Hence, any hegemon with size

$$w_h > \left(\frac{m-k}{\Pi} \right)^{\frac{c-b}{c}} - \mathcal{W} \equiv \overline{h}$$

will strictly prefer paying the exclusion cost to building an institution. Note now that:

$$\underline{h} < \overline{h} \Leftrightarrow k < \frac{m(\Pi_S - \Pi)}{\Pi},$$

where $\Pi_S - \Pi = \hat{\Pi} > 0$ and $\hat{\Pi}$ is defined in (14). In this case, all $w_h \in [\underline{h}, \overline{h}]$ prefer building an institution to discriminating, and prefer discriminating to living without an institution. $\square$