Technical Appendices to Cason & Mui (2013), "Coordinating Resistance through Communication and Repeated Interaction"

These Appendices are included for the reader's reference in case they are interested. The paper is self-contained and can be read without referring to the Appendices.

We consider the no deviation conditions of the leader, the beneficiary, and the victim in each Appendix below. The most interesting parts of these proofs, however, are the no deviation conditions of the beneficiary in some key cases, and the reader may want to simply skim the parts of proofs that contain these materials. To facilitate such a quick reading, we provide information about the "key points" of each Appendix in the Table of Contents below.

Table of Contents

	in the DAC subgame changes as a function of the value of p (pp.21-24)
Appendix D (pp.26-31)	The proof regarding how to support the No Transgression outcome as equilibrium in the indefinitely repeated CR game with social preferences and incomplete information using the trigger strategies in Weingast (1995, 1997). • Key no deviation condition for the *SP*-type beneficiary (p.27)
Appendix E (pp.32-63)	The proof that shows how to support (eventual) No Transgression as a separating equilibrium in the indefinitely repeated CR game with social preferences and incomplete information. The main result here is stated as Result 2 in these appendices. • Intuitive explanation of Result 2 (pp.33-35) • Formal description of strategies (pp.35-36) • Formal description of belief updating rules (pp.36-38) • Formal discussion and intuitive explanation of the no deviation condition for the *SP*-type beneficiary in period 1 (pp.41-44)
Appendix F (pp.69-94)	Part 1 is the proof that cheap talk can be informative in the one-shot CR game with social preferences and communication. Part 2 shows that cheap talk can also be informative in the indefinitely repeated CR game with social preferences and incomplete information and communication. • Formal definition of equilibrium in the one-shot CR game with social preferences and incomplete information and communication (p.64) • Formal statement and intuitive explanation of how cheap talk can be informative in the one-shot CR game with social preferences and incomplete information and communication (pp.65-66) • Key no deviation condition for the *SP*-type beneficiary at the *responder action stage* (pp.66-67) • Key no deviation condition for the *SP*-type beneficiary at the *responder communication stage* (p.71) • Formal statement and intuitive explanation of how cheap talk can be informative and can help support (eventual) No Transgression as a separating equilibrium in the indefinitely repeated CR game with social preferences and incomplete information and communication (pp.78-82) • Formal discussion and intuitive explanation of the key no deviation condition for the *SP*-type beneficiary at the responder action stage in period 1 (pp.82-86) • Discussion and intuitive explanation of the key no

Appendix A

Supporting the No Transgression Outcome in the Indefinitely Repeated CR Game with Standard Preferences Using the Trigger Strategies in Weingast (1995, 1997) Discussed in Footnote 4 in the Text.

Let DAC denote Divide-and-Conquer transgression, TAB denote Transgression against Both responders, TA denotes Divide-and-Conquer transgression against A, TB denotes Divide-and-Conquer transgression against B, NT denote No Transgression, AC denotes acquiesce, and CH denotes challenge. We consider an indefinitely repeated CR game with standard preferences with a continuation probability of δ.

1. No deviation Conditions for the Responders

We first consider whether a responder may want to deviate. There are five different cases to consider:

1.1 There Has Been no Acquiescence to Earlier Transgression and the Leader Chooses DAC Transgression in the Current Period

1.1.1 No Deviation Condition for the Beneficiary

No deviation - (CH)

$$7+8(\delta+\delta^2+.........) \qquad (1)$$

Deviation - (AC)

$$9+2(\delta+\delta^2+.........) \qquad (2)$$

$$\therefore 7+8\frac{\delta}{1-\delta}\geq 9+2\frac{\delta}{1-\delta}$$

$$\Rightarrow 6\frac{\delta}{1-\delta}\geq 2$$

$$\Rightarrow 6\delta\geq 2-2\delta \qquad (3)$$

$$\Rightarrow \delta\geq\frac{1}{4}$$

1.1.2 No Deviation Condition for the Victim

No deviation - (CH)

$$7+8(\delta+\delta^2+.........) \qquad (4)$$

Deviation - (AC)

$$2+2(\delta+\delta^2+.........) \qquad (5)$$

$$\therefore \textit{No deviation}$$

1.2 There Has Been No Acquiescence to Earlier Transgression and the Leader Chooses TAB in the Current period

No Deviation Condition for a Responder

No deviation - (CH)

$$7+8\left(\frac{\delta}{1-\delta}\right) \tag{6}$$

Deviation - (AC)

$$2+2\left(\frac{\delta}{1-\delta}\right) \tag{7}$$

$\therefore \textit{No deviation}$

1.3 There Has Been No Acquiescence to Earlier Transgression and the Leader Chooses NT in the Current Period

No Deviation Condition for a Responder

No deviation - (AC)

$$8+8\left(\frac{\delta}{1-\delta}\right) \tag{8}$$

Deviation - (CH)

$$7+8\left(\frac{\delta}{1-\delta}\right) \tag{9}$$

$\therefore \textit{No deviation}$

1.4 There Has Been Acquiescence to Earlier Transgression

In this case, on the equilibrium path the players play (TAB, AC, AC) thereafter. If the leader plays *TAB*, the no deviation condition for a responder will be $2+2\left(\frac{\delta}{1-\delta}\right) > 1+2\left(\frac{\delta}{1-\delta}\right)$, which is satisfied.

1.5 There Has Been Acquiescence to Earlier Transgression but the Leader Plays Something other than TAB in the Current Period

Then it is straightforward to verify that the responder does not want to deviate.

To see this, consider the case when the leader plays DAC transgression. Then the no deviation condition for the beneficiary will be $9+2\left(\frac{\delta}{1-\delta}\right)>8+2\left(\frac{\delta}{1-\delta}\right)$, which is satisfied. The no deviation condition for the victim will be $2+2\left(\frac{\delta}{1-\delta}\right)>1+2\left(\frac{\delta}{1-\delta}\right)$, which is satisfied.

Consider the case when the leader plays NT. Then the no deviation condition for the responder will be $8+2\left(\frac{\delta}{1-\delta}\right)>7+2\left(\frac{\delta}{1-\delta}\right)$, which is satisfied.

2. No deviation Conditions for the Leader

2.1 There Has Been No Acquiescence to Earlier Transgression
No Deviation Condition for the leader

$$\text{No deviation (play } NT\text{): } 6+6\left(\frac{\delta}{1-\delta}\right) \tag{10}$$

$$TAB: 0+6\left(\frac{\delta}{1-\delta}\right) \tag{11}$$

$$DAC: 0+6\left(\frac{\delta}{1-\delta}\right) \tag{12}$$

$\therefore No\, deviation$

2.2 There Has Been Acquiescence to Some Transgression

No Deviation Condition for the leader:

$$\text{No deviation (play } TAB\text{) : } 12+12\left(\frac{\delta}{1-\delta}\right) \tag{13}$$

$$DAC: 8+12\left(\frac{\delta}{1-\delta}\right) \tag{14}$$

$$NT: 6+12\left(\frac{\delta}{1-\delta}\right) \tag{15}$$

$\therefore No\, deviation$

In conclusion, the critical discount factor that will support cooperation is $\delta^* = \frac{1}{4}$. Since the continuation probability in our indefinite repetition treatments is $\frac{7}{8} > \frac{1}{4}$, no transgression can be supported as an equilibrium by the trigger strategy in the indefinite repetition treatments.

Appendix B

Supporting the No Transgression Outcome in the Finitely Repeated CR Game with Standard Preferences Using the Trigger Strategies in Weingast (1995, 1997)

Consider a *T* period finitely repeated CR Game, in which it is common knowledge that the game will end for certain at period *T*.

We focus on the beneficiary's no deviation conditions, since the other no deviation conditions are easily verified.

We shall assume that in the last period, if neither responder has acquiesced previously the leader will randomize between DAC transgression against A (denoted TA) and DAC transgression against B (denoted TB) with probability 0.5, and that every responder will acquiesce when a DAC transgression occurs. Expecting this, in the second to the last period, if a beneficiary of a DAC transgression does not deviate—that is, if she challenges—her expected payoff is $(7+5.5)$. If she deviates—that is, if she acquiesces— her expected payoff is $(9+2)$. Since $(7+5.5)>(9+2)$, she will not deviate. That is, the beneficiary expects that cooperation will break down in the last period. However, deviation will trigger the (subgame perfect) equilibrium that involves the leader successfully transgressing against both responders. From the beneficiary's perspective, this is worse than the outcome of having an equal chance of being a beneficiary or a victim of a successful DAC transgression. Therefore, the beneficiary will still cooperate with the victim in the second to the last period.

We now show that if the players will play according to the equilibrium strategies starting from period $T-k$ onward, then the beneficiary will not deviate at period $T-k-1$. Expecting that every player will play according to the equilibrium strategies starting from period $T-k$ onward, if the beneficiary does not deviate at period $T-k-1$, she receives expected payoff

k terms

$$7+(8+8+...+8)+5.5=7+k\cdot 8+5.5 \tag{1}$$

If she deviates, she gets

k terms

$$9+(2+2+...+2)+2=9+k\cdot 2+2 \tag{2}$$

Since $7+k\cdot 8+5.5>9+k\cdot 2+2$, the beneficiary will not deviate.

Appendix C

Equilibria in the One-Shot CR Game with Social Preferences and Incomplete Information

This appendix sketches a model of the one-shot CR game with social preferences and incomplete information, in which an agent can either have standard own-money maximizing preferences or social preferences, and her type is her private information. Appendices D-F will then extend this model to consider repetition and communication. In Appendix G, we will discuss a model of the CR game with social preferences and complete information, in which an agent has social preferences with certainty. This model in Appendix G, which was discussed in the text, is the special case of the model discussed here in Appendix D.

Consider a model in which all agents are of two types. With probability p an agent has standard preferences, and with probability $(1-p)$ the agent has social preferences. An agent's type is her private information. Cox et al. (2007) assume that in a (two-player) sequential move game, when a second-mover with social preferences makes her decision after observing the action chosen by the first-mover, the second-mover's marginal rate of substitution between her income and the income of the first-mover depends on her emotional state toward the first-mover. Following this approach, and for simplicity we assume that the only emotional reaction that can be triggered in the CR game is the negative reaction toward a transgressing leader by the responders. Qualitatively, our main results hold for any social preferences models in which a social preference type beneficiary prefers that DAC be defeated.

If agent *i* is a Social Preferences type (hereafter the *SP*-type) and a responder, she regards a DAC transgression by the leader as undesirable, modeled with the utility function

$$U_i\left(y_L, y_i, y_j\right) = \left\{\frac{1}{\alpha}\left[y_i^{\alpha} + \theta y_L^{\alpha})\right], \quad \begin{matrix} \theta \in (-1,0) \text{ if } a_L \in \{TAB, TA, TB\} \\ \theta = 0 \text{ if } a_L = NT \end{matrix}\right. \qquad \text{(A1)}$$

Here, y_i is agent *i*'s income, y_L is the leader's income, y_j is the income of the other responder, θ is the (conditional) emotional state variable, and $\alpha \leq 1$ (and $\alpha \neq 0$) is an elasticity of substitution parameter. TAB denotes transgression against both responders, NT denotes No Transgression, and TA and TB denote divide-and-conquer transgression against A and B, respectively. If an agent is a Standard type (hereafter the *S*-type), then regardless of whether she is a leader or a responder, she has a utility function

$$U_i\left(y_i\right) = \frac{1}{\alpha} y_i^{\alpha}. \qquad \text{(A2)}$$

Because the only emotional reaction we focus on is the negative reaction by a responder towards a transgressing leader, an *SP*-type leader will also have a utility function of $U_i\left(y_i\right) = \frac{1}{\alpha} y_i^{\alpha}$.

Result 1: If $|\theta| > \left(9^{\alpha} - 7^{\alpha}\right)/8^{\alpha}$ and $\frac{7^{\alpha} - 9^{\alpha} - \theta 8^{\alpha}}{7^{\alpha} - 8^{\alpha} - \theta 8^{\alpha}} < p < \frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1}$, then each of the three following strategy profiles constitutes a Bayesian Nash Equilibrium in the DAC subgame with social preferences:

> (i) Both the *S*-type victim and the *SP*-type victim acquiesce, and both the *S*-type beneficiary and the *SP*-type beneficiary acquiesce.
> (ii) Both the *S*-type victim and the *SP*-type victim challenge. The *S*-type beneficiary always acquiesces, and the *SP*-type beneficiary challenges.
> (iii)The *S*-type victim challenges with a probability $\beta = \frac{(1-p)(9^{\alpha} + \theta 8^{\alpha} - 7^{\alpha}) + p(9^{\alpha} - 8^{\alpha})}{p(7^{\alpha} - 8^{\alpha} - \theta 8^{\alpha})} \in (0,1)$, and the *SP*-type victim always challenges. The *S*-type beneficiary always acquiesces, and the *SP*-type beneficiary challenges with a probability $\gamma = \frac{2^{\alpha} - 1}{(1-p)\left(7^{\alpha} - 1\right)} \in (0,1)$.

Result 1 shows that when social preferences are sufficiently strong and that there is a sufficiently high probability that a beneficiary is an *SP*-type, then social preferences transform the DAC subgame into a stag-hunt game for the responders, with multiple (and Pareto-ranked) equilibria. The model also implies that victims will challenge more than beneficiaries, which is consistent with the empirical findings in Cason and Mui (2007). Although joint resistance can be supported as an equilibrium, incomplete information about the types of other responders and multiple equilibria can prevent joint resistance from occurring.

1. **Candidate Equilibria**

We first explain the condition $|\theta| > \frac{9^{\alpha} - 7^{\alpha}}{8^{\alpha}}$. Note that an *S*-type agent will never challenge a DAC transgression when she is a beneficiary. If the social preferences of the *SP*-type are not too strong, then she will not be willing to challenge a DAC transgression when she is a beneficiary. This will be the case if $9^{\alpha} + \theta 8^{\alpha} \geq 7^{\alpha}$, or $|\theta| \leq \frac{9^{\alpha} - 7^{\alpha}}{8^{\alpha}}$ (assuming that an agent will not challenge if she is indifferent between challenging or acquiescing). When $|\theta| \leq \frac{9^{\alpha} - 7^{\alpha}}{8^{\alpha}}$, the DAC subgame has a unique equilibrium in which both types of victim and both types of beneficiary acquiesce, and NT cannot be supported as an equilibrium.

If $|\theta| > \frac{9^{\alpha} - 7^{\alpha}}{8^{\alpha}}$, then an *SP*-type agent will prefer to challenge a DAC transgression even when she is a beneficiary if both types of victim will challenge.

The following are candidate equilibria considered in Result 1 above:

(i) Both the *S*-type victim and the *SP*-type victim acquiesce, and both the *S*-type beneficiary and the *SP*-type beneficiary acquiesce.
(ii) Both the *S*-type victim and the *SP*-type victim challenge. An *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges.
(iii) An *S*-type victim challenges with a probability $\beta = \frac{(1-p)(9^\alpha + \theta 8^\alpha - 7^\alpha) + p(9^\alpha - 8^\alpha)}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)} \in (0,1)$, an *SP*-type victim always challenges. An *S*-type beneficiary always acquiesces, an *SP*-type beneficiary challenges with a probability $\gamma = \frac{2^\alpha - 1}{(1-p)(7^\alpha - 1)} \in (0,1)$.

We shall also consider additional candidate equilibria:

(ii)' When DAC transgression occurs, the *S*-type victim acquiesces and the *SP*-type victim challenges. An *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges.

(iii)' When DAC transgression occurs, an *S*-type victim acquiesces, an *SP*-type victim challenges with a probability $\tau = \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} \in (0,1)$. An *S*-type beneficiary always acquiesces, an *SP*-type beneficiary challenges with a probability $\gamma = \frac{2^\alpha - 1}{(1-p)(7^\alpha - 1 - \theta 8^\alpha)} \in (0,1)$.

2. Critical Probabilities

We first define the following critical probabilities. They are derived below in this section.

Define:

$$\hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} \qquad (1)$$

If $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, then the *S*-type victim will challenge if the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges.

Define:

$$\hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha} \qquad (2)$$

If $p < \hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$, then the *SP*-type victim will challenge if the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges.

Define:

$$\hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} \qquad (3)$$

If $p < \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, then the *SP*-type beneficiary will challenge if the *S*-type victim acquiesces and the *SP*-type victim challenges.

It is straightforward to verify that

$$\hat{p}_V^{S} = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} < \hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha} \qquad (4)$$

(4) reflects the fact that the *SP*-type victim has a stronger incentive to challenge than then *S*-type victim.

We, however, find that depending on the values of θ and α, we can have:

(I) $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} < \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$, or

(II) $\frac{7^\alpha - 2^\alpha}{7^\alpha - 1} < \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$, or

(III) $\frac{7^\alpha - 2^\alpha}{7^\alpha - 1} < \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha} < \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$.

A straightforward computational exercise shows that (I) is far more common than the other two cases, so we focus on (I) here. However, using the results below, one can derive (qualitatively similar) results for cases (II) and (III).

We first derive $\hat{p}_V^{S} = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$.

Suppose the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges. If the *S*-type victim acquiesces, she gets

$$U_V^S(AC) = \frac{1}{\alpha} p\left(2^\alpha\right) + \frac{1}{\alpha}(1-p)\left(2^\alpha\right) = \frac{1}{\alpha} 2^\alpha \qquad (5)$$

If the *S*-type victim challenges, she gets

$$U_V^S(CH) = \frac{1}{\alpha} p(1) + \frac{1}{\alpha}(1-p)\left(7^\alpha\right) \tag{6}$$

The *S*-type victim will challenge if

$$\frac{1}{\alpha} p(1) + \frac{1}{\alpha}(1-p)\left(7^\alpha\right) > \frac{1}{\alpha} 2^\alpha \Rightarrow$$

$$p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} \tag{7}$$

It is obvious that $0 < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} < 1$.

To derive $\hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$, suppose the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges. If the *SP*-type victim acquiesces, she gets

$$U_V^{SP}(AC) = \frac{1}{\alpha} p\left[2^\alpha + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left(2^\alpha + \theta 8^\alpha\right) = \frac{1}{\alpha}\left(2^\alpha + \theta 8^\alpha\right) \tag{8}$$

If she challenges, she gets

$$U_V^{SP}(CH) = \frac{1}{\alpha} p\left[1 + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left(7^\alpha\right) \tag{9}$$

The *SP*-type victim will challenge if

$$\frac{1}{\alpha} p\left[1 + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left(7^\alpha\right) > \frac{1}{\alpha}\left(2^\alpha + \theta 8^\alpha\right) \Rightarrow$$

$$p < \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha} \tag{10}$$

It is obvious that $0 < \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha} < 1$.

Now consider $\hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$. Suppose the *S*-type victim acquiesces and the *SP*-type victim challenges.

If the *SP*-type beneficiary acquiesces, she gets

$$U_B^{SP}(AC) = \frac{1}{\alpha} p\left[9^\alpha + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left(9^\alpha + \theta 8^\alpha\right) = \frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right) \tag{11}$$

If she challenges, she gets

$$U_B^{SP}(CH) = \frac{1}{\alpha} p\left[8^\alpha + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left(7^\alpha\right) \tag{12}$$

The *SP*-type beneficiary will challenge if

$$\frac{1}{\alpha} p\left[8^\alpha + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left(7^\alpha\right) > \frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right) \Rightarrow$$

$$p < \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} \tag{13}$$

It is obvious that $0 < \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < 1$.

3. Best Response Functions

We first state the relevant best response functions that will determine the equilibria. We then derive these best response functions below, later in this subsection.

First, observe that Acquiesce is the dominant strategy for the *S*-type beneficiary. Therefore, we do not discuss the no deviation conditions for the *S*-type beneficiary below.

Define

$$\hat{\gamma}^S = \frac{2^\alpha - 1}{(1-p)\left(7^\alpha - 1\right)} \tag{14}$$

If the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability γ, the *S*-type victim will

$$\begin{array}{l} \text{Challenge} \\ \text{Randomize between AC and CH} \\ \text{Acquiesce} \end{array} \text{ if } \gamma \begin{array}{c} > \\ = \\ < \end{array} \hat{\gamma}^{S} = \frac{2^{\alpha}-1}{(1-p)(7^{\alpha}-1)} \tag{15}$$

Define

$$\hat{\gamma}^{SP} = \frac{2^{\alpha}-1}{(1-p)(7^{\alpha}-1-\theta 8^{\alpha})} \tag{16}$$

If the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability γ, the *SP*-type victim will

$$\begin{array}{l} \text{Challenge} \\ \text{Randomize between AC and CH} \\ \text{Acquiesce} \end{array} \text{ if } \gamma \begin{array}{c} > \\ = \\ < \end{array} \hat{\gamma}^{SP} = \frac{2^{\alpha}-1}{(1-p)(7^{\alpha}-1-\theta 8^{\alpha})} \tag{17}$$

We shall show below that

$$\hat{\gamma}^{SP} < \hat{\gamma}^{S} \tag{18}$$

Define

$$\hat{\beta}^{SP} = \frac{(1-p)(9^{\alpha}+\theta 8^{\alpha}-7^{\alpha})+p(9^{\alpha}-8^{\alpha})}{p(7^{\alpha}-8^{\alpha}-\theta 8^{\alpha})} \tag{19}$$

If the *S*-type victim challenges with a probability β, and the *SP*-type victim challenges, the *SP*-type beneficiary will

$$\begin{array}{l} \text{Challenge} \\ \text{Randomize between AC and CH} \\ \text{Acquiesce} \end{array} \text{ if } \beta \begin{array}{c} > \\ = \\ < \end{array} \hat{\beta}^{SP} = \frac{(1-p)(9^{\alpha}+\theta 8^{\alpha}-7^{\alpha})+p(9^{\alpha}-8^{\alpha})}{p(7^{\alpha}-8^{\alpha}-\theta 8^{\alpha})} \tag{20}$$

Define

$$\hat{\tau}^{SP} = \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} \tag{21}$$

If the *S*-type victim acquiesces, and the *SP*-type victim challenges with a probability τ, the *SP*-type beneficiary will

$$\begin{array}{c} \text{Challenge} \\ \text{Randomize between AC and CH} \\ \text{Acquiesce} \end{array} \text{ if } \tau \begin{array}{c} > \\ = \\ < \end{array} \hat{\tau}^{SP} = \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} \tag{22}$$

We now derive (15). Suppose the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability γ. If the *S*-type victim acquiesces, she gets

$$U_V^S(AC) = p\left(\frac{1}{\alpha}2^\alpha\right) + (1-p)\left[\gamma\frac{1}{\alpha}2^\alpha + (1-\gamma)\frac{1}{\alpha}2^\alpha\right] = \frac{1}{\alpha}2^\alpha \tag{23}$$

If the *S*-type victim challenges, she gets

$$U_V^S(CH) = p\left(\frac{1}{\alpha}1\right) + (1-p)\left[\gamma\frac{1}{\alpha}7^\alpha + (1-\gamma)\frac{1}{\alpha}1\right] \tag{24}$$

The *S*-type victim will randomize if

$$\frac{1}{\alpha}2^\alpha = p\left(\frac{1}{\alpha}1\right) + (1-p)\left[\gamma\frac{1}{\alpha}7^\alpha + (1-\gamma)\frac{1}{\alpha}1\right] \Rightarrow$$

$$\gamma = \frac{2^\alpha - 1}{(1-p)7^\alpha - (1-p)} = \frac{2^\alpha - 1}{(1-p)(7^\alpha - 1)}$$

Define $\hat{\gamma}^S = \frac{2^\alpha - 1}{(1-p)(7^\alpha - 1)}$. It is obvious for any $p \in (0,1)$, $\hat{\gamma}^S > 0$. Furthermore, observe that $\hat{\gamma}^S < 1$ if and only if $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$. We conclude that:

For any $p \in (0,1)$, $\hat{\gamma}^S > 0$, and $\hat{\gamma}^S < 1$ if and only if $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ (25)

and (15) is the best response function for the *S*-type victim.

Note that because $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ implies $\hat{\gamma}^S < 1$, consistent with our discussion of (1), (15) implies that when $p < \hat{p}_V^S$, if the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability $\gamma = 1 > \hat{\gamma}^S = \frac{2^\alpha - 1}{(1-p)(7^\alpha - 1)}$, the *S*-type victim will challenge. Furthermore, because $0 < \hat{\gamma}^S = \frac{2^\alpha - 1}{(1-p)(7^\alpha - 1)} < 1$ when $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, when $p < \hat{p}_V^S$, the *S*-type victim will randomize when the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges with probability $\hat{\gamma}^S$. Finally, when $p \geq \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, $\hat{\gamma}^S \geq 1$. Therefore, (15) implies that if $p \geq \hat{p}_V^S$, with the *S*-type beneficiary acquiescing, even if the *SP*-type beneficiary challenges with a probability $\gamma = 1$, the *S*-type victim will acquiesce because $1 \leq \hat{\gamma}^S$.

Now consider (17). Suppose the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability γ. If the *SP*-type victim acquiesces, she gets

$$U_V^{SP}(AC) = \frac{1}{\alpha} p\left[2^\alpha + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left\{\gamma\left[2^\alpha + \theta 8^\alpha\right] + (1-\gamma)\left[2^\alpha + \theta 8^\alpha\right]\right\} \qquad (26)$$

If she challenges, she gets

$$U_V^{SP}(CH) = \frac{1}{\alpha} p\left[1 + \theta 8^\alpha\right] + \frac{1}{\alpha}(1-p)\left\{\gamma\left(7^\alpha\right) + (1-\gamma)\left[1^\alpha + \theta 8^\alpha\right]\right\} \qquad (27)$$

The *SP*-type victim randomizes if $U_V^{SP}(AC) = U_V^{SP}(CH)$, which implies

$$\gamma = \frac{2^\alpha - 1}{(1-p)\left(7^\alpha - 1 - \theta 8^\alpha\right)}.$$

Define $\hat{\gamma}^{SP} = \dfrac{2^{\alpha} - 1}{(1-p)\left(7^{\alpha} - 1 - \theta 8^{\alpha}\right)}$. Since $\theta < 0$, $2^{\alpha} - 1 > 0$, and $7^{\alpha} - 1 > 0$, $\hat{\gamma}^{SP} > 0$ for any *p*.

Furthermore, it is straightforward to observe that

$$\hat{\gamma}^{SP} = \frac{2^{\alpha} - 1}{(1-p)\left(7^{\alpha} - 1 - \theta 8^{\alpha}\right)} < \hat{\gamma}^{S} = \frac{2^{\alpha} - 1}{(1-p)\left(7^{\alpha} - 1\right)} \tag{28}$$

and that $\hat{\gamma}^{SP} < 1$ if and only if $p < \hat{p}_V^{SP} = \dfrac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{7^{\alpha} - 1 - \theta 8^{\alpha}}$.

Note that $\hat{\gamma}^{SP} < \hat{\gamma}^{S}$ is consistent with the intuition that the *SP*-type victim has a stronger incentive to challenge than the *S*-type victim. Therefore, if the *SP*-type beneficiary challenges with a probability $\hat{\gamma}^{S} = \dfrac{2^{\alpha} - 1}{(1-p)\left(7^{\alpha} - 1\right)} > \hat{\gamma}^{SP} = \dfrac{2^{\alpha} - 1}{(1-p)\left(7^{\alpha} - 1 - \theta 8^{\alpha}\right)}$, the *S*-type victim will just be willing to randomize between AC and CH, while the *SP*-type victim will strictly prefer to challenge.

(28) implies that the assumption that $p < \hat{p}_V^{S} = \dfrac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1}$ (which implies that $p < \hat{p}_V^{SP} = \dfrac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{7^{\alpha} - 1 - \theta 8^{\alpha}}$ by (4)), by guaranteeing that $\hat{\gamma}^{S} < 1$, will also be sufficient to ensure that $\hat{\gamma}^{SP} < 1$. We conclude that:

$$\text{For any } p \in (0,1),\ \hat{\gamma}^{SP} > 0, \text{ and } \hat{\gamma}^{SP} < 1 \text{ if and only if } p < \hat{p}_V^{SP} = \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{7^{\alpha} - 1 - \theta 8^{\alpha}} \tag{29}$$

and (17) is the best response function for the *SP*-type victim.

In summary, (17) has the following implications. Because $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} \Rightarrow p < \hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$, when $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, (29) implies that $\hat{\gamma}^{SP} < 1$. When $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, if the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability $\gamma = 1 > \hat{\gamma}^{SP}$, the *SP*-type victim will challenge. Furthermore, because $0 < \hat{\gamma}^{SP} = \frac{2^\alpha - 1}{(1-p)\left(7^\alpha - 1 - \theta 8^\alpha\right)} < 1$ when $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, when $p < \hat{p}_V^S$, the *SP*-type victim will randomize when the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges with probability $\hat{\gamma}^{SP}$. Finally, when $p \geq \hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$, $\hat{\gamma}^{SP} \geq 1$. Therefore, (17) implies that if $p \geq \hat{p}_V^{SP}$, with the *S*-type beneficiary acquiescing, even if the *SP*-type beneficiary challenges with a probability $\gamma = 1$, because $1 \leq \hat{\gamma}^{SP}$, the *SP*-type victim will acquiesce.

Now consider (20). If the *S*-type victim challenges with a probability β, and the *SP*-type victim always challenges, if the *SP*-type beneficiary acquiesces, she gets

$$U_B^{SP}(AC) = \frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right) \tag{30}$$

If she challenges, she gets

$$U_B^{SP}(CH) = \frac{1}{\alpha} p\left\{\beta(7^\alpha) + (1-\beta)\left[8^\alpha + \theta 8^\alpha\right]\right\} + \frac{1}{\alpha}(1-p)(7^\alpha) \tag{31}$$

The *SP*-type beneficiary will randomize if $U_B^{SP}(AC) = U_B^{SP}(CH)$, which implies

$$\beta = \frac{(1-p)(9^\alpha + \theta 8^\alpha - 7^\alpha) + p(9^\alpha - 8^\alpha)}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)}$$

Define

$$\hat{\beta}^{SP} = \frac{(1-p)(9^\alpha + \theta 8^\alpha - 7^\alpha) + p(9^\alpha - 8^\alpha)}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)} \quad (32)$$

Rewrite (32) as

$$\hat{\beta}^{SP} = 1 - \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)}. \quad (33)$$

Since $\theta < 0$ and $|\theta| > \frac{9^\alpha - 7^\alpha}{8^\alpha}$, $\left(7^\alpha - 8^\alpha - \theta 8^\alpha\right) > \left(7^\alpha - 9^\alpha - \theta 8^\alpha\right) > 0$, so $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} > 0$, and $\hat{\beta}^{SP} < 1$ for any $p > 0$. To ensure that $\hat{\beta}^{SP} > 0$, we need $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)} < 1$, which implies that

$$p > \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} \quad (34)$$

We conclude that:

For any $p \in (0,1)$, $\hat{\beta}^{SP} < 1$, and $\hat{\beta}^{SP} > 0$ if and only if $p > \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$ (35)

and (20) is the *SP*-type beneficiary's best response function when the *S*-type victim challenges with a probability β, and the *SP*-type victim always challenges.

The best response function (20) implies the following. When $p > \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, $0 < \hat{\beta}^{SP} = 1 - \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)} < 1$. When $p > \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, if the *S*-type victim challenges with a probability $\beta = 1 > \hat{\beta}^{SP}$, and the *SP*-type victim always challenges, the *SP*-type beneficiary will challenge. Furthermore, because $0 < \hat{\beta}^{SP} = 1 - \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)} < 1$ when $p > \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, when $p > \hat{p}_B^{SP}$, the *SP*-type beneficiary will randomize when the *S*-type victim challenges with probability $\hat{\beta}^{SP}$ and the *SP*-type victim always challenges. Finally, when $p \le \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, $\hat{\beta}^{SP} \le 0$. Therefore, (20) implies that if $p \le \hat{p}_B^{SP}$, with the *S*-type victim acquiescing—that is, challenging with a probability $\beta = 0 \le \hat{\beta}^{SP}$ --even if the *SP*-type victim always challenges, the *SP*-type beneficiary will acquiesce.

Now consider (22). Suppose the *S*-type victim acquiesces, and the *SP*-type victim challenges with probability τ. If the *SP*-type beneficiary acquiesces, she gets

$$U_B^{SP}(AC) = \frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right) \tag{36}$$

If she challenges, she gets

$$U_B^{SP}(CH) = \frac{1}{\alpha}(p)(8^\alpha + \theta 8^\alpha) + \frac{1}{\alpha}(1-p)\left\{\tau(7^\alpha) + (1-\tau)\left(8^\alpha + \theta 8^\alpha\right)\right\} \tag{37}$$

The *SP*-type beneficiary will randomize if $U_B^{SP}(AC) = U_B^{SP}(CH)$, which implies

$$\tau = \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)}$$

Define

$$\hat{\tau}^{SP} = \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} \tag{38}$$

Since $|\theta| > \frac{9^\alpha - 7^\alpha}{8^\alpha}$, $\hat{\tau}^{SP} > 0$ for any *p*. Furthermore, $\hat{\tau}^{SP} < 1 \Rightarrow \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} < 1$, which implies that

$$p < \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} \tag{39}$$

We conclude that:

For any $p \in (0,1), \hat{\tau}^{SP} > 0,$ and $\hat{\tau}^{SP} < 1$ if and only if $p < \hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$ (40)

and (22) is the *SP*-type beneficiary's best response function when the *S*-type victim acquiesces, and the *SP*-type victim challenges with a probability τ.

The best response function (22) implies the following. When $p < \hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, $0 < \hat{\tau}^{SP} = \dfrac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} < 1$. When $p < \hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, if the *S*-type victim always acquiesces, and the *SP*-type victim challenges with probability $\tau = 1 > \hat{\tau}^{SP}$, the *SP*-type beneficiary will challenge. Furthermore, because $0 < \hat{\tau}^{SP} = \dfrac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} < 1$ when $p < \hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, when $p < \hat{p}_B^{SP}$, the *SP*-type beneficiary will randomize when the *S*-type victim always acquiesces and the *SP*-type victim challenges with probability $\hat{\tau}^{SP} = \dfrac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)}$. Finally, when $p \geq \hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, $\hat{\tau}^{SP} \geq 1$. Therefore, (22) implies that if $p \geq \hat{p}_B^{SP} = \dfrac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, with the *S*-type victim acquiescing, even if the *SP*-type victim challenges with probability $\tau = 1$, the *SP*-type beneficiary will acquiesce because $\tau = 1 \leq \hat{\tau}^{SP}$.

4. Equilibria in the DAC Subgame

Figure 1 summarizes the equilibria of the **DAC sub**game as a function of the parameter values. We first note that for any values of *p*, (i)—the case in which both types of victim and both types of beneficiary acquiesce—is an equilibrium. It is obvious that if both types of beneficiary acquiesce, then the best response of both types of victim will be to acquiesce, and vice versa. [1]

[1] For example, consider the no deviation condition of the *S*-type victim. Formally, (15) implies that when $p < \hat{p}_V^S = \dfrac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, $\hat{\gamma}^S = \dfrac{2^\alpha - 1}{(1-p)(7^\alpha - 1)} \in (0,1)$. If the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary

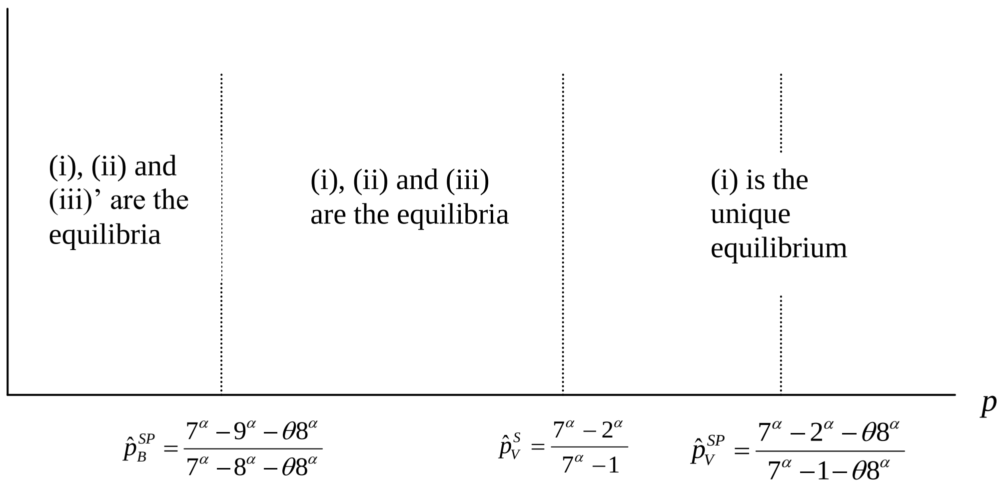


Figure 1. Equilibrium in the DAC Subgame as a Function of p

We first discuss the case of $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ --that is, the case covered by Result 1 in the text—and then briefly comment on the cases of $p \le \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{(7^\alpha - 8^\alpha - \theta 8^\alpha)}$ and $\frac{7^\alpha - 2^\alpha}{\left(7^\alpha - 1\right)} \le p$.

Consider candidate equilibria (ii) when $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$. In this case, (25) and (29) imply that $0 < \hat{\gamma}^{SP} < \hat{\gamma}^S < 1$, and (35) implies that $0 < \hat{\beta}^{SP} < 1$. When the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability $\gamma = 1 > \hat{\gamma}^S > \hat{\gamma}^{SP}$, (15) and (17) (and the discussion in the paragraph immediately below (25) and (29), respectively) imply that both the *S*-type victim and the *SP*-type victim will challenge. When the *S*-type victim challenges with a probability $\beta = 1 > \hat{\beta}^{SP}$, and the *SP*-type victim always challenges, (20) (and the discussion in the paragraph immediately below (35)) implies that the *SP*-type beneficiary will

---

challenges with a probability $\gamma = 0 < \hat{\gamma}^S = \frac{2^\alpha - 1}{\left(1 - p\right)\left(7^\alpha - 1\right)}$, (15) implies that the *S*-type victim will acquiesce. When $p \ge \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, $\hat{\gamma}^S = \frac{2^\alpha - 1}{\left(1 - p\right)\left(7^\alpha - 1\right)} \ge 1$. Therefore, if the *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges with a probability $\gamma = 0$, we again have $\gamma = 0 < \hat{\gamma}^S = \frac{2^\alpha - 1}{\left(1 - p\right)\left(7^\alpha - 1\right)}$ and the *S*-type victim will acquiesce. Similar arguments establish the cases regarding the no deviation conditions of the *SP*-type victim and the *SP*-type beneficiary.

challenge. Therefore, (ii) is an equilibrium. The result that the candidate equilibrium (iii) that involves randomization is also an equilibrium follows from (15), (17), and (20).

We have established that (i), (ii), and (iii) are equilibria when $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$. One can also show that there are no other equilibria in this case. First, using (15), (17), (20), (22), the fact that the *S*-type beneficiary has a dominant strategy to Acquiesce, and the observation that regardless of whether she is a victim or a beneficiary, an *SP*-type responder has a stronger incentive to challenge than her *S*-type counterpart, one can show that for any value of *p*, the only candidate equilibria in this one-shot CR game with social preferences are precisely (i), (ii), (iii), (ii)', and (iii)'. Since we now have $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, (15) (and the discussion in the paragraph immediately below (25)) implies that the *S*-type victim will challenge if the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges, hence (ii)' cannot be an equilibrium. Since $p > \hat{p}_B^{SP}$, (40) implies that $\hat{\tau}^{SP} = \frac{9^\alpha - 8^\alpha}{(1-p)(7^\alpha - 8^\alpha - \theta 8^\alpha)} > 1$, so (iii)' cannot be an equilibrium either. We therefore conclude that when $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha} < p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, (i), (ii), (iii) are the only equilibria and Result 1 is established.

Note that while the conditions of $|\theta| > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ and $\frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{(7^\alpha - 8^\alpha - \theta 8^\alpha)} < p < \frac{7^\alpha - 2^\alpha}{(7^\alpha - 1)}$ are needed for the hybrid strategy equilibrium in (iii), only the conditions $|\theta| > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ and $p < \frac{7^\alpha - 2^\alpha}{(7^\alpha - 1)}$ are needed for the pure strategy equilibrium in (ii), while (i)—the case of everyone acquiescing—is obviously an equilibrium for any p and θ.

Finally, we examine the incidence of challenge by the victim and the beneficiary. In the pure strategy equilibrium in (i), no one challenges the transgression. In the pure strategy equilibrium in (ii), the victim challenges with a probability one, while the beneficiary only challenges with a probability $(1-p)$. In the hybrid equilibrium in (iii), the victim challenges with a probability $p\beta + (1-p)$, while the beneficiary challenges with a probability $(1-p)\gamma$. Since $1 > (1-p)$ and $p\beta + (1-p) > (1-p)\gamma$, in both equilibria (ii) and (iii), the victim challenges with a higher probability than the beneficiary.

We now discuss briefly the cases of $p \le \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$ and $p \ge \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$.

When $p \ge \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$, (15) (and the discussion in the paragraph immediately below (25)) implies that if the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges, the *S*-type victim will acquiesce. Hence there exists no equilibrium in which the *S*-type victim challenges. However, as illustrated in Figure, I, because $p \geq \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} \Rightarrow p > \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, (22) (and the discussion in the paragraph immediately below (40)) implies that if the *S*-type victim acquiesces, even if the *SP*-type victim challenges, the *SP*-type beneficiary will acquiesce. Hence, there exists no equilibrium in which any challenge occurs and (i) is the unique equilibrium in this case.

When $p \leq \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, as shown in Figure I, we also have $p < \hat{p}_V^S = \frac{7^\alpha - 2^\alpha}{7^\alpha - 1} < \hat{p}_V^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{7^\alpha - 1 - \theta 8^\alpha}$. Therefore, (25) and (29) imply that $0 < \hat{\gamma}^S < \hat{\gamma}^{SP} < 1$, (35) implies that $\hat{\beta}^{SP} \leq 0$, and (40) implies that $0 < \hat{\tau}^{SP} < 1$. That (ii) and (iii)' are the equilibria follows from (15), (17), (20) and (22), and $0 < \hat{\gamma}^S < \hat{\gamma}^{SP} < 1$, $\hat{\beta}^{SP} \leq 0$, and $0 < \hat{\tau}^{SP} < 1$.[2] Since we now have $0 < \hat{\gamma}^S < 1$ as $p < \hat{p}_V^S$, if the *S*-type beneficiary acquiesces and the *SP*-type beneficiary challenges, (15) implies that the *S*-type victim will challenge instead of acquiesce. Hence, (ii)' cannot be an equilibrium. Finally, since we now have $\hat{\beta}^{SP} = \frac{(1-p)(9^\alpha + \theta 8^\alpha - 7^\alpha) + p(9^\alpha - 8^\alpha)}{p(7^\alpha - 8^\alpha - \theta 8^\alpha)} \leq 0$ when $p \leq \hat{p}_B^{SP} = \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, (iii) cannot be an equilibrium. Hence, when $p \leq \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{7^\alpha - 8^\alpha - \theta 8^\alpha}$, (i), (ii) and (iii)' are the only equilibria.

5. Supporting No Transgression as Equilibrium in the One-Shot CR Game with Social Preferences

Using Result 1, one can determine the set of all equilibria in the one-shot CR game with social preferences. For space considerations, we only illustrate here how No Transgression can be supported as an equilibrium.
Assume for simplicity that the leader will not transgress when indifferent between transgressing and not transgressing. In this case we have

> **Corollary 1.** If $p \leq \left(\frac{3}{4}\right)^\alpha$, then the following constitutes an equilibrium in the one-shot CR game with social preferences: (i) the leader chooses No Transgression (ii) Regardless of their type, both responders challenge if the leader chooses to transgress against both of them (iii) When DAC transgression occurs, both the *S*-type victim and the *SP*-type victim challenge.

---

[2] For example, to verify that the *SP*-type beneficiary will not deviate in (ii) when $p \leq \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{(7^\alpha - 8^\alpha - \theta 8^\alpha)}$, observe that because $\hat{\beta}^{SP} \leq 0$ when $p \leq \frac{7^\alpha - 9^\alpha - \theta 8^\alpha}{(7^\alpha - 8^\alpha - \theta 8^\alpha)}$, we have $\hat{\beta}^{SP} < 1$. Therefore, (20) implies that if the *S*-type victim challenges with probability $\beta = 1 > \hat{\beta}^{SP}$, and the *SP*-type victim challenges, then the *SP*-type beneficiary will challenge. Similar arguments establish the other no deviation conditions for the claim in the text.

An *S*-type beneficiary acquiesces, and the *SP*-type beneficiary challenges (iv) Regardless of their type, both responders acquiesce if the leader chooses No Transgression

It is obvious that the leader will not transgress against both responders given their strategies. If he practices DAC transgression, he receives $p\frac{1}{\alpha}(8)^{\alpha}$. If he chooses No Transgression, he receives $\frac{1}{\alpha}(6)^{\alpha}$. He will therefore choose No Transgression if $p\frac{1}{\alpha}(8)^{\alpha} \leq \frac{1}{\alpha}(6)^{\alpha}$, which implies $p \leq \left(\frac{3}{4}\right)^{\alpha}$.

Appendix D

Supporting the No Transgression Outcome in the Indefinitely Repeated CR Game with Social Preferences and Incomplete Information Using the Trigger Strategies in Weingast (1995, 1997)

Suppose that regardless of his/her type, a player plays the trigger strategy specified in Weingast (1995, 1997), which can support NT as equilibrium in the indefinitely repeated game with standard preferences as showed in Appendix A.

In the presence of private information about a player's type, instead of considering subgame perfect equilibrium, we consider Perfect Bayesian equilibrium, which requires that players update their beliefs according to Bayes' rule using other players' strategies on the equilibrium path. For this specific trigger strategy equilibrium, however, Bayesian updating does not have any bite since the strategy specifies that an *S*-type responder and an *SP*-type responder will adopt the same action for any history. We therefore specify that every player's posterior beliefs regarding the types of other players are the same as his/her prior beliefs for any history. This is consistent with Bayesian updating on the equilibrium path as players are always pooling, and Bayesian updating places no restrictions on beliefs off-the equilibrium path. Our goal is to simply demonstrate that No Transgression can still be supported as an equilibrium in the repeated CR game with social preferences. Because transgression never occurs in equilibrium, this trigger strategy equilibrium does not provide theoretical insight regarding whether victims will challenge more than the beneficiaries in the repeated game. In Appendix E, we construct another equilibrium in which signalling of types occurs in equilibrium and helps coordinate collective resistance. This equilibrium also provides useful guidance for statistical analysis of subjects' behavior because unlike the trigger strategy discussed here, DAC transgression and challenge can occur on the equilibrium path.

We now show that NT can be supported as equilibrium using the trigger strategies.

1. No deviation Conditions for the Responders

(1.1) There has been no acquiescence to earlier transgression and the leader chooses DAC transgression in the current period:

(1.1.1) No Deviation Condition for the *S*-type Beneficiary:

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{1}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{2}$$

$$\frac{1}{\alpha}\left[9^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right]\geq\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right]$$

$$\Rightarrow\delta\geq\frac{9^{\alpha}-7^{\alpha}}{\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)} \tag{3}$$

(1.1.2) No Deviation Condition for the *SP*-type Beneficiary:

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{4}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 12^{\alpha}\right)(\delta+\delta^{2}+.........)\right] \tag{5}$$

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right]\geq\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 12^{\alpha}\right)(\delta+\delta^{2}+.........)\right]$$

$$\Rightarrow\delta\geq\frac{9^{\alpha}-7^{\alpha}+\theta 8^{\alpha}}{\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)-\theta 12^{\alpha}} \tag{6}$$

If $|\theta|>\frac{9^{\alpha}-7^{\alpha}}{8^{\alpha}}$, then R.H.S. of (3) is negative and the *SP*-type beneficiary will not deviate even if she is completely impatient. When $|\theta|>\frac{9^{\alpha}-7^{\alpha}}{8^{\alpha}}$, the *SP*-type beneficiary prefers the transgression to fail and will challenge even in the one-shot game if both types of victim challenges (Appendix C). It is straightforward to observe that even if $|\theta|<\frac{9^{\alpha}-7^{\alpha}}{8^{\alpha}}$, $\frac{9^{\alpha}-7^{\alpha}+\theta 8^{\alpha}}{\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)-\theta 12^{\alpha}}\in\left(0,1\right)$ and hence, there always exists high enough discount factor that can deter the *SP*-type beneficiary from deviating even when her social preferences are "mild" so that she has a dominant strategy of acquiescing in the one-shot game. [3]

(1.1.3) No Deviation Condition for the *S*-type Victim:

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{7}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{8}$$

[3] For brevity, we usually ignore knife-edge conditions such as $|\theta|=\frac{9^{\alpha}-7^{\alpha}}{8^{\alpha}}$ throughout the appendix.

$\therefore No\,deviation$

(1.1.4) No Deviation Condition for the *SP*-type Victim:

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{9}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right] \tag{10}$$

$\therefore No\,deviation$

(1.2) There has been no acquiescence to earlier transgression and the leader chooses TAB transgression in the current period:

(1.2.1) No Deviation Condition for an *S*-type Responder:

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{11}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{12}$$

$\therefore No\,deviation$

(1.2.2) No Deviation Condition for an *SP*-type Responder:

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{13}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right] \tag{14}$$

$\therefore No\,deviation$

(1.3) There has been no acquiescence to earlier transgression and the leader chooses NT in the current period:

(1.3.1) No Deviation Condition for a Responder of both types:

No deviation - (AC)

$$\frac{1}{\alpha}\left[8^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{15}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \qquad (16)$$

$\therefore$ *No deviation*

(1.4) There has been acquiescence against earlier transgression and the leader plays TAB in the current period:

In this case, the leader is supposed to play TAB thereafter, while each type of responder is supposed to play AC thereafter.

If the leader plays TAB, the no deviation condition for an *S*-type responder will be $\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right]>\frac{1}{\alpha}\left[1+2^{\alpha}(\delta+\delta^{2}+.........)\right]$, which is satisfied.

If the leader plays TAB, the no deviation condition for an *SP*-type responder will be $\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right]>\frac{1}{\alpha}\left[1+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right]$, which is satisfied.

(1.5) There has been acquiescence against earlier transgression, and the leader has plays something other than TAB in the current period.

Consider first the case when the leader plays DAC transgression. Then the no deviation condition for the S-type beneficiary will be $\frac{1}{\alpha}\left[9^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right]>\frac{1}{\alpha}\left[8^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right]$, and the no deviation condition for the SP-type beneficiary will be $\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right]>\frac{1}{\alpha}\left[8^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right]$. Both conditions are satisfied.

When the leader plays DAC transgression, the no deviation condition for the S-type victim will be $\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}(\delta+\delta^{2}+.........)\right]>\frac{1}{\alpha}\left[1+2^{\alpha}(\delta+\delta^{2}+.........)\right]$, and the no deviation condition for the SP-type victim will be $\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right]>\frac{1}{\alpha}\left[1+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)(\delta+\delta^{2}+.........)\right]$. Both are satisfied.

Consider the case when the leader plays NT. Then the no deviation condition for the responder will be $8^{\alpha}+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)>7^{\alpha}+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)$, which is satisfied.

2. No deviation Conditions for the Leader:

(2.1) There has been no acquiescence against any transgression:

(2.1.1) No Deviation Condition for the leader:

No deviation (play *NT*): $\frac{1}{\alpha}\left[6^{\alpha}+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ (17)

$TAB:\frac{1}{\alpha}\left[0+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ (18)

$DAC:\frac{1}{\alpha}\left[0+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ (19)

$\therefore No\,deviation$

(2.2) There has been acquiescence against some transgression:

(2.2.1) No Deviation Condition for the leader:

No deviation (play *TAB*) : $\frac{1}{\alpha}\left[12^{\alpha}+12^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ (20)

$DAC:\frac{1}{\alpha}\left[8^{\alpha}+12^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ (21)

$NT:\frac{1}{\alpha}\left[6^{\alpha}+12^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ (22)

$\therefore No\,deviation$

In conclusion, because $\frac{9^{\alpha}-7^{\alpha}+\theta 8^{\alpha}}{\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)-\theta 12^{\alpha}}<\frac{9^{\alpha}-7^{\alpha}}{\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)}$, the critical discount factor that will support cooperation is $\delta^{*}\left(\alpha\right)=\frac{9^{\alpha}-7^{\alpha}}{\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)}$.

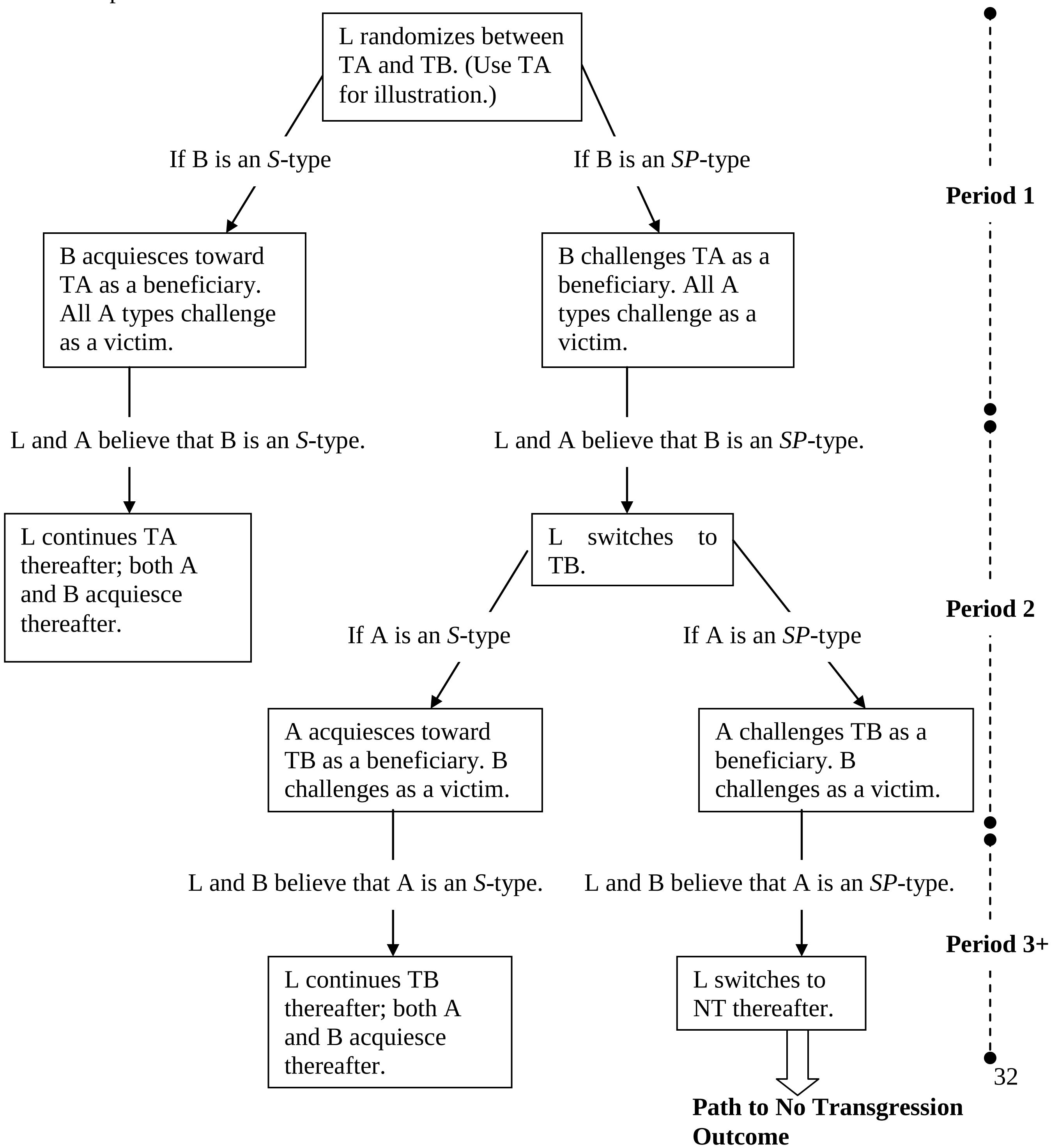
Recall that the continuation probability in the indefinite repetition treatments is $\frac{7}{8}$. When $\alpha=1$, $\delta^{*}\left(1\right)=\frac{2}{8}=\frac{1}{4}<\frac{7}{8}$. Furthermore, $\frac{d\delta^{*}\left(\alpha\right)}{d\alpha}=\frac{\left(8^{\alpha}-2^{\alpha}\right)\left(9^{\alpha}\ln 9-7^{\alpha}\ln 7\right)-\left(9^{\alpha}-7^{\alpha}\right)\left(8^{\alpha}\ln 8-2^{\alpha}\ln 2\right)}{\left[\left(8^{\alpha}-2^{\alpha}\right)+\left(9^{\alpha}-7^{\alpha}\right)\right]^{2}}$. This result does not help us sign $\frac{d\delta^{*}\left(\alpha\right)}{d\alpha}$. Computational results, however, show that for $\alpha\in\left(0,1\right)$, $\delta^{*}\left(\alpha\right)<0.25$, and

$\frac{d\delta^*(\alpha)}{d\alpha} < 0$. Hence, we conclude that no transgression can be supported as an equilibrium by the trigger strategy in the indefinite repeated CR game with social preferences.

Appendix E

An Equilibrium in the Indefinitely Repeated CR Game with Social Preferences and Incomplete Information in Which the Leader Punishes the Challenging Beneficiary

Result 2: If $\frac{9^\alpha-7^\alpha}{8^\alpha} < -\theta < \frac{\left(9^\alpha-7^\alpha\right)+\delta\left(9^\alpha-1\right)+\delta^2\left(9^\alpha-2^\alpha\right)}{8^\alpha}$ and $\frac{6^\alpha\left(1-\delta\right)}{8^\alpha-\delta 6^\alpha} < p <$ $\min\left(\frac{7^\alpha-2^\alpha}{7^\alpha-1}, \frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha+\delta^2 8^\alpha-9^\alpha-\theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha+\delta 8^\alpha-\left(1-\delta\right)\left(1+\theta 8^\alpha\right)-\delta\left(2^\alpha+\theta 8^\alpha\right)\right]}\right)$, then the following strategies constitute a Perfect Bayesian equilibrium in the indefinitely repeated CR game with social preferences:



Remark 1: Before proceeding with the proof, we provide interpretations of the conditions that ensure that Result 2 holds to convey the main intuition.

Result 2 states that in the indefinitely repeated CR game with social preferences and incomplete information, a separating equilibrium exists in which both DAC transgression and revelation of preference types occur in equilibrium.

The condition $-\theta > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ is the (now familiar) condition that says that an *SP*-type beneficiary prefers a DAC transgression to be defeated even though it increases her material payoff.

To interpret the condition $-\theta < \frac{\left(9^\alpha - 7^\alpha\right) + \delta\left(9^\alpha - 1\right) + \delta^2\left(9^\alpha - 2^\alpha\right)}{8^\alpha}$, observe that if the leader chooses TA in period 1, since both types of A (the victim) will challenge in period 1, by challenging in period 1 B (the beneficiary) will ensure successful joint resistance. This leads *SP*-type B to obtain a utility gain of $\frac{1}{\alpha}\left[7^\alpha - \left(9^\alpha + \theta 8^\alpha\right)\right]$. In equilibrium, B's challenge in period 1 will cause the leader to switch to TB in period 2. In period 2, the *SP*-type B will challenge as a victim. If, however, A is the *S*-type, then A will acquiesce as a beneficiary. In this case, B's material payoff is 1 and she gets a utility of $\frac{1}{\alpha}$. Furthermore, the leader will practice TB from period 3 onward and B will get a utility of $\frac{1}{\alpha}2^\alpha$ each period from period 3 onward. If B acquiesces in period 1, she will get the lower utility of $\frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right) < \frac{1}{\alpha}7^\alpha$ in period 1, but she ensures that she will get $\left(9^\alpha + \theta 8^\alpha\right)$ every subsequent period in this repeated game. Thus, if an *SP*-type B knows that A is an *S*-type for certain, then by choosing CH instead of AC in period 1, she gets a utility gain $\frac{1}{\alpha}\left[7^\alpha - \left(9^\alpha + \theta 8^\alpha\right)\right]$ in period 1, but will incur future utility losses with a present discounted value given by $\frac{1}{\alpha}\left[\delta\left(9^\alpha - 1\right) + \frac{\delta^2}{1-\delta}\left(9^\alpha - 7^\alpha\right)\right]$. This present discounted value reflects a utility loss of $\frac{1}{\alpha}\left[\left(9^\alpha + \theta 8^\alpha\right) - \left(1 + \theta 8^\alpha\right)\right] = \frac{1}{\alpha}\left(9^\alpha - 1\right)$ in period 2, and a utility loss of $\frac{1}{\alpha}\left[\left(9^\alpha + \theta 8^\alpha\right) - \left(2^\alpha + \theta 8^\alpha\right)\right] = \frac{1}{\alpha}\left(9^\alpha - 2^\alpha\right)$ each period from period 3 onward. The condition $-\theta < \frac{\left(9^\alpha - 7^\alpha\right) + \delta\left(9^\alpha - 1\right) + \delta^2\left(9^\alpha - 2^\alpha\right)}{8^\alpha}$ can be re-written as $\frac{1}{\alpha}\left[7^\alpha - \left(9^\alpha + \theta 8^\alpha\right)\right] < \frac{1}{\alpha}\left[\delta\left(9^\alpha - 1\right) + \frac{\delta^2}{1-\delta}\left(9^\alpha - 7^\alpha\right)\right]$, which simply says that if an *SP*-type B knows that A is an *S*-type for certain, then she will choose AC instead of CH in period 1, because the present discounted value of the future utility losses of being "perpetually trapped" as a victim starting from period 2 is larger than the one period gain of successful coordinated resistance in period 1. Of course, when the discount factor is $\delta = 0$, this condition will not be satisfied as the present discounted value of future utility losses is zero. This condition, however, will be satisfied for a large range of discount factors. That is, except for exceptionally low discount factors, when the leader practices the strategy of punishing the non-cooperating beneficiary and chooses TA in period 1, an *SP*-type B in period 1 will not challenge in period 1 if she knows that A is an *S*-type who will "sell her out" in the future and cause her to suffer as a victim of persistent DAC.

This observation also implies that if an *SP*-type B in period 1 believes that there is a sufficiently high probability that the victim is also an *SP*-type who will cooperate with her to challenge in future periods even when the leader practices TB, the *SP*-type B will challenge TA in period 1. By choosing CH instead of AC, she gets a current utility gain $\frac{1}{\alpha}\left[7^\alpha - \left(9^\alpha + \theta 8^\alpha\right)\right]$ in period 1. If A turns out to be an *SP*-type , she will also get a utility gain $\frac{1}{\alpha}\left[7^\alpha - \left(9^\alpha + \theta 8^\alpha\right)\right]$ in period 2 due to successful coordinated resistance with A, as well as a utility gain $\frac{1}{\alpha}\left[8^\alpha - \left(9^\alpha + \theta 8^\alpha\right)\right]$ each period from period 3 onward as the leader will be deterred from practicing any transgression from period 3 onward. If p is sufficiently small, then the current gain from successful coordination in the current period and the expected future gain due to successful coordinated resistance with an *SP*-type A will be larger than the future expected loss due to being trapped in persistent DAC against her due to the acquiescence by an *S*-type A. This condition is given by $p < \frac{(1-\delta)(1+\delta)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[(1-\delta)7^\alpha + \delta 8^\alpha - (1-\delta)\left(1+\theta 8^\alpha\right) - \delta\left(2^\alpha + \theta 8^\alpha\right)\right]}$.

When deciding whether to challenge TA as a victim in period 1, an *S*-type A will do so if and only if she expects that there is a high enough probability that B is an *SP*-type who will challenge. The condition $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ guarantees that this will be the case. Similarly, when deciding whether to challenge TA as a victim, an *SP*-type A will do so if and only if she expects that there is a high enough probability B is an *SP*-type who will challenge. Note that compared to an *S*-type A, an *SP*-type A suffers from the extra utility loss of $\theta 8^\alpha$ when the leader succeeds in practicing TA. Hence, an *SP*-type A has a stronger incentive to challenge than an *S*-type A, which means that the condition $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ is also sufficient to ensure that an *SP*-type A will challenge TA in period 1.

When deciding whether to practice DAC, the leader will do so if he believes that there is a high enough probability that the beneficiary is an *S*-type so that he has a high enough probability to succeed. This is given by the condition $p < \frac{6^{\alpha}(1-\delta)}{8^{\alpha}-\delta 6^{\alpha}}$. These observations explain the condition

$$\frac{6^{\alpha}(1-\delta)}{8^{\alpha}-\delta 6^{\alpha}} < p < \min\left(\frac{7^{\alpha}-2^{\alpha}}{7^{\alpha}-1}, \frac{(1-\delta)(1+\delta)7^{\alpha}+\delta^{2}8^{\alpha}-9^{\alpha}-\theta 8^{\alpha}}{\delta\left[(1-\delta)7^{\alpha}+\delta 8^{\alpha}-(1-\delta)(1+\theta 8^{\alpha})-\delta(2^{\alpha}+\theta 8^{\alpha})\right]}\right).$$

We now proceed to a formal discussion of Result 2, and then will present the complete proof.

1. Strategies

Consider the following strategy profile:

(1.1) The leader randomizes between TA and TB in the first period.

(1.2) If the leader has practiced a particular DAC transgression in the previous period, he continues to practice the same DAC (the “incumbent DAC”) in the current period unless this incumbent DAC has always been challenged by its beneficiary (for example, if the leader practiced TA in the previous period and B did not challenge TA, then he continues to practice TA in the current period). If the incumbent DAC transgression has always been challenged by its beneficiary, the leader switches to the “alternative” DAC in the current period (that is, punish the responder who challenged the incumbent DAC as a beneficiary) unless the alternative DAC has always been challenged by its beneficiary. If the alternative DAC has always been challenged by its beneficiary as well, switches to playing NT thereafter.

(1.3) If the leader has played NT in the previous period, and there exists a DAC transgression that has not been challenged by its beneficiary, he switches to such a DAC transgression in the current period regardless of the responders’ past actions. If both TA and TB had always been challenged by their respective beneficiaries, the leader plays NT thereafter.

(1.4) If the leader has played TAB in the previous period, then he continues to play TAB if any responder acquiesces in the previous period. Otherwise, he switches to DAC if there exists a DAC transgression that has not been challenged by its beneficiary before, or to NT if both TA and TB have always been challenged by their respective beneficiaries.

(1.5) An *S*-type B adopts the following strategy: Always acquiesce toward TA (as a beneficiary). If TB occurs for the first time in the current period and no TA has occurred before, she challenges. If TB occurs for the first time in the current period and TA occurred before, she acquiesces except when she herself has challenged all TA before (as a beneficiary). If TB occurs in the current period and TB occurred before, she challenges if and only both A and B have always challenged TB before.

(1.6) An *SP*-type B adopts the following strategy: If TA occurs for the first time in the current period and no TB has occurred before, she challenges. If TA occurs for the first time in the current period and TB occurred before, she challenges if and only if A has challenged all previous TB. If TA occurs in the current period and TA has occurred before, she challenges if and only TA has always been challenged by both A and B. If TB occurs for the first time in the current period and no TA occurred before, she challenges. If TB occurs for the first time in the current period and TA occurred before, she challenges if and only if B has challenged all TA before. If TB occurs in the current period and TB had occurred before, she challenges if and only both A and B have challenged all previous TB.

(1.7) Regardless of whether she is an *S*-type or an *SP*-type, B always acquiesces toward NT, but always challenges TAB.

(1.8) An *S*-type A adopts the mirror image of the *S*-type B's strategy, and an *SP*-type A adopts the mirror image of the *SP*-type B's strategy.

2. Beliefs

Perfect Bayesian equilibrium requires that along the equilibrium path, a player updates his/her belief about other players' types using the equilibrium strategies according to Bayes' rule. For illustration, consider some examples of how responder A updates her posterior beliefs about responder B.

Suppose the leader plays TA in the first period so that B is the beneficiary. Given that the *SP*-type B will challenge and the *S*-type B will acquiesce, if A observes that B challenges, A should conclude that B is an *SP*-type, but should conclude that B is an *S*-type if B acquiesces.

On the other hand, suppose that the leader plays TB in the first period so that B is the victim. Given that both types of B are supposed to challenge, observing that B challenges does not give A new information, so her posterior belief should be the same as her prior.

Although the responders are adopting a "separating" strategy, because this is a repeated game there are still various kinds of off equilibrium histories that need to be considered. For such histories, when applicable we shall choose posterior beliefs that capture the Intuitive Criterion (Cho and Kreps, 1989) in this context.[4] For an example in which the Intuitive Criterion imposes restrictions on beliefs, consider the case of (2.3) discussed in Section 2 below: in period *t*, the leader plays TA for the first time, and TB has occurred prior to period *t* and A has acquiesced toward TB in some period prior to period *t*. If B challenges TA in period *t*, then we shall assume that A believes that B is an *SP*-type. Because A has already revealed that she herself is an *S*-type, according to the equilibrium strategies, both the *S*-type B and the *SP*-type B should acquiesce (as a beneficiary) when TA occurs. Therefore, B challenging TA is a "probability zero" event that should not happen in equilibrium. Because an *S*-type B can never gain by challenging while an *SP*-type B can be better off if her

[4] As Govindan and Wilson (2009) describe succinctly, the Intuitive Criterion is satisfied "if no type of sender could obtain a payoff higher than his equilibrium payoff were he to choose a nonequilibrium message and the receiver responds with an action that is an optimal reply to a belief that imputes zero probability to Nature's choice of those types that cannot gain from such a deviation regardless of the receiver's play" (Govindan and Wilson, 2009, p. 3).

challenge eventually leads to the leader playing NT, we assume that A believes that B is an *SP*-type when such an off equilibrium event occurs.

Furthermore, for simplicity we shall assume that whenever a beneficiary of a DAC transgression has acquiesced when she is "supposed" to challenge, this will lead the other players to conclude that she is an *S*-type. Consider, for example, the history that in the first period, the leader plays TA, both A and B challenge, and then in period 2, the leader plays TA again, but this time B acquiesces. In this case, although A believes that B is an *SP*-type in the beginning of period 2, we shall assume that A concludes that B is an *S*-type at the beginning of period 3 upon observing B's acquiescence in period 2. Note that this history is off the equilibrium path because according to his equilibrium strategies, the leader should switch to playing TB in the second period after TA was challenged in the first period.

These are the key ideas employed to update players' beliefs when we check the no deviation conditions below. Let p_t^{ij} denote *i*'s posterior belief that *j* is the standard type at the beginning of period $t+1$; for example, p_t^{AB} is A's belief that B is the *S*-type at the beginning of period *t*. Because we focus on pure strategies, p_t^{AB} can only take on the values of 0, *p*, and 1 (recall that *p* is the unconditional probability that an agent is an *S*-type). For illustration, we state explicitly below in (2.1)-(2.10) the rules that describe how responder A updates her posterior beliefs regarding the type of responder B for various histories. Responder B updates her beliefs regarding responder A's type using the mirror image of these rules, and the leader uses similar rules to update his beliefs about the types of each responder.

(2.1) Suppose that in period *t*, the leader plays TA for the first time, and TB has never occurred prior to period *t*. If B challenges TA in period *t*, then $p_t^{AB} = 0$. If B acquiesces, then $p_t = 1$.

(2.2) Suppose that in period *t*, the leader plays TA for the first time, and TB has occurred prior to period *t* and A has always challenged TB previously: If B challenges TA in period *t* then $p_t^{AB} = 0$. If B acquiesces, then $p_t = 1$.

(2.3) Suppose that in period *t*, the leader plays TA for the first time, and TB has occurred prior to period *t* and A has acquiesced toward TB in some period prior to period *t*: If B challenges TA in period *t*, then $p_t^{AB} = 0$. If B acquiesces, then $p_t^{AB} = 1$.

(2.4) Suppose that in period *t*, the leader plays TA, and the leader has played TA prior to period *t*, and both A and B have challenged all TA before. If B challenges TA in period *t*, then $p_t^{AB} = 0$. If B acquiesces, then $p_t^{AB} = 1$.

(2.5) Suppose that in period *t*, the leader plays TA, and the leader has played TA prior to period *t*, and B has acquiesced toward TA in some period prior to period *t*: Regardless of whether B challenges or acquiesces, $p_t^{AB} = 1$.

(2.6) Suppose that in period *t*, the leader plays TA, and the leader has played TA prior to period *t*, and A has acquiesced toward TA in some period prior to period *t*, while B has challenged all TA before: If B challenges TA in period *t*, then $p_t^{AB} = 0$. If B acquiesces, then $p_t^{AB} = 0$.

(2.7) Suppose that the leader plays TB in period *t* for the first time, and TA never occurred prior to period *t*: Regardless of whether B challenges or acquiesces, $p_t^{AB} = p$.

(2.8) Suppose that the leader plays TB in period *t* for the first time, and TA has occurred prior to period *t* and B has always challenged TA previously: If B challenges TB in period *t* (as a victim), then $p_t^{AB} = 0$. If B acquiesces, then $p_t^{AB} = 1$.

(2.9) Suppose that the leader plays TB in period *t* for the first time, and TA has occurred prior to period *t* and B has acquiesced toward TA in in some period prior to *t*: Regardless of whether B challenges or acquiesces, $p_t^{AB} = 1$.

(2.10) Suppose that the leader plays either NT or TAB, then regardless of whether B challenges or acquiesces, $p_t^{AB} = p_{t-1}^{AB}$.

We now verify that no player wants to deviate given others' strategies and her posterior beliefs that are updated according to the rules discussed above. We first check the no deviation conditions of the responders, and then discuss the no deviation conditions of the leader. For earlier cases, after stating the equations describing the relevant payoffs, we provide detailed comments that explain how these payoffs are determined for the interested reader . For brevity, we do not provide such comments for all cases.

3. No Deviation Conditions of the Responders

We first define the following critical probabilities. They will be derived below in this section.

Define:

$$p_{BR}^{SP} = \frac{(1-\delta)(1+\delta)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[(1-\delta)7^\alpha + \delta 8^\alpha - (1-\delta)\left(1+\theta 8^\alpha\right) - \delta\left(2^\alpha + \theta 8^\alpha\right)\right]} \qquad (1)$$

When $p < \hat{p}_{BR}^{SP}$, if a DAC occurs in the first period and its beneficiary is an *SP*-type, then this *SP*-type beneficiary will be willing to take the risk to challenge and reveal her type. This is because there is a sufficiently high probability (as $(1-p)$ is sufficiently large) that the other responder will also be an *SP*-type, so that A and B can succeed in coordinating collective resistance and, after enduring some DAC transgression, can deter the leader from future transgression.

Define:

$$p_{VR}^{S} = \frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1} \tag{2}$$

When $p < \hat{p}_{R}^{S}$, if a DAC occurs in the first period, then even if its victim is an *S*-type, the victim will be willing to challenge because there is a sufficiently high probability that the beneficiary is an *SP*-type.

Define:

$$p_{VR}^{SP} = \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{\left(7^{\alpha} - 1\right) - \theta 8^{\alpha}} \tag{3}$$

When $p < p_{VR}^{SP}$, if an DAC occurs in the first period, then an *SP*-type victim will be willing to challenge because there is a sufficiently high probability that the beneficiary is an *SP*-type.

Define:

$$p_{LR} = \frac{6^{\alpha} - \delta 6^{\alpha}}{8^{\alpha} - \delta 6^{\alpha}} \tag{4}$$

When $p > p_{LR} = \frac{6^{\alpha}\left(1-\delta\right)}{8^{\alpha} - \delta 6^{\alpha}}$, then the leader will randomize between TA and TB in the first period

3.1 TA Occurs in the First period

3.1.1 No Deviation Condition for the *S*-type B (Who is the Beneficiary)

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha} + 9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{5}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left\{7^{\alpha}+p\delta\left[1+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\delta\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\} \tag{6}$$

The *S*-type B will not deviate because $(5)>(6)$. The payoffs in (5) and (6) are determined as follows.

Explanations for (5): If the *S*-type B acquiesces toward TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}9^{\alpha}$ in period 1, and reveals that she is the *S*-type. The leader will continue to practice TA in period 2, both A and B will acquiesce, and from period 2 onward the players will be playing (TA, AC, AC) (this means that the leader plays TA, responder A plays AC, and responder B plays AC) every period, and B will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ in every period starting from period 2. This explains why playing AC in period 1 generates the (discounted intertemporal) payoff of (5) for the *S*-type B. Note that (5) is the highest possible payoff that the *S*-type B can get. Hence, she cannot gain from deviating. Nevertheless, we now explain how (6) is determined.

Explanations for (6): If the *S*-type B deviates to challenge TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 1, and both the leader and A will believe that she is the *SP*-type. The leader will switch to practice TB in period 2, and the *S*-type B will challenge as a victim. If A is the *S*-type, A will acquiesce in period 2 (so that B will get a payoff of $\frac{1}{\alpha}1^{\alpha}=\frac{1}{\alpha}\cdot 1$ in period 2), which will lead both the leader and B conclude that A is an *S*-type. The leader now believes that A is the *S*-type and B is the *SP*-type at the beginning of period 3, and will continue to practice TB in period 3, and the players will be playing (TB, AC, AC) thereafter. This implies that starting from period 3, the *S*-type B will get a payoff of $\frac{1}{\alpha}2^{\alpha}$ every period. On the other hand, if A is the *SP*-type, A will challenge in period 2 (so that B will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2), which will lead both the leader and B conclude that A is an *SP*-type. The leader now believes that both A and B are the *SP*-type at the beginning of period 3, and will switch to play NT in period 3. Starting from period 3, the players will be playing (NT, AC, AC) thereafter, and the *S*-type B will get a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period starting from period 3. This explains why playing CH in period 1 generates the (discounted intertemporal) payoff of (6) for the *S*-type B.

Summary: Acquiescence in the current period will ensure that the *S*-type B gets the highest payoff $\frac{1}{\alpha}9^{\alpha}$ in the current period, as well as in every period thereafter so she cannot gain from deviating. In fact, deviating to challenging TA in the current period is costly in the current period for the *S*-type B, and will cause the *S*-type B to always get a payoff lower than $\frac{1}{\alpha}9^{\alpha}$ in the second period regardless of the type of A. Furthermore, if A is the *S*-type, challenging TA in the first period can cause B to suffer from perpetual DAC transgression against her starting from period 2. On the other hand, if A is the *SP*-type, challenging TA in the first period can cause the leader to play NT in period 3 and thereafter. Since deviating to challenging in the first period involves a current loss, and will always lead to future losses, the *S*-type B will never want to deviate.

3.1.2 No Deviation Condition for the *SP*-type B (Who is the Beneficiary)

No deviation - (CH), she gets

$$\frac{1}{\alpha}\left\{7^{\alpha}+p\delta\left[\left(1+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+\left(1-p\right)\delta\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\} \qquad (7)$$

Deviation - (AC), she gets

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \qquad (8)$$

The *SP*-type B will not deviate iff $(7)>(8)$.

Define

$$f\left(p\right)=\frac{1}{\alpha}\left\{7^{\alpha}+p\delta\left[\left(1+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+\left(1-p\right)\delta\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}-\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \qquad (9)$$

This is the (expected) intertemporal utility difference between choosing CH and AC in period 1 for the *SP*-type B when facing TA. Note that

$$\begin{aligned} f\left(0\right)&=\frac{1}{\alpha}\left\{7^{\alpha}+\delta\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}-\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \\ &=\frac{1}{\alpha}\left\{\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]+\delta\left[\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]+\left(\frac{\delta}{1-\delta}\right)\left[8^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]\right]\right\}>0 \end{aligned} \qquad (10)$$

because $\frac{9^{\alpha}-7^{\alpha}}{8^{\alpha}}<-\theta$. Equation (10) reflects the following observation discussed in the beginning of this Appendix E: if an *SP*-type B in period 1 believes that there is a sufficiently high probability that the victim is also an *SP*-type who will cooperate with her to challenge in future periods when the leader practices TB, the *SP*-type B will challenge TA in period 1. By choosing CH instead of AC, she gets a current utility gain $\frac{1}{\alpha}\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]$ in period 1. If A turns out to be an *SP*-type , she will also get a utility gain $\frac{1}{\alpha}\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]$ in period 2 due to successful coordinated resistance with A, as well as a utility gain $\frac{1}{\alpha}\left[8^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]$ each period from period 3 onward as the leader will be deterred from practicing any transgression from period 3 onward.

On the other hand,

$$\begin{aligned} f(1) &= \frac{1}{\alpha}\left\{7^{\alpha}+\delta\left[\left(1+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]\right\}-\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \\ &= \frac{1}{\alpha}\left\{\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]+\delta\left[\left[\left(1+\theta 8^{\alpha}\right)-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]+\left(\frac{\delta}{1-\delta}\right)\left[\left(2^{\alpha}+\theta 8^{\alpha}\right)-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]\right]\right\}<0 \end{aligned} \tag{11}$$

because $-\theta<\frac{\left(9^{\alpha}-7^{\alpha}\right)+\delta\left(9^{\alpha}-1\right)+\delta^{2}\left(9^{\alpha}-2^{\alpha}\right)}{8^{\alpha}}$. Equation (11) reflects the following observation discussed at length in the beginning of this Appendix E: if an *SP*-type B knows that A is an *S*-type for certain, then she will choose AC instead of CH in period 1, because the present discounted value of the future utility losses of being "perpetually trapped" as a victim starting from period 2 is larger than the one period gain of successful coordinated resistance in period 1.

Furthermore,

$$f^{'}(p)=\frac{1}{\alpha}\delta\left\{\left[\left(1+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]-\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}<0 \tag{12}$$

Equation (12) reflects the fact that from the *SP*-type beneficiary's perspective, the higher the probability that the victim is an *S*-type, the more attractive that AC will be compared to CH. Since $f(1)<0<f(0)$, by the Intermediate Value Theorem, there exists $p\in(0,1)$ such that $f(p)=0$. Furthermore, because $f^{'}(p)<0$, p must be unique. Suppose not, so that there exist $p_1, p_2\in(0,1)$, $p_1<p_2$, $f(p_1)=f(p_2)=0$. Because $f^{'}(p)<0$, however, we must have $f(p_2)<f(p_1)==0$, which contradicts with $f(p_2)=0$. From (10), setting $f(p)=0$, simple algebra shows that $p=\frac{(1-\delta)(1+\delta)7^{\alpha}+\delta^{2}8^{\alpha}-9^{\alpha}-\theta 8^{\alpha}}{\delta\left[(1-\delta)7^{\alpha}+\delta 8^{\alpha}-(1-\delta)\left(1+\theta 8^{\alpha}\right)-\delta\left(2^{\alpha}+\theta 8^{\alpha}\right)\right]}$

Define:

$$p_{BR}^{SP} = p = \frac{(1-\delta)(1+\delta)7^{\alpha} + \delta^2 8^{\alpha} - 9^{\alpha} - \theta 8^{\alpha}}{\delta\left[(1-\delta)7^{\alpha} + \delta 8^{\alpha} - (1-\delta)\left(1+\theta 8^{\alpha}\right) - \delta\left(2^{\alpha} + \theta 8^{\alpha}\right)\right]} \tag{13}$$

When $p < \hat{p}_{BR}^{SP}$, if a DAC occurrs in the first period and its beneficiary is an *SP*-type, then this *SP*-type beneficiary will be willing to take the risk to challenge and reveal her type. This is because there is a sufficiently high probability (as $(1-p)$ is sufficiently large) that the other responder will also be an *SP*-type, so that A and B can succeed in coordinating collective resistance and, after enduring some DAC transgression, can deter the leader from future transgression.

In defining critical probabilities in this appendix, the superscript refers to the type of the agent (*SP* or *S*), the first subscript refers to the role (beneficiary or victim), and the second subscript *R* refers to the fact that these are critical probabilities for the repeated game.

Explanations for (7): If the *SP*-type B challenges TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}7^{\alpha}$ in the current period, and reveals that she is the *SP*-type. The leader will switch to playing TB in period 2, and the *SP*-type B will challenge as a victim. If A is the *S*-type, A will acquiesce in period 2 (so that B will get a payoff of $\frac{1}{\alpha}\left(1+\theta 8^{\alpha}\right)$ in period 2), which will lead both the leader and B conclude that A is an *S*-type. At the beginning of period 3, the leader now believes that A is the *S*-type and B is the *SP*-type, and he will continue to practice TB in period 3, and the players will be playing (TB, AC, AC) thereafter. The *SP*-type B will therefore be “trapped” in the outcome of (TB, AC, AC) with the payoff of $\left(2^{\alpha} + \theta 8^{\alpha}\right)$ every period starting from period 3. On the other hand, if A is the *SP*-type, A will challenge in period 2 (so that B will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2), which will lead both the leader and B conclude that A is an *SP*-type. At the beginning of period 3, the leader now believes that both A and B are the *SP*-type, and he will switch to playing NT in period 3. Starting from period 3, the players will be playing (NT, AC, AC) thereafter, and the *SP*-type B will get a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period.

Explanations for (8): If the *SP*-type B deviates and acquiesces toward TA as a beneficiary in period 1, she gets a payoff of $\frac{1}{\alpha}\left(9^{\alpha} + \theta 8^{\alpha}\right)$ in period 1, and both the leader and responder A now believe that B is the *S*-type. The leader will continue to practice TA in period 2, both A and B will acquiesce, and from period 2 onward the players will be playing (TA, AC, AC) every period. By deviating to acquiesce toward TA in period 1, the *SP*-type B induces the leader and responder A to believe that she is an *S*-type, and the *SP*-type B will receive a payoff of $\frac{1}{\alpha}\left(9^{\alpha} + \theta 8^{\alpha}\right)$ in every period thereafter.

Summary: The act of challenging TA in the first period is a *risky* action for the *SP*-type B, and when the probability that A is an *S*-type is lower it is less risky—and hence more attractive—for the *SP*-type B to challenge TA. Note that when $|\theta| > \left(9^\alpha - 7^\alpha\right)/8^\alpha$, the *SP*-type B's social preferences are sufficiently strong so that she prefers successful coordinated resistance against TA with a utility of $\frac{1}{\alpha}7^\alpha$ to getting the highest material payoff of 9 with a utility of $\frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right)$. Since A will always challenge TA as a victim regardless of her type in the current period, comparing to acquiescing, challenging TA in period 1 always brings an increase in utility from $\frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right)$ to $\frac{1}{\alpha}7^\alpha$ in the current period. If A is an *S*-type, however, challenging TA in period 1 will cause the *SP*-type B to get a utility of $\frac{1}{\alpha}\left(1 + \theta 8^\alpha\right)$ in the second period, which is lower than $\frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right)$, and will also cause her to be trapped in the outcome of (TB, AC, AC) with a utility of $\frac{1}{\alpha}\left(2^\alpha + \theta 8^\alpha\right) < \frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right)$ every period starting from period 3. On the other hand, if A is an *SP*-type, challenging TA in period 1 will ensure that in the second period, the *SP*-type B gets a utility of $\frac{1}{\alpha}7^\alpha$--which is higher than $\frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right)$. Furthermore, the successful coordinated resistance with the *SP*-type A in both periods 1 and 2--by revealing that both A and B are the *SP*-type--will deter the leader from practicing any transgression starting from period 3, and will give the *SP*-type B a utility of $\frac{1}{\alpha}8^\alpha > \frac{1}{\alpha}\left(9^\alpha + \theta 8^\alpha\right)$ every period starting from period 3. Equation (13) defines the critical probability that will ensure that the *SP*-type B is willing to take the risk to challenge TA in the first period.

3.1.3 No Deviation Condition for the *S*-type A (Who is the Victim)

No deviation - (CH)

$$\frac{1}{\alpha}\left\{p\left[1^\alpha + 2^\alpha\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\left[7^\alpha + 9^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\} \qquad (14)$$

Deviation - (AC)

$$\frac{1}{\alpha}\left\{p\left[2^\alpha + 2^\alpha\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\left[2^\alpha + 9^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\} \qquad (15)$$

No deviation iff $(14) > (15)$, which can be simplified to $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$

Define:

$$p_{VR}^{S}=\frac{7^{\alpha}-2^{\alpha}}{7^{\alpha}-1} \tag{16}$$

When $p<\hat{p}_{R}^{S}$, if a DAC occurs in the first period, then even if its victim is an *S*-type, the victim will be willing to challenge because there is a sufficiently high probability that the beneficiary is an *SP*-type. Note that $p_{VR}^{S}=\frac{7^{\alpha}-2^{\alpha}}{\left(7^{\alpha}-1\right)}$ has the same value as p_{V}^{S}, defined in Appendix C. We shall explain this observation below.

<u>Explanations for (14)</u>: Suppose the *S*-type A challenges TA as a victim. If B is the *S*-type, B will acquiesce, and the *S*-type A gets a payoff of $\frac{1}{\alpha}\cdot 1$ in period 1, and both the leader and A will believe that B is the *S*-type. The leader will continue to practice TA in period 2, and both A and B will acquiesce. In this case, starting from period 2, the outcome will be (TA, AC, AC), and the *S*-type A will get a payoff of $\frac{1}{\alpha}2^{\alpha}$ every period starting from period 2. If B is the *SP*-type, B will challenge in period 1, and the *S*-type A will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 1. B's challenge will lead both the leader and A conclude that B is an *SP*-type. The leader will switch to play TB in period 2, B will challenge as a victim, but the *S*-type A will acquiesce as a beneficiary, and will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ in period 2. At the beginning of period 3, the leader believes that A is the *S*-type and B is the *SP*-type, and starting from period 3, the players will be playing (TB, AC, AC) thereafter. That is, starting from period 2, the *S*-type A will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ every period.

<u>Explanations for (15)</u>: Now suppose the *S*-type A acquiesces toward TA as a victim in period 1. If B is the *S*-type, B will acquiesce, and the *S*-type A will get a payoff of $\frac{1}{\alpha}2^{\alpha}$ in period 1, and both the leader and A will believe that B is the *S*-type. The leader will continue to practice TA in period 2, and both A and B will acquiesce. In this case, starting from period 2, the outcome will be (TA, AC, AC) every period, and the *S*-type A will get a payoff of $\frac{1}{\alpha}2^{\alpha}$ every period. If B is the *SP*-type, B will challenge in period 1, but the *S*-type A will still get a payoff of $\frac{1}{\alpha}2^{\alpha}$ in period 1, since the *S*-type A acquiesces. B's challenge will lead both the leader and A conclude that B is an *SP*-type. The leader will switch to playing TB in period 2, B will challenge as a victim, but the *S*-type A will acquiesce as a beneficiary, and will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ in period 2. At the beginning of period 3, the leader believes that A is the *S*-type and B is the *SP*-type, and

starting from period 3, the players will be playing (TB, AC, AC) thereafter and the *S*-type A will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ every period.

<u>Summary</u>: When TA occurs in the first period, the leader's decision regarding whether to continue to practice TA or switch to TB in the second period depends on whether B--the beneficiary of TA in the first period--challenges. In this repeated game equilibrium, given the leader's strategy, B's strategy, and A's own strategy, the *S*-type A's expected intertemporal payoff starting from the second period is essentially determined by B's type, and is not affected by the *S*-type A's choice between acquiescing or challenging in the first period. Therefore, in deciding whether to challenge as a victim in the first period, the *S*-type A is comparing the expected payoff of challenging in the first period with the expected payoff of acquiescing in the first period. This decision problem is similar to the one she faces when TA occurs in the one-shot CR game with social preferences and she expects that the *SP*-type B will challenge as a beneficiary while the *S*-type B will acquiesce. Therefore, the *S*-type A will challenge iff $p < \frac{7^{\alpha} - 2^{\alpha}}{\left(7^{\alpha} - 1\right)}$, which explain why $p_{VR}^{S} = \frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1} = p_{V}^{S}$

3.1.4 No Deviation Condition for the *SP*-type A (Who is the Victim)

No deviation - (CH)

$$\frac{1}{\alpha}\left\{ p\left[1 + \theta 8^{\alpha} + \left(2^{\alpha} + \theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] + \left(1-p\right)\left[7^{\alpha} + \delta 7^{\alpha} + 8^{\alpha}\left(\frac{\delta^{2}}{1-\delta}\right)\right]\right\} \tag{17}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left\{ p\left[2^{\alpha} + \theta 8^{\alpha} + \left(2^{\alpha} + \theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] + \left(1-p\right)\left[2^{\alpha} + \theta 8^{\alpha} + \delta 7^{\alpha} + 8^{\alpha}\left(\frac{\delta^{2}}{1-\delta}\right)\right]\right\} \tag{18}$$

No deviation iff $(17) > (18)$, which can be simplified to $p < \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{\left(7^{\alpha} - 1\right) - \theta 8^{\alpha}}$.

Define:

$$p_{VR}^{SP} = \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{\left(7^{\alpha} - 1\right) - \theta 8^{\alpha}} \tag{19}$$

The *SP*-type A will not deviate as a victim if $p < p_{VR}^{SP} = \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{\left(7^{\alpha} - 1\right) - \theta 8^{\alpha}}$.

Explanation for (17): Suppose the *SP*-type A challenges TA as a victim. If B is the *S*-type, B will acquiesce, and the *SP*-type A gets a payoff of $\frac{1}{\alpha}\left(1+\theta 8^{\alpha}\right)$ in period 1, and both the leader and A will believe that B is the *S*-type. The leader will continue to practice TA in period 2, and both A and B will acquiesce. Starting from period 2, the outcome will be (TA, AC, AC), and the *S*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ every period. If B is the *SP*-type, B will challenge in period 1, and the *SP*-type A will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 1. B's challenge will lead both the leader and A conclude that B is an *SP*-type. The leader will switch to play TB in period 2, B will challenge as a victim, and the *SP*-type A will challenge as a beneficiary, and the *SP*-type A will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2. At the beginning of period 3, the leader believes that both A and B are the *SP*-type, and the leader will switch to play NT in period 3. Starting from period 3, the players will be playing (NT, AC, AC) thereafter, which will give the *SP*-type a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period.

Explanation for (18): Suppose the *SP*-type A acquiesces toward TA as a victim in period 1. If B is the *S*-type, B will acquiesce, and the *S*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ in period 1, and both the leader and A will believe that B is the *S*-type. The leader will continue to practice TA in period 2, and both A and B will acquiesce. Starting from period 2, the outcome will be (TA, AC, AC), and the *SP*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ every period. If B is the *SP*-type, B will challenge in period 1, but the *SP*-type A will still get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ in period 1 since she acquiesces. Both the leader and A now believe that B is an *SP*-type. The leader will switch to play TB in period 2, B will challenge as a victim, and the *SP*-type A will challenge as a beneficiary, so that the *SP*-type A will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2. At the beginning of period 3, the leader believes that both A and B are the *SP*-type, and the leader will switch to play NT in period 3. Starting from period 3, the players will be playing (NT, AC, AC) thereafter, which will give the *SP*-type A a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period.

Summary: When TA occurs in the first period, the leader's decision regarding whether to continue to practice TA or switch to TB in the second period depends on whether B--the beneficiary of TA in the first period--challenges. As in (3.1.3), in this repeated game equilibrium, given the leader's strategy, B's strategy, and A's own strategy, the *SP*-type A's expected intertemporal payoff starting from the second period is essentially determined by B's type, and is not affected by the *SP*-type A's choice between acquiescing or challenging in the first period. Therefore, in deciding whether to challenge as a victim in the first period, the *SP*-type A is comparing the expected payoff of challenging in the first period with the expected payoff of acquiescing in the first period. This decision problem is similar to the one she faces when TA occurs in the one-shot CR game with social preferences and she expects that the *SP*-type B will challenge as a beneficiary while the *S*-type B will acquiesce. Therefore, the *SP*-type A will challenge iff $p < \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{\left(7^\alpha - 1\right) - \theta 8^\alpha}$, which explain why $p_{VR}^{SP} = \frac{7^\alpha - 2^\alpha - \theta 8^\alpha}{\left(7^\alpha - 1\right) - \theta 8^\alpha} = p_V^{SP}$.

The comments above explain how to determine a responder's payoff for different courses of action, taking into account how her decision can affect others' beliefs and future plays in this repeated CR game with incomplete information about social preferences. Similar explanations can be provided using the belief updating rules and equilibrium strategies for each sub-case considered below. To limit the size of this already lengthy Appendix, we shall only provide explanations for a selective number of cases.

3.2 TA Occurs for the First Time in period *t* > 1 and No TB Occurred Before

The decision problems for each type of responder at this off equilibrium information set will be the same as those faced in 3.1, so the no deviation conditions will hold.

3.3 No deviation Conditions in period *t* When TA Occurs for the First Time in Period *t* and TB Occurred Prior to Period *t*

3.3.1 No Deviation Condition for the *S*-type B (Who is the Beneficiary)

In this case, because an *S*-type A and *SP*-type A choose different actions when TB occurred prior to Period *t*, B has learned the type of A and hence, depending on her posterior belief, her no deviation conditions are as follows:

3.3.1.1 No Deviation Condition for the *S*-type B (Who is the Beneficiary) When A is the *S*-type

In this case, Nature chooses A to be the *S*-type, B observes that A acquiesces toward TB prior to period *t*, and believes that *A* is the *S*-type, that is, $p_t^{BA} = 1$. Note that this is an information set that is off the equilibrium path: Given that A acquiesces toward TB prior to period *t*, in equilibrium, the leader will continue to practice TB and TA should not be observed in period *t*.

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^\alpha + 9^\alpha\left(\frac{\delta}{1-\delta}\right)\right] \tag{20}$$

The *S*-type B will not deviate because (20) is the highest payoff that the *S*-type B can earn in this repeated game.

3.3.1.2 No Deviation Condition for the *S*-type B (Who is the Beneficiary) When A is the *SP*-type

B observes that A challenges TB prior to period *t*, and believes that *A* is the *SP*-type.

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{21}$$

Again, (21) is the highest payoff that the *S*-type B can earn in this repeated game and she cannot gain from deviating.

3.3.2 No Deviation Condition for the *SP*-type B (Who is the Beneficiary)

3.3.2.1 No Deviation Condition for the *SP*-type B (Who is the Beneficiary) When A is the *S*-type

B observes that A acquiesces toward TB prior to period *t*, and believes that *A* is the *S*-type. This is another information set that is off the equilibrium path: Given that A acquiesces toward TB prior to period *t*, in equilibrium, the leader will continue to practice TB and TA should not be observed in period *t*.

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{22}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[8^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{23}$$

The *SP*-type B will not deviate because $(22)>(23)$.

3.3.2.2 No Deviation Condition for the *SP*-type B (Who is the Beneficiary) When A is the *SP*-type

B observes that A challenges TB, and believes that *A* is the *SP*-type.

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{24}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{25}$$

The *SP*-type B will not deviate because $(24) > (25)$.

3.3.3 No Deviation Condition for the *S*-type A (Who is the Victim)

In this case, because an *S*-type B and *SP*-type B choose the same action (as a victim) when TB occurred prior to period *t*, A's posterior belief will be the same as her prior belief.

No deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{26}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[1+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{27}$$

The *S*-type A will not deviate because $(26) > (27)$.

3.3.4 No Deviation Condition for the *SP*-type A (Who is the Victim)

Because an *S*-type B and *SP*-type B choose the same action (as a victim) when TB occurred prior to period *t*, A's posterior belief regarding B's type and L's posterior belief will be the same as their prior beliefs. However, because the *SP*-type A challenged TB prior to period *t*, both B and L believe that A is the *SP*-type.

No deviation - (CH)

$$\frac{1}{\alpha}\left\{p\left[1+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta^{2}}{1-\delta}\right)\right]\right\} \tag{28}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left\{p\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[2^{\alpha}+\theta 8^{\alpha}+8^{\alpha}\left(\frac{\delta^{2}}{1-\delta}\right)\right]\right\} \tag{29}$$

Explanations for (28): Suppose the *SP*-type A challenges TA as a victim in period t. If B is the *S*-type, B will acquiesce, and the *SP*-type A will get a payoff of $\frac{1}{\alpha}\left(1+\theta 8^{\alpha}\right)$ in period *t*, and both the leader and A will believe that B is the *S*-type. The leader will continue to practice TA in period $t+1$, and both A and B will acquiesce. Starting from period $t+1$, the outcome will be (TA, AC, AC), and the *SP*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ every period. If B is the *SP*-type, B will challenge in period t, and the *SP*-type A will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period t. At the beginning of period $t+1$, the leader believes that both A and B are the *SP*-type, and the leader will switch to playing NT in period $t+1$. Starting from period $t+1$, the players will be playing (NT, AC, AC) thereafter, which will give the *SP*-type A a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period.

Explanations for (29): Suppose the *SP*-type A acquiesces toward TA as a victim in period t. If B is the *S*-type, B will acquiesce, and the *SP*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ in period t, and both the leader and A will believe that B is the *S*-type. The leader will continue to practice TA in period $t+1$, and both A and B will acquiesce. Starting from period $t+1$, the outcome will be (TA, AC, AC), and the *SP*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ every period. If B is the *SP*-type, B will challenge in period t, but the *SP*-type A will get a payoff of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ in period t since she acquiesces. At the beginning of period $t+1$, the leader believes that both A and B are the *SP*-type, and the leader will switch to play NT in period $t+1$. Starting from period $t+1$, the players will be playing (NT, AC, AC) thereafter, which will give the *SP*-type A a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period starting from period $t+1$.

Summary: In this case, the decision problem faced by the *SP*-type A as a victim in period *t* regarding whether to challenge is similar to the decision problem she faces in the one-shot game when she expects that the *SP*-type B will challenge as a beneficiary while the *S*-type B will acquiesce. The *SP*-type A will challenge if $p<p_{VR}^{SP}=\frac{7^{\alpha}-2^{\alpha}-\theta 8^{\alpha}}{\left(7^{\alpha}-1\right)-\theta 8^{\alpha}}=p_{V}^{SP}$.

3.4 No deviation Conditions in period *t* When TA Occurs in Period *t*, TA Occurred Prior to Period *t* and No TB Occurred Prior to *t*

3.4.1 No Deviation Condition for the *S*-type B (who is the Beneficiary)

In this case, because an *S*-type A and *SP*-type A choose the same action (as a victim) when TA occurred prior to period *t*, B's posterior belief will be the same as her prior belief. Because playing AC will give the *S*-type B the highest possible payoff of $\frac{1}{\alpha}\left[9^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$, she will not deviate.

3.4.2 No Deviation Condition for the *SP*-type B (Who is the Beneficiary)

Because an *S*-type A and *SP*-type A choose the same action (as a victim) when TA occurred prior to period *t*, B's posterior belief will be the same as her prior belief. Note that this is another off equilibrium information set: in equilibrium, the leader will not continue to practice TA in period *t* given that B has challenged TA prior to period *t*.

No deviation (CH)

$$\frac{1}{\alpha}\left\{7^{\alpha}+p\delta\left[\left(1+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+\left(1-p\right)\delta\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\} \qquad (30)$$

Deviation - (AC), she gets

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \qquad (31)$$

This no deviation condition for the *SP*-type B is the same as the one in (3.1.2), which will hold when $p<\hat{p}_{BR}^{SP}=\frac{\left(1-\delta\right)\left(1+\delta\right)7^{\alpha}+\delta^{2}8^{\alpha}-9^{\alpha}-\theta 8^{\alpha}}{\delta\left[\left(1-\delta\right)7^{\alpha}+\delta 8^{\alpha}-\left(1-\delta\right)\left(1+\theta 8^{\alpha}\right)-\delta\left(2^{\alpha}+\theta 8^{\alpha}\right)\right]}$.

3.4.3 No Deviation Condition for the *S*-type A (Who is the Victim)

Because an *S*-type B and *SP*-type B choose different actions when TA occurred prior to period *t*, A has learned the type of B and hence, depending on her posterior belief, her no deviation conditions are as follows:

3.4.3.1 No Deviation Condition for the *S*-type A (Who is the Victim) When B is the *S*-type

In this case, A observes that B acquiesced toward TA, and believes that B is the *S*-type.

No deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \qquad (32)$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[1+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{33}$$

The *S*-type A will not deviate because $(32) > (33)$.

3.4.3.2 No Deviation Condition for the *S*-type A (Who is the Victim) When B is the *SP*-type

In this case, A observed that B challenged TA, and believes that B is the *SP*-type. Note that this is another off equilibrium information set: in equilibrium, the leader will not continue to practice TA in period *t* given that B has challenged TA prior to period *t*.

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{34}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{35}$$

The *S*-type A will not deviate because $(34) > (35)$.

3.4.4 No Deviation Condition for the *SP*-type A (Who is the Victim)

Because an *S*-type B and *SP*-type B choose different actions when TA occurred earlier, A has learned the type of B and hence, depending on her posterior belief, her no deviation conditions are as follows:

3.4.4.1 No Deviation Condition for the *SP*-type A (Who is the Victim) When B is the *S*-type

A observes that B acquiesced toward TA, and believes that *B* is the *S*-type.

No deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{36}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[1+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{37}$$

The *SP*-type A will not deviate because $(36) > (37)$.

3.4.4.2 No Deviation Condition for the *SP*-type A (Who is the Victim) When B is the *SP*-type

In this case, A observed that B challenged TA, and believes that *B* is the *SP*-type. Note that this is another off equilibrium information set: in equilibrium, the leader will not continue to practice TA in period *t* given that B has challenged TA prior to period *t*.

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+7^{\alpha}\delta+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{38}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+7^{\alpha}\delta+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{39}$$

The *SP*-type A will not deviate because $(38) > (39)$.

3.5 TA Occurs in Period *t*, TA Also Occurred in a Period $k < t$, the First time TB Occurred Was in a Period $h < k$, and TA Did Not Occur Prior to Period *h*

3.5.1 No Deviation Condition for the *S*-type B (Who is the Beneficiary)

In this case, because an *S*-type A and *SP*-type B choose different actions when TB occurred prior to period *t*, B has learned the type of A and hence, depending on her posterior belief, her no deviation conditions are as follows:

3.5.1.1 When A is the *S*-type

In this case, nature chooses A to be the *S*-type, B observes that A acquiesces toward TB at *h*, and learned that *A* is the *S*-type, that is, $p_t^{BA}=1$. Note that this is an information set that is off the equilibrium path: Given that A acquiesces toward TB prior to period *t*, in equilibrium, the leader will continue to practice TB and the history described in (3.5) should not occur.

If the *S*-type B acquiesces, she get her highest possible payoff $\frac{1}{\alpha}\left[9^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$ so she will not deviate.

3.5.1.2 When A is an *SP*-type

B observes that A challenges TB at *h*, and believes that *A* is the *SP*-type. Again, the *S*-type B will not deviate because she gets the highest possible payoff by acquiescing.

3.5.2 No Deviation Condition for the *SP*-type B (Who is the Beneficiary)

3.5.2.1 When A is an *S*-type

B observes that A acquiesces toward TB prior to period *t*, and believes that A is the *S*-type.

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{40}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[8^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{41}$$

The *SP*-type B will not deviate because $(40)>(41)$.

3.5.2.2 When A is an *SP*-type

B observes that A challenges TB, and believes that *A* is the *SP*-type.

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{42}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{43}$$

The *SP*-type B will not deviate because $(42)>(43)$. By deviating, B will trigger the leader to practice TA from period $t+1$ onward, with no prospect of successful challenge by the responders. This is worse for her than getting the No Transgression outcome thereafter given that she is the *SP*-type.

3.5.3 No Deviation Condition for the *S*-type A (Who is the Victim)

In this case, when TB occurred in period *h*, both the *S*-type B and the *SP*-type B challenge (as a victim), and the *S*-type A responder acquiesces (as the beneficiary). This implies that when TA occurs in period *k*, both the *S*-type B and the *SP*-type B acquiesce (as the beneficiary). Because both types of B choose the same action when TB occurred in period *h* and when TA occurred in period *k*, A learns nothing about B's type.[5] A's posterior belief regarding B's type will be the same as her prior belief. This is an information set that is off the equilibrium path: Given that A acquiesces toward TB in period *h*, in equilibrium, the leader will continue to practice TB and the history described in (3.5) should not occur.

No deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \qquad (44)$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[1+2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \qquad (45)$$

The *S*-type A will not deviate because $(44)>(45)$.

3.5.4 No Deviation Condition for the *SP*-type A (Who is the Victim)

In this case, when TB occurred in period *h*, both the *S*-type B and the *SP*-type B challenge as a victim, and the *SP*-type A responder challenges (as the beneficiary). This implies that when TA occurs in period *k*, the *S*-type B acquiesce and the *SP*-type B challenges (as the beneficiary). Because an *S*-type B and *SP*-type B choose different actions when TA occurred in period *k*, A has learned the type of B. Depending on her posterior belief, her no deviation conditions are as follows.

3.5.4.1 When B is the *S*-type

In this case, A observed that B acquiesced toward TA in period *h*, and believes that *B* is the *S*-type.

No deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \qquad (46)$$

[5] If the leader chooses TAB or NT prior to period *k*, both types of B will choose the same action as well so A again learns nothing from those observed actions.

Deviation - (CH)

$$\frac{1}{\alpha}\left[1+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{47}$$

The *SP*-type A will not deviate because $(46) > (47)$.

3.5.4.2 When B is the *SP*-type

In this case, A observed that B challenged TB in period *h*, and believe that B is the *SP*-type. This is another information set that is off the equilibrium path: Given that both A and B are the *SP*-type, they both challenged TB in period *h*, and they both challenge TA in period *k*. The successful collective resistance against TB in period *h* and against TA in period *k*, respectively, would have induced the leader to switch to playing NT thereafter starting from period *k* +1. Thus, the history described in (3.5) should not occur.

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{48}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{49}$$

The *SP*-type A will not deviate because $(48) > (49)$.

One can verify similarly that the responders will not deviate for other histories when DAC transgression occurs in the current period, which will feature no deviation conditions similar to those considered above. For example, one can consider the case of TB occurring in the first period (which is simply the mirror image of (3.1) and hence is satisfied), or other off - equilibrium histories, such as TA occurs in period *t*, TB occurred in a period $k < t$ the first time, and TA occurred a period $h < k$ the last time. For brevity, we do not discuss these other cases here.

We now comment briefly on the simpler cases when either NT or TAB occurs in the current period. For illustration, consider the following:

3.6 NT Occurs in Period *t*, Both TA and TB Occurred Prior to Period *t* and Both TA and TB Have Always Been Challenged by their Respective Beneficiaries

3.6.1 No Deviation Conditions for both Responder Types

No deviation - (AC)

$$\frac{1}{\alpha}\left[8^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{50}$$

Deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+8^{\alpha}(\delta+\delta^{2}+.........)\right] \tag{51}$$

Since $(50)>(51)$, the responder will not deviate.

It is straightforward to verify that the responders will not deviate for other histories when NT or TAB occurs in the current period, such as the history of NT occurring in period *t*, acquiescence by A towards TB occurred prior to period *t*, and TA had never occurred; or the history of TAB occurring in period *t*, both TA and TB had occurred prior to period *t*, and both TA and TB have always been challenged by their respective beneficiaries.

4. No Deviation Conditions for the Leader (L)

4.1 L Will Randomize between TA and TB in the First Period

We now show that if $p>p_{LR}=\dfrac{6^{\alpha}(1-\delta)}{8^{\alpha}-\delta 6^{\alpha}}\in(0,1)$, then the leader will randomize between TA and TB in the first period

If L plays TA, he gets

$$\frac{1}{\alpha}\left\{p\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left\{0+p\delta\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\delta\left[0+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}\right\} \tag{52}$$

Let

$$X(p,\alpha,\delta)$$
$$=\left\{p\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left\{0+p\delta\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\delta\left[0+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}\right\} \tag{53}$$

It is easy to see that if L plays TB in the first period, he gets the same expected payoff as in (52), so L will be indifferent between TA and TB.

If L plays TAB in the first period, he gets 0 in the first period as TAB will be challenged by both A and B. He will switch to play TA or TB in the second period and get an expected payoff given by (52). So playing TAB in the first period is dominated by either playing either TA or TB in the first period.

If he plays NT in the first period, he gets $\frac{1}{\alpha}6^{\alpha}$ in the first period. He will then switch to playing TA or TB in the second period, and will get an expected payoff given by (52). That is, if L deviates to play NT in the first period, he gets

$$\frac{1}{\alpha}6^{\alpha} + \frac{1}{\alpha}\delta X$$

He will NOT deviate iff $\frac{1}{\alpha}6^{\alpha} + \frac{1}{\alpha}\delta X < \frac{1}{\alpha}X \Leftrightarrow$

$$(1-\delta)X > 6^{\alpha} \tag{54}$$

Consider the equation $(1-\delta)X = 6^{\alpha}$, which is a quadratic equation in p given by:

$$(1-\delta)\left\{p\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left\{0+p\delta\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\delta\left[0+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}\right\}=6^{\alpha} \tag{55}$$

Simplifying, (55) becomes

$$\delta\left(\delta 6^{\alpha}-8^{\alpha}\right)p^{2}+\left[(1+\delta)8^{\alpha}-2\delta^{2}6^{\alpha}\right]p+6^{\alpha}\left(\delta^{2}-1\right)=0 \tag{56}$$

(56) can be re-written as

$$\left[\left(\delta 6^{\alpha}-8^{\alpha}\right)p-6^{\alpha}(\delta-1)\right]\left[\delta p-(\delta+1)\right]=0 \tag{57}$$

This equation has two distinct roots,

$$p_1=\frac{6^{\alpha}(1-\delta)}{8^{\alpha}-\delta 6^{\alpha}}=\frac{6^{\alpha}-\delta 6^{\alpha}}{8^{\alpha}-\delta 6^{\alpha}}=p_{LR} \text{ and } p_2=\frac{\delta+1}{\delta}=1+\frac{1}{\delta}.$$

Define:

$$p_{LR}=\frac{6^{\alpha}-\delta 6^{\alpha}}{8^{\alpha}-\delta 6^{\alpha}} \tag{58}$$

It is obvious that $0 < p_1 < 1$, and $2 < p_2$. Furthermore, note that

$$(1-\delta)X(p,\alpha,\delta)-6^{\alpha}=\left[\left(\delta 6^{\alpha}-8^{\alpha}\right)p-6^{\alpha}(\delta-1)\right]\left[\delta p-(\delta+1)\right]$$

For any $p \in (p_1, p_2)$, $\left[\left(\delta 6^\alpha - 8^\alpha\right) p - 6^\alpha (\delta - 1)\right] < 0$, $\left[\delta p - (\delta + 1)\right] < 0$, so that $(1-\delta) X(p,\alpha,\delta) - 6^\alpha = \left[\left(\delta 6^\alpha - 8^\alpha\right) p - 6^\alpha (\delta - 1)\right]\left[\delta p - (\delta + 1)\right] > 0$. This implies that for any $p \in (p_1, 1]$, $(1-\delta) X(p,\alpha,\delta) - 6^\alpha > 0$. Therefore, the condition $p > p_{LR} = \dfrac{6^\alpha (1-\delta)}{\left(8^\alpha - \delta 6^\alpha\right)}$ will be sufficient to guarantee that L will randomize between TA and TB in the first period.

4.2 L Played TA for the First Time in Period *t*-1, B Did Not Challenge TA in Period *t*-1, and TB Never Occurred Before

In this case, L believes that B is an *S*-type. Consider the leader's decision in period *t*. If he plays TA, he gets $\frac{1}{\alpha} 8^\alpha \left(\frac{1}{1-\delta}\right)$. This is the highest possible payoff he can get, so he cannot gain by deviating.

4.3 L Played TA in Period *t*-1, B Challenged TA in Period *t*-1, and TB Never Occurred Before

In this case, L believes that B is the *SP*-type.

If he plays TB (no deviation), he gets

$$\frac{1}{\alpha}\left\{p\left[8^\alpha + 8^\alpha\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\left[0 + 6^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\} \tag{59}$$

If he plays TA, he gets

$$\begin{aligned}&\frac{1}{\alpha}\left\{p\left[0 + 8^\alpha\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\left[0 + \delta\left[0 + 6^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right]\right\}\\ &= \frac{1}{\alpha}\left\{0 + \delta\left\{p\left[8^\alpha + 8^\alpha\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\left[0 + 6^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\}\right\}\end{aligned} \tag{60}$$

If he plays TAB, he gets

$$\frac{1}{\alpha}\left\{0 + \delta\left\{p\left[8^\alpha + 8^\alpha\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\left[0 + 6^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\}\right\} \tag{61}$$

If he plays NT, he gets

$$\frac{1}{\alpha}\left\{6^{\alpha}+\delta\left\{p\left[8^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[0+6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}\right\}$$
$$=p\left(\frac{8^{\alpha}}{1-\delta}\right)+(1-p)6^{\alpha}\left(\frac{\delta}{1-\delta}\right) \tag{62}$$

It is obvious that $(59)>(60)$ and $(59)>(61)$ so that L will never deviate to TA or TAB. L will prefer TB to NT if and only if

$$\frac{1}{\alpha}6^{\alpha}+\frac{1}{\alpha}\delta\left[p\left(\frac{8^{\alpha}}{1-\delta}\right)+(1-p)6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]<\frac{1}{\alpha}\left[p\left(\frac{8^{\alpha}}{1-\delta}\right)+(1-p)6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\Leftrightarrow$$
$$(1-\delta)\left[p\left(\frac{8^{\alpha}}{1-\delta}\right)+(1-p)6^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]>6^{\alpha}.$$

Simplifying, this implies that $p>p_{LR}=\dfrac{6^{\alpha}(1-\delta)}{(8^{\alpha}-\delta 6^{\alpha})}$. This is the same condition derived in (4.1.1).

4.4 L Played TA for the First Time in Period *t*-1, B Challenged TA in Period *t*-1, TB occurred in a period *k* < *t* -1 and TB Had Been Challenged by A

In this case, L believes that both A and B are the *SP*-type.

If he plays NT (no deviation), he gets

$$\frac{1}{\alpha}6^{\alpha}\left(\frac{1}{1-\delta}\right) \tag{63}$$

If he plays TA, TB, or TAB he gets

$$\frac{1}{\alpha}\left[0+6^{\alpha}\left(\frac{1}{1-\delta}\right)\right] \tag{64}$$

Therefore, L will not deviate.

4.5 L Played TA for the First Time in Period *t*-1, B Challenged TA in Period *t*-1, TB Occurred in a Period *k* < *t* -1 and Had Not Been Challenged by A

This is an off-equilibrium history. If A has acquiesced toward TB in a period $k<t-1$, A has revealed to both B and L that she is the *S*-type, and L should have continued to practice TB and this history would not have occurred. At this off-equilibrium information set, L believes that A is the *S*-type, and that B is the *SP*-type (see (2.3) in section 2).

If L plays TB (no deviation), he gets $\frac{1}{\alpha}8^{\alpha}\left(\frac{1}{1-\delta}\right)$. This is the highest possible payoff L can get, so he cannot gain by deviating.

One can further check that the leader will not deviate for other histories when DAC transgression occurred in the previous period, which will feature no deviation conditions similar to those considered above. For example, one can consider the no deviation condition in period t when L played TA in period $t-1$ and in period $k<t-1$, B acquiesced in period $t-1$, and TB never occurred prior to period $t-1$. One can also consider other off-equilibrium history such as L played TA in period $t-1$ as well as in period $k<t-1$, and all TA had been challenged by B, and TB occurred in a period $h<k$ the last time and TB had been challenged by A

We now discuss briefly histories that involve L playing either NT or TAB in the previous period.

4.6 L Played NT in Period *t*-1, and There Exists a DAC Transgression That Had Occurred and Was Not Challenged by Its Beneficiary

For illustration, supposed TA had occurred and had not been challenged by B.

In this case, L believes that B is the *S*-type.

If L plays TA, he gets $\frac{1}{\alpha}8^{\alpha}\left(\frac{1}{1-\delta}\right)$. This is the highest possible payoff L can get, so he cannot gain by deviating.

4.7 L Played NT in Period *t*-1, and Both DAC Transgression Had Occurred and Had Always Been Challenged by Their Respective Beneficiaries

In this case, L believes that both A and B are the *SP*-type. So he should not deviate.

Formally, if L plays NT (no deviation), he gets

$$\frac{1}{\alpha}6^{\alpha}\left(\frac{1}{1-\delta}\right) \tag{65}$$

If he plays TA, TB, or TAB he gets

$$\frac{1}{\alpha}\left[0+6^{\alpha}\left(\frac{1}{1-\delta}\right)\right] \tag{66}$$

Therefore, L will not deviate.

4.8 L Played TAB in Period *t*-1, and There Exists a DAC Transgression That Had Occurred and Was Not Challenged by Its Beneficiary

For illustration, supposed TA had occurred and had not been challenged by B.

In this case, L believes that B is the *S*-type.

If L plays TA, he gets $\frac{1}{\alpha}8^{\alpha}\left(\frac{1}{1-\delta}\right)$. This is the highest possible payoff L can get, so he cannot gain by deviating.

4.9 L Played TAB in Period *t*-1, and Both DAC Transgression Had Occurred and Had Always Been Challenged by Their Respective Beneficiaries

In this case, L believes that both A and B are the *SP*-type. This case is similar to (4.7) above, and L will not deviate.

5. Conclusion

From the analysis above, and in particular from (13), (16), (19), and (58), if the conditions $p < p_{BR}^{SP} = \frac{(1-\delta)(1+\delta)7^{\alpha} + \delta^2 8^{\alpha} - 9^{\alpha} - \theta 8^{\alpha}}{(1-\delta)7^{\alpha} + \delta 8^{\alpha} - (1-\delta)(1+\theta 8^{\alpha}) - \delta(2^{\alpha} + \theta 8^{\alpha})}$, $p < p_{VR}^{S} = \frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1}$, $p < p_{VR}^{SP} = \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{(7^{\alpha} - 1) - \theta 8^{\alpha}}$, and $p > p_{LR} = \frac{6^{\alpha}(1-\delta)}{8^{\alpha} - \delta 6^{\alpha}}$ hold, the strategies described in Section 1 constitute an equilibrium. Because $p < p_{VR}^{S} = \frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1} < p_{VR}^{SP} = \frac{7^{\alpha} - 2^{\alpha} - \theta 8^{\alpha}}{(7^{\alpha} - 1) - \theta 8^{\alpha}}$, these conditions are equivalent to the condition

$$\frac{6^{\alpha}(1-\delta)}{8^{\alpha} - \delta 6^{\alpha}} < p < \min\left(\frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1}, \frac{(1-\delta)(1+\delta)7^{\alpha} + \delta^2 8^{\alpha} - 9^{\alpha} - \theta 8^{\alpha}}{(1-\delta)7^{\alpha} + \delta 8^{\alpha} - (1-\delta)(1+\theta 8^{\alpha}) - \delta(2^{\alpha} + \theta 8^{\alpha})}\right).$$

This establishes Result 2. Computational exercises show that there exists values of α, δ, and θ such that this condition is satisfied.

Appendix F

Informative Cheap Talk in the CR Game with Social Preferences and Incomplete Information

In this appendix, we show how cheap talk can be informative and help coordinate resistance in the CR game with social preferences. We first show that a pure-strategy informative cheap talk equilibrium exists in the one-shot CR game with social preferences and incomplete information.We then describe a pure-strategy informative cheap talk equilibrium in the repeated CR game with social preferences that supports separation between the *SP*-type and the *S*-type beneficiary. To keep this Appendix F at a reasonable length, we shall primarily focus on showing that an *SP*-type beneficiary will not deviate from this informative cheap talk equilibrium in the repeated CR game with social preferences and with communication, which is the most important and interesting no deviation condition for this equilibrium.

1. Informative Cheap Talk Equilibrium in the One-shot CR Game with Social Preferences

Consider first the one-shot CR game with social preferences and with communication. The timing of events is as follows:

1. Nature chooses the type of each player.
2. The leader, L, chooses his action $a_L \in A_L = \{TAB, TA, TB, NT\}$.
3. At the *responder communication stage,* responders A and B observe the leader's action, and then simultaneously choose a message $m_i \in M_i = \{AC, CH\}, i = A, B$.
4. Observing the message sent by the other responder, a responder updates her belief about the type of the other responder.
5. At the *responder action stage*, each responder chooses her action $a_i \in A_i = \{AC, CH\}, i = A, B$, as a function of the observed message profile $(m_A, m_B) \in M_A \times M_B$ sent by the responders.
6. Payoffs are realized for the leader and the two responders.

In this game, the players' strategies constitute a Perfect Bayesian equilibrium if:

(E1) Given the responders' strategies, L's chosen action maximizes his expected utility.

(E2) At the responder communication stage, for any responder, given the leader's chosen action and the other responder's strategy, each type of this responder's chosen message maximizes her expected utility.

(E3) Having observed the message sent by the other responder, a responder updates her belief about the type of the other responder based on the other responder's strategy according to Bayes' rule.

(E4) At the responder action stage, for any responder and for any observed message profile, each type of this responder's action maximizes her expected utility given the leader's chosen action and the other responder's strategy.

Note that (E1) reflects the assumption that an *S*-type leader and an *SP*-type leader will behave the same in this model.

Result F1: If $-\theta > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ and $p < \left(\frac{3}{4}\right)^\alpha$, then the following strategy profiles constitute a Perfect Bayesian equilibrium in the CR game with social preferences:

(i) The leader chooses NT. (ii) If the leader chooses TA, then both the *SP*-type and the *S*-type of responder A (who is the victim) will choose the message CH, while the *SP*-type of responder B (who is the beneficiary) will choose the message CH, and the *S*-type of responder B will choose the message AC. Both types of A will choose the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$. The *SP*-type of B will choose the action CH iff the message profile is $(m_A, m_B) = (CH, CH)$, and the *S*-type of B will choose the action AC for any message profile (m_A, m_B). (iii) if the leader chooses TB, then A (who is now the beneficiary) and B (who is now the victim) will adopt strategies that are mirror images of the strategies just described when the leader chooses TA. (iv) if the leader chooses TAB, then regardless of their type, both responders will choose the message CH, and will choose the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$. (v) if the leader chooses NT, then regardless of their type, both responders will choose the message AC, and will choose the action AC for any observed message profile (m_A, m_B).

Remark 1: Before proceeding to the proof, we provide interpretations of the conditions that ensure that Result F1 holds.

The condition $-\theta > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ is the (now familiar) condition that social preferences are sufficiently strong so that an *SP*-type beneficiary prefers that divide-and-conquer be defeated despite the fact that successful DAC transgression will increase her material payoff. As we shall show in the proof below, only the condition of $-\theta > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ is required for cheap talk to be informative in equilibrium. The condition $p < \left(\frac{3}{4}\right)^\alpha$ says that the probability that the beneficiary is an *SP*-type is sufficiently high from the leader's perspective, so that the leader will be deterred from practicing DAC. This condition, however, is not required for cheap talk to be informative in equilibrium.

Note that compared to Result 1 in Appendix C, Result 1 here does not require the condition of $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$. The intuition for this difference is as follows.

Consider the pure-strategy equilibrium in Result 1 in Appendix C in which both types of victim challenge, the *SP*-type beneficiary challenges, and the *S*-type beneficiary acquiesces when there is no communication. In the absence of communication, an *S*-type victim will be willing to challenge only if she believes that there is a sufficiently high probability that the beneficiary is an

SP-type beneficiary who will challenge. The condition $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ guarantees that this will be the case. Similarly, when deciding whether to challenge, an *SP*-type victim will do so if and only if she expects that there is a high enough probability that the beneficiary is an *SP*-type who will challenge. Note that compared to an *S*-type victim, an *SP*-type victim suffers from the extra utility loss of $\theta 8^\alpha$ when the leader succeeds in practicing DAC. Hence, an *SP*-type victim has a stronger incentive to challenge than an *S*-type victim, which means that the condition $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ is also sufficient to ensure that an *SP*-type A will challenge TA.

In the pure-strategy informative equilibrium in Result F1 here, informative cheap talk by the beneficiary allows the victim to learn the beneficiary's type. Hence, when making decision during the responder action stage, a victim will challenge only if the beneficiary sent a message of CH in the earlier communication stage. Informative cheap talk eliminates the victim's uncertainty about the beneficiary's type, so the condition $p < \frac{7^\alpha - 2^\alpha}{7^\alpha - 1}$ is no longer needed to support the pure-strategy informative equilibrium in Result F1 that also involves both types of victim challenge, the *SP*-type beneficiary challenges, and the *S*-type beneficiary acquiesces.

In this informative equilibrium, both types of victim will indicate Challenge and will challenge if and only if both the victim and the beneficiary have indicated Challenge in the communication stage. An *SP*-type beneficiary will indicate Challenge, and will challenge if and only if both the victim and the beneficiary have indicated Challenge in their messages. An *S*-type beneficiary will indicate Acquiesce, and will acquiesce regardless of the messages sent by the victim and beneficiary. These strategies constitute an equilibrium because an *SP*-type beneficiary prefers that a DAC transgression be defeated and has the incentive to send a message of CH to indicate that she is an *SP*-type so as to coordinate resistance with the victim. On the other hand, an *S*-type beneficiary prefers that a DAC transgression succeeds, and has no incentive to deviate to send a message of Challenge. Thus, communication can help coordinate resistance against DAC in this environment.

We now proceed to the proof.

Proof: We first consider the most interesting case when the leader practices DAC--illustrated by L's choice of TA. In this case, A is the victim and B is the beneficiary.

1.1 The no deviation conditions for the responders when the leader practices DAC against A

1.1.1 Consider first the no deviation conditions at the responder action stage in the TA subgame.

1.1.1.1 Suppose that the observed message profile is $(m_A, m_B) = (CH, CH)$.

We need to check that both types of victim and the *SP*-type beneficiary will challenge, and that the *S*-type beneficiary will acquiesce (this is condition (E4) in the definition of equilibrium above).

We first consider the no deviation conditions of a beneficiary.

Because both types of A (as a victim) will send a message of CH, observing a message of CH sent by the victim does not give the beneficiary new information about the victim's type. The beneficiary's posterior belief will be the same as her prior belief, which assigns probability p to the event that the victim is an *S*-type (this is E3 in the equilibrium conditions above).

Expecting that both types of victim will challenge, if the *SP*-type beneficiary challenges, she gets

$$U_B^{SP}(A_B = CH) = \frac{1}{\alpha} 7^{\alpha} \qquad (1)$$

If she acquiesce, she gets

$$U_B^{SP}(A_B = AC) = \frac{1}{\alpha}\left(9^{\alpha} + \theta 8^{\alpha}\right) \qquad (2)$$

So the *SP*-type beneficiary will challenge if

$$\frac{1}{\alpha} 7^{\alpha} > \frac{1}{\alpha}\left(9^{\alpha} + \theta 8^{\alpha}\right) \qquad (3)$$

which can be re-written as $-\theta > \dfrac{9^{\alpha} - 7^{\alpha}}{8^{\alpha}}$.

Now consider the no-deviating condition for the *S*-type beneficiary. An *S*-type beneficiary will only send a message of AC in equilibrium, so the event that nature has chosen the beneficiary to be an *S*-type and this *S*-type beneficiary observing a message profile $(m_A, m_B) = (CH, CH)$ is a probability zero event that should not happen in equilibrium. If the *S*-type beneficiary finds herself at this off-equilibrium information set, the *S*-type beneficiary is supposed to acquiesce. She will do so because AC is her dominant strategy. Formally, expecting that the victim is an *S*-type with probability p and that both types of victim will challenge, if the *S*-type beneficiary acquiesces, she gets $\frac{1}{\alpha} 9^{\alpha}$. If she challenges, she gets $\frac{1}{\alpha} 7^{\alpha}$. She will acquiesce.

We now check that both types of victim will indeed choose CH when the observed message profile is $(m_A, m_B) = (CH, CH)$.

Because only an *SP*-type beneficiary will send a message of CH, having observed a message of CH by the beneficiary, the victim's posterior belief is that the beneficiary is an *SP*-type with certainty.

Expecting that the beneficiary is an *SP*-type who will challenge, if the *SP*-type victim challenges, she gets $\frac{1}{\alpha}7^{\alpha}$. If she acquiesces, she gets $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$, so she will challenge.

Expecting that the beneficiary is an *SP*-type who will challenge, if the *S*-type victim challenges, she gets gets $\frac{1}{\alpha}7^{\alpha}$. If she acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$, so she will challenge.

1.1.1.2 Suppose that the observed message profile is $\left(m_A, m_B\right)=\left(CH, AC\right)$.

In this case, we need to check that both types of victim and both types of beneficiary will acquiesce.

We first consider the no deviation conditions of a beneficiary.

Because both types of victim will send a message of CH, observing a message of CH sent by the victim does not give the beneficiary new information about the victim's type. The beneficiary's posterior belief will be the same as her prior belief, which assigns probability p to the event that the victim is an *S*-type.

An *SP*-type beneficiary will only send a message of CH in equilibrium, so the event that nature has chosen the beneficiary to be an *SP*-type and this *SP*-type beneficiary observing a message profile $\left(m_A, m_B\right)=\left(CH, AC\right)$ is a probability zero event that should not happen in equilibrium. If the *SP*-type beneficiary finds himself at this off-equilibrium information set, she is supposed to choose AC.

Recall that both types of victim will challenge iff the observed message profile is $\left(m_A, m_B\right)=\left(CH, CH\right)$. Since the observed message profile is $\left(m_A, m_B\right)=\left(CH, AC\right)$, expecting that both types of victim will acquiesce, if the *SP*-type beneficiary acquiesces, she gets $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$. If she challenges, she gets $\frac{1}{\alpha}\left(8^{\alpha}+\theta 8^{\alpha}\right)$. She will acquiesce.

Now consider the no-deviating condition for the *S*-type beneficiary. That the *S*-type beneficiary will acquiesce when the observed message profile is $\left(m_A, m_B\right)=\left(CH, AC\right)$ follows from the fact that AC is the dominant strategy for an *S*-type beneficiary.

We now check that both types of victim will indeed choose AC when the observed message profile is $\left(m_A, m_B\right)=\left(CH, AC\right)$.

Because only an *S*-type beneficiary will send a message of AC, having observed a message of AC by the beneficiary, the victim's posterior belief is that the beneficiary is an *S*-type with certainty.

Expecting that the beneficiary is an *S*-type who will acquiesce, if the *SP*-type victim challenges, she gets $\frac{1}{\alpha}\left(1+\theta 8^{\alpha}\right)$. If she acquiesces, she gets $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$, so she will acquiesce.

Expecting that the beneficiary is an *S*-type who will acquiesce, if the *S*-type victim challenges, she gets $\frac{1}{\alpha}$. If she acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$, so she will acquiesce.

1.1.1.3 Suppose that the observed message profile is $\left(m_A, m_B\right)=\left(AC, CH\right)$.

In this case, we need to check that both types of victim and both types of beneficiary will acquiesce.

We first consider the no deviation conditions of the beneficiary.

Because both types of victim will send a message of CH in equilibrium, the event that the victim sends a message of AC is a probability zero event that should not happen in equilibrium. Bayes' rule now imposes no restriction on the beneficiary's posterior belief, and any belief is admissible. Note that in this case, as both types of victim will acquiesce when the observed message profile is $\left(m_A, m_B\right)=\left(AC, CH\right)$, it actually does not matter what posterior belief the beneficiary holds in this case. We shall simply assume that the beneficiary's posterior belief will be the same as the prior.

Expecting that both types of victim will acquiesce, if the *SP*-type beneficiary acquiesces, she gets $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$. If she challenges, she gets $\frac{1}{\alpha}\left(8^{\alpha}+\theta 8^{\alpha}\right)$. She will acquiesce.

Now consider the no-deviating condition for the *S*-type beneficiary. That the *S*-type beneficiary will acquiesce when the observed message profile is $\left(m_A, m_B\right)=\left(AC, CH\right)$ follows from the fact that AC is the dominant strategy for an *S*-type beneficiary.

We now consider the no deviation conditions of a victim.

Because only an *SP*-type beneficiary will send a message of CH, having observed a message of CH by the beneficiary, the victim's posterior belief is that the beneficiary is an *SP*-type with certainty.

Expecting that the beneficiary is an *SP*-type who will acquiesce (because the observed message profile is $(m_A, m_B) = (AC, CH)$), if the *SP*-type victim challenges, she gets $\frac{1}{\alpha}(1+\theta 8^{\alpha})$. If she acquiesces, she gets $\frac{1}{\alpha}(2^{\alpha} + \theta 8^{\alpha})$, so she will acquiesce.

Expecting that the beneficiary is an *SP*-type who will acquiesce, if the *S*-type victim challenges, she gets $\frac{1}{\alpha}$. If she acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$, so she will acquiesce.

1.1.1.4 Suppose that the observed message profile is $(m_A, m_B) = (AC, AC)$.

In this case, we need to check that both types of victim and both types of beneficiary will acquiesce.

We first consider the no deviation conditions of the beneficiary.

Because both types of victim will send a message of CH in equilibrium, the event that the victim sends a message of AC is a probability zero event that should not happen in equilibrium. Bayes' rule now imposes no restriction on the beneficiary's posterior belief, and any belief is admissible. Note again that in this case, as both types of victim will acquiesce when the observed message profile is $(m_A, m_B) = (AC, AC)$, it actually does not matter what posterior belief the beneficiary holds in this case. We shall simply assume that the beneficiary's posterior belief will be the same as the prior.

Expecting that both types of victim will acquiesce, if the *SP*-type beneficiary acquiesces, she gets $\frac{1}{\alpha}(9^{\alpha} + \theta 8^{\alpha})$. If she challenges, she gets $\frac{1}{\alpha}(8^{\alpha} + \theta 8^{\alpha})$. She will acquiesce.

Now consider the no-deviating condition for the *S*-type beneficiary. The fact that the *S*-type beneficiary will acquiesce when the observed message profile is $(m_A, m_B) = (AC, AC)$ follows from the fact that AC is the dominant strategy for an *S*-type beneficiary.

We now consider the no deviation conditions of a victim.

Because only an *S*-type beneficiary will send a message of AC in equilibrium, having observed a message of AC by the beneficiary, the victim's posterior belief is that the beneficiary is an *S*-type with certainty.

Expecting that the beneficiary is an *S*-type who will acquiesce, if the *SP*-type victim challenges, she gets $\frac{1}{\alpha}(1+\theta 8^{\alpha})$. If she acquiesces, she gets $\frac{1}{\alpha}(2^{\alpha} + \theta 8^{\alpha})$, so she will acquiesce.

Expecting that the beneficiary is an *SP*-type who will challenge, if the *S*-type victim challenges, she gets $\frac{1}{\alpha}$. If she acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$, so she will acquiesce.

1.1.2 Now consider the no deviation conditions at the responder communication stage in the TA subgame.

We need to show that both types of victim will choose the message CH, while the *SP*-type beneficiary will choose the message CH, and the *S*-type beneficiary will choose the message AC (this is condition E2 in the definition of equilibrium above).

Recall that when the leader chooses TA, A is the victim and B is the beneficiary.

We first show that the *SP*-type beneficiary will send a message of CH. If the *SP*-type beneficiary sends a message of CH, the observed message profile will be $(m_A, m_B) = (CH, CH)$. The *SP*-type beneficiary reveals to the victim that she is an *SP*-type beneficiary, and induces both types of victim to challenge in the responder action stage. The *SP*-type beneficiary herself will also challenge in the responder action stage. Thus, by sending a message of CH, the *SP*-type beneficiary gets $\frac{1}{\alpha}7^{\alpha}$. If the *SP*-type beneficiary instead sends a message of AC, the observed message profile will be $(m_A, m_B) = (CH, AC)$, and the *SP*-type beneficiary induces both types of victim to acquiesce in the responder action stage, and the *SP*-type beneficiary herself will also acquiesce in the responder action stage. Thus, by sending a message of AC, the *SP*-type beneficiary gets $\frac{1}{\alpha}\left(9^{\alpha} + \theta 8^{\alpha}\right) < \frac{1}{\alpha}7^{\alpha}$. So the *SP*-type beneficiary will send a message of CH.

We now show that the *S*-type beneficiary will send a message of AC. If the *S*-type beneficiary sends a message of CH, the observed message profile will be $(m_A, m_B) = (CH, CH)$. The *S*-type beneficiary will actually lead the victim to believe that she is an *SP*-type, and thus induces both types of victim to challenge in the responder action stage. The *S*-type beneficiary herself, however, will acquiesce in the responder action stage. Thus, by sending a message of CH, the *S*-type beneficiary gets $\frac{1}{\alpha}9^{\alpha}$. If the *S*-type beneficiary sends a message of AC, the observed message profile will be $(m_A, m_B) = (CH, AC)$. The *S*-type beneficiary induces both types of victim to acquiesce in the responder action stage, and the *S*-type beneficiary herself will also acquiesce in the responder action stage. Thus, by sending a message of AC, the *S*-type beneficiary also gets $\frac{1}{\alpha}9^{\alpha}$. So the *S*-type beneficiary has no incentive to deviate from the equilibrium strategy of sending a message of AC.

We now show that the *S*-type victim will send a message of CH. Suppose the *S*-type victim sends a message of CH. If nature chooses the beneficiary to be an *SP*-type (which occurs with probability $(1-p)$), the observed message profile will be $(m_A, m_B) = (CH, CH)$, and both the victim and the beneficiary will challenge in the responder action stage. The *S*-type victim gets $\frac{1}{\alpha}7^{\alpha}$ in this case. On the other hand, if nature chooses the beneficiary to be an *S*-type (which occurs with probability p), the observed message profile will now be $(m_A, m_B) = (CH, AC)$, and both the victim and the beneficiary will acquiesce in the responder action stage. The *S*-type victim gets $\frac{1}{\alpha}2^{\alpha}$ in this case. Therefore, by sending a message of CH, the *S*-type victim gets

$$U_A^S(m_A = CH) = p\frac{1}{\alpha}2^{\alpha} + (1-p)\frac{1}{\alpha}7^{\alpha} \qquad (4)$$

Suppose the *S*-type victim instead sends a message of AC. By doing so, the *S*-type victim ensures that the observed message profile (m_A, m_B) will differ from (CH, CH). Hence, both types of beneficiary will acquiesce and the *S*-type victim herself will also acquiesce.

Therefore, by sending a message of AC, the *S*-type victim gets

$$U_A^S(m_A = AC) = \frac{1}{\alpha}2^{\alpha} \qquad (5)$$

So the *S*-type victim will send a message of CH because

$$p\frac{1}{\alpha}2^{\alpha} + (1-p)\frac{1}{\alpha}7^{\alpha} > \frac{1}{\alpha}2^{\alpha} \qquad (6)$$

We now show that the *SP*-type victim will send a message of CH. Suppose the *SP*-type victim sends a message of CH. If nature chooses the beneficiary to be an *SP*-type, the observed message profile will be $(m_A, m_B) = (CH, CH)$, and both the victim and the beneficiary will challenge in the responder action stage. The *SP*-type victim gets $\frac{1}{\alpha}7^{\alpha}$ in this case. On the other hand, if nature chooses the beneficiary to be an *S*-type, the observed message profile will be $(m_A, m_B) = (CH, AC)$, and both the victim and the beneficiary will acquiesce in the responder action stage. The *SP*-type victim gets $\frac{1}{\alpha}\left(2^{\alpha} + \theta 8^{\alpha}\right)$ in this case.

Therefore, by sending a message of CH, the *SP*-type victim gets

$$U_A^{SP}(m_A = CH) = p\frac{1}{\alpha}\left(2^{\alpha} + \theta 8^{\alpha}\right) + (1-p)\frac{1}{\alpha}7^{\alpha} \qquad (7)$$

Suppose the *SP*-type victim instead sends a message of AC. By doing so, the *SP*-type victim ensures that the observed message profile (m_A, m_B) will differ from (CH, CH). Hence, both types of beneficiary will acquiesce and the *SP*-type victim herself will also acquiesce.

Therefore, by sending a message of AC, the *SP*-type victim gets

$$U_A^{SP}(m_A = AC) = \frac{1}{\alpha}\left(2^\alpha + \theta 8^\alpha\right) \quad (8)$$

So the *SP*-type victim will send a message of CH because

$$p\frac{1}{\alpha}\left(2^\alpha + \theta 8^\alpha\right) + (1-p)\frac{1}{\alpha}7^\alpha > \frac{1}{\alpha}\left(2^\alpha + \theta 8^\alpha\right) \quad (9)$$

1.2 The no deviation conditions for the responders when the leader practices DAC against B

This is the mirror image of case (1) and hence a proof is unnecessary.

1.3 The no deviation conditions for the responders when the leader chooses TAB (*T*ransgression *A*gainst *B*oth)

1.3.1 Consider first the no deviation conditions at the responder action stage in the TA subgame.

1.3.1.1 Suppose that the observed message profile is $(m_A, m_B) = (CH, CH)$.

In this case, we need to check that both types of A and both of B will challenge. Given the symmetry of the TAB subgame, suffice to check that both types of A will challenge in this case.

Because both types of B will send a message of CH, observing a message of CH sent by B does not give A new information about B's type, and A's posterior belief will be the same as her prior belief, which assigns probability p to the event that B is an *S*-type.

Expecting that both types of B will challenge, if the *SP*-type A challenges, she gets

$$U_B^{SP}(A_B = CH) = \frac{1}{\alpha}7^\alpha \quad (10)$$

If she acquiesces, she gets

$$U_B^{SP}(A_B = AC) = \frac{1}{\alpha}\left(2^\alpha + \theta 12^\alpha\right) \quad (11)$$

So the *SP*-type A will challenge if

$$\frac{1}{\alpha}7^{\alpha} > \frac{1}{\alpha}\left(2^{\alpha} + \theta 12^{\alpha}\right) \qquad (12)$$

which is always satisfied since $\theta < 0$.

Expecting that both types of B will challenge, if the *S*-type A challenges, she gets $\frac{1}{\alpha}7^{\alpha}$. If she acquiesce, she gets $\frac{1}{\alpha}2^{\alpha}$. So the *S*-type A will challenge.

1.3.1.2 Suppose that the observed message profile is $\left(m_A, m_B\right) = \left(CH, AC\right)$.

In this case, we need to check that both types of A and both types of B will acquiesce.

Because both types of B will send a message of CH in equilibrium, the event that B sends a message of AC is a probability zero event that should not happen in equilibrium. Bayes' rule imposes no restriction on A's posterior belief, and any belief is admissible. Note that in this case, as both types of B will acquiesce when the observed message profile is $\left(m_A, m_B\right) = \left(CH, AC\right)$, it actually does not matter what posterior belief A holds in this case. We shall simply assume that A's posterior belief will be the same as the prior.

Expecting that both types of B will acquiesce (because the observed message profile is $\left(m_A, m_B\right) = \left(CH, AC\right)$), if the *SP*-type A acquiesces, she gets $\frac{1}{\alpha}\left(2^{\alpha} + \theta 8^{\alpha}\right)$. If she challenges, she gets $\frac{1}{\alpha}\left(1 + \theta 8^{\alpha}\right)$. She will acquiesce.

Expecting that both types of B will acquiesce if the *S*-type A acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$. If she challenges, she gets $\frac{1}{\alpha}$. She will acquiesce.

Because both types of A will send a message of CH, observing a message of CH sent by A does not give B new information about A's type, and B's posterior belief will be the same as her prior belief, which assigns probability p to the event that A is an *S*-type.

Expecting that both types of A will acquiesce, if the *SP*-type B acquiesces, she gets $\frac{1}{\alpha}\left(2^{\alpha} + \theta 8^{\alpha}\right)$. If she challenges, she gets $\frac{1}{\alpha}\left(1 + \theta 8^{\alpha}\right)$. The *SP*-type B will acquiesce.

Expecting that both types of A will acquiesce, if the *S*-type B acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$. If she challenges, she gets $\frac{1}{\alpha}$. The *S*-type B will acquiesce.

1.3.1.3 Suppose that the observed message profile is $(m_A, m_B) = (AC, CH)$. This is the mirror image of (3.1.2) and hence a proof is unnecessary.

1.3.1.4 Suppose that the observed message profile is $(m_A, m_B) = (AC, AC)$.

In this case, we need to check that both types of A and both types of B will acquiesce. Because of symmetry, suffice to check that both types of A will acquiesce in this case.

Because both types of B will send a message of CH in equilibrium, the event that B sends a message of AC is a probability zero event that should not happen in equilibrium. Bayes' rule imposes no restriction on A's posterior belief, and any belief is admissible. Note that in this case, as both types of B will acquiesce when the observed message profile is $(m_A, m_B) = (CH, AC)$, it actually does not matter what posterior belief A holds in this case. We shall simply assume that A's posterior belief will be the same as the prior.

Expecting that both types of B will acquiesce, if the *SP*-type A acquiesces, she gets $\frac{1}{\alpha}\left(2^{\alpha} + \theta 8^{\alpha}\right)$. If she challenges, she gets $\frac{1}{\alpha}\left(1 + \theta 8^{\alpha}\right)$. She will acquiesce.

Expecting that both types of B will acquiesce, if the *S*-type A acquiesces, she gets $\frac{1}{\alpha}2^{\alpha}$. If she challenges, she gets $\frac{1}{\alpha}$. She will acquiesce.

1.3.2 Now consider the no deviation conditions at the responder communication stage in the TAB subgame.

We need to show that both types of A and both types of B will choose the message CH. Because of symmetry, suffice to check that both types of A will send the message CH.

Because both types of A will send a message of CH, observing a message of CH sent by A does not give B new information about A's type, and B's posterior belief will be the same as her prior belief, which assigns probability p to the event that A is an *S*-type.

Although a message of CH does not help reveal her type, by sending a message of CH, the *SP*-type A ensures that the observed message profile is $(m_A, m_B) = (CH, CH)$, which will induce both types of B to challenge in the responder action stage. The *SP*-type A will also challenge herself in the responder action stage. Thus, by sending a message of CH, the *SP*-type A gets $\frac{1}{\alpha}7^{\alpha}$. If the *SP*-type A instead sends a message of AC, the observed message profile will be $\left(m_A, m_B\right) = \left(AC, CH\right)$, and the *SP*-type A induces both types of B to acquiesce in the responder action stage, and the *SP*-type A herself will also acquiesce in the responder action stage. Thus, by sending a message of AC, the *SP*-type A gets $\frac{1}{\alpha}\left(2^{\alpha} + \theta 12^{\alpha}\right) < \frac{1}{\alpha}7^{\alpha}$. So the *SP*-type A will send a message of CH.

Although a message of CH does not help reveal her type, by sending a message of CH, the *S*-type A ensures that the observed message profile is $\left(m_A, m_B\right) = \left(CH, CH\right)$, which will induce both types of B to challenge in the responder action stage. The *S*-type A will also challenge herself in the responder action stage. Thus, by sending a message of CH, the *S*-type A gets $\frac{1}{\alpha}7^{\alpha}$. If the *S*-type A sends a message of AC, the observed message profile will be $\left(m_A, m_B\right) = \left(AC, CH\right)$. The *S*-type A induces both types of B to acquiesce in the responder action stage, and the *S*-type A herself will also acquiesce in the responder action stage. Thus, by sending a message of AC, the *S*-type A gets $\frac{1}{\alpha}2^{\alpha}$. So the *S*-type A will send a message of CH.

1.4 The no deviation conditions for the responders when the leader chooses NT (No Transgression)

1.4.1 Consider first the no deviation conditions in the responder action stage in the NT subgame.

We need to show that any type of any responder will choose to acquiesce. This follows simply from the fact that, for any type of any responder, acquiesce is the dominant strategy.

1.4.2 Consider first the no deviation conditions at the responder communication stage in the NT subgame.

We need to show that any type of any responder will send the message AC. This follows from the fact that, for any type of any responder, sending a message of CH or AC will both lead to the other responder and herself choose AC in the responder action stage. Hence, she does not have any incentive to deviate to sending CH.

Remark 2: As indicated earlier and demonstrated in the proof above, only the condition $-\theta > \frac{9^{\alpha} - 7^{\alpha}}{8^{\alpha}}$ is necessary for cheap talk to be informative.

1.5 The leader will in fact choose NT in equilibrium when $p < \left(\frac{3}{4}\right)^{\alpha}$

Finally, we show that when cheap talk is informative, the leader will not transgress in equilibrium if $p < \left(\frac{3}{4}\right)^{\alpha}$.

Suppose the leader practices divide-and-conquer against A. In this case, both types of victim will send a message of CH and will challenge iff the observed message profile is $(m_A, m_B) = (CH, CH)$, the *SP*-type beneficiary will send a message of CH and will challenge iff the observed message profile is $(m_A, m_B) = (CH, CH)$, the *S*-type beneficiary will send a message of AC and will acquiesce for any message profile (m_A, m_B). Coordinated resistance—which will give the leader a zero payoff--occurs if and only if nature chooses the beneficiary to be an *SP*-type so that the observed message profile will be $(m_A, m_B) = (CH, CH)$, which will induce both the victim and the beneficiary to challenge upon observing this message profile. On the other hand, when nature chooses the beneficiary to be an *S*-type, the *S*-type beneficiary will send a message of AC, the observed message profile will be $(m_A, m_B) = (CH, AC)$, and both the victim and the beneficiary will acquiesce upon observing this message profile. The leader will thus succeed in his divide-and-conquer if and only if the beneficiary (which is B) is an *S*-type in this case. Similarly, the leader's DAC against B will succeed if and only if the beneficiary (which will be A in this case) is an *S*-type.

Hence, if the leader practices divide-and-conquer, for example, if she chooses TA, she gets a payoff

$$U_L^S(A_L = TA) = p\frac{1}{\alpha}\left(8^{\alpha}\right) + (1-p)0 = p\frac{1}{\alpha}\left(8^{\alpha}\right) \qquad (13)$$

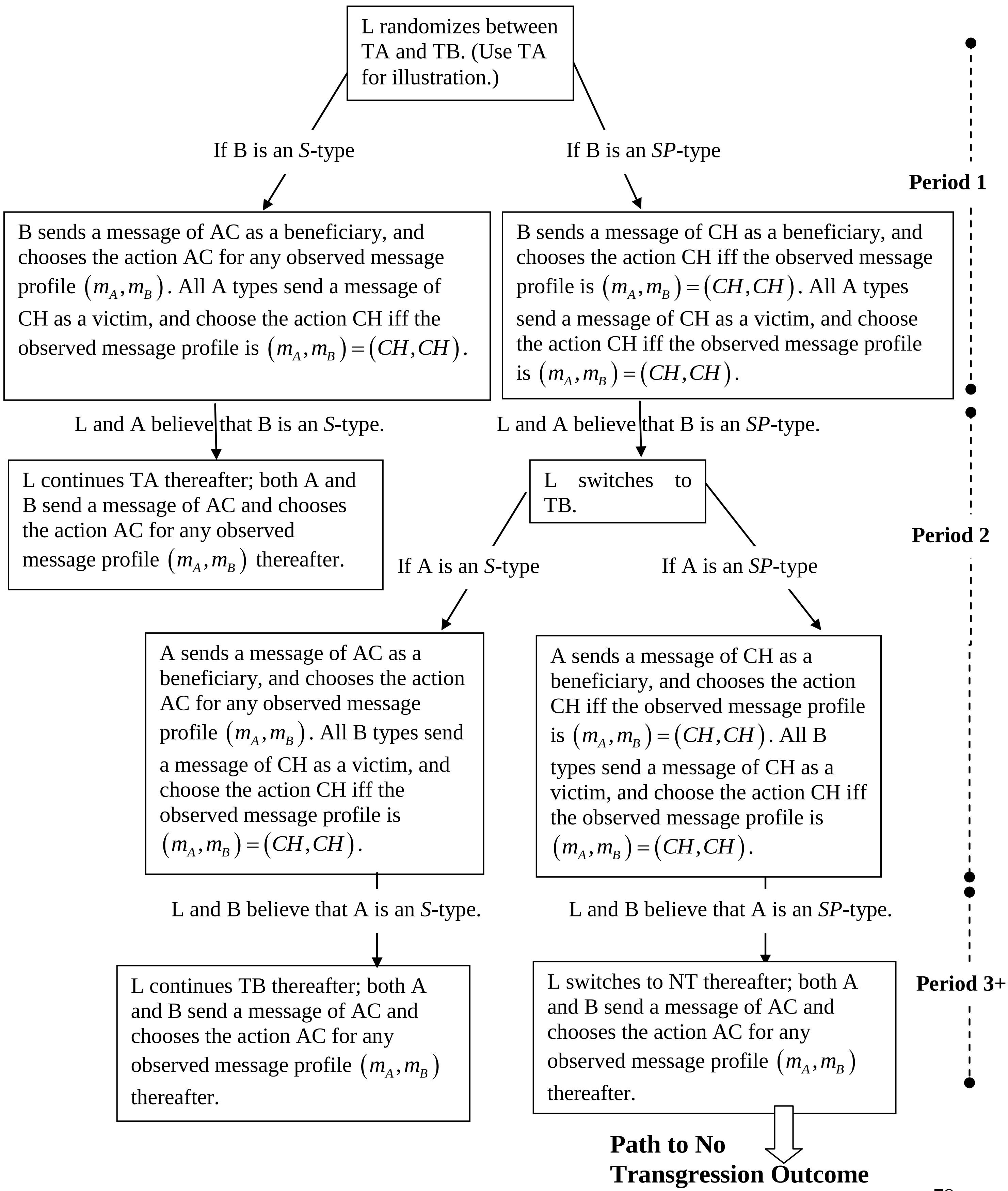
Suppose the leader chooses to transgress against both. In this case, regardless of their type, both responder will send a message of CH, and will also challenge upon observing the message profile $(m_A, m_B) = (CH, CH)$. Coordinated resistance will always occur, and the leader will get a zero payoff by choosing TAB.

Suppose the leader chooses not to transgress against any responder. In this case, regardless of their type, both responders will send a message of AC, and will also acquiesce at the responder action stage. The leader will thus get $\frac{1}{\alpha}\left(6^{\alpha}\right)$ by choosing NT. The leader will choose NT instead of TAB or DAC if $\frac{1}{\alpha}\left(6^{\alpha}\right) > p\frac{1}{\alpha}\left(8^{\alpha}\right)$, which is equivalent to $p < \left(\frac{3}{4}\right)^{\alpha}$.

2. Informative Cheap Talk Equilibrium in the Indefinitely Repeated CR Game with Social Preferences

We now proceed to show that there exists a pure-strategy informative cheap talk equilibrium in the indefinitely repeated CR game with social preferences. The three players now play an indefinitely repeated version of the game with social preferences and with communication described in Part 1.

Result F2: If $\frac{9^\alpha - 7^\alpha}{8^\alpha} < -\theta < \frac{\left(9^\alpha - 7^\alpha\right) + \delta\left(9^\alpha - 1\right) + \delta^2\left(9^\alpha - 2^\alpha\right)}{8^\alpha}$ and $\frac{6^\alpha\left(1-\delta\right)}{8^\alpha - \delta 6^\alpha} < p < \frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(2^\alpha + \theta 8^\alpha\right)\right]}$, then the following strategies constitute a Perfect Bayesian equilibrium in the indefinitely repeated CR game with social preferences and with communication:


L randomizes between TA and TB. (Use TA for illustration.)
If B is an *S*-type
If B is an *SP*-type
Period 1
B sends a message of AC as a beneficiary, and chooses the action AC for any observed message profile (m_A, m_B). All A types send a message of CH as a victim, and choose the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$.
B sends a message of CH as a beneficiary, and chooses the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$. All A types send a message of CH as a victim, and choose the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$.
L and A believe that B is an *S*-type.
L and A believe that B is an *SP*-type.
L continues TA thereafter; both A and B send a message of AC and chooses the action AC for any observed message profile (m_A, m_B) thereafter.
L switches to TB.
Period 2
If A is an *S*-type
If A is an *SP*-type
A sends a message of AC as a beneficiary, and chooses the action AC for any observed message profile (m_A, m_B). All B types send a message of CH as a victim, and choose the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$.
A sends a message of CH as a beneficiary, and chooses the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$. All B types send a message of CH as a victim, and choose the action CH iff the observed message profile is $(m_A, m_B) = (CH, CH)$.
L and B believe that A is an *S*-type.
L and B believe that A is an *SP*-type.
Period 3+
L continues TB thereafter; both A and B send a message of AC and chooses the action AC for any observed message profile (m_A, m_B) thereafter.
L switches to NT thereafter; both A and B send a message of AC and chooses the action AC for any observed message profile (m_A, m_B) thereafter.
Path to No Transgression Outcome

Remark 3: Before proceeding to the proof, we provide interpretations of the conditions that ensure that Result F2 holds.

The condition $\frac{9^\alpha - 7^\alpha}{8^\alpha} < -\theta < \frac{\left(9^\alpha - 7^\alpha\right) + \delta\left(9^\alpha - 1\right) + \delta^2\left(9^\alpha - 2^\alpha\right)}{8^\alpha}$ was the same condition that supports the separating equilibrium for the repeated game without communication (see the text and the proof in Appendix E). The condition $-\theta > \frac{9^\alpha - 7^\alpha}{8^\alpha}$ is the familiar condition that says that social preferences are sufficiently strong so that an *SP*-type beneficiary prefers that divide-and-conquer be defeated despite the fact that successful DAC transgression will increase her material payoff. The condition $-\theta < \frac{\left(9^\alpha - 7^\alpha\right) + \delta\left(9^\alpha - 1\right) + \delta^2\left(9^\alpha - 2^\alpha\right)}{8^\alpha}$ says that if an *SP*-type B knows that A is an *S*-type for certain, then she will choose AC instead of CH in period 1, because the present discounted value of the future utility losses of being "perpetually trapped" as a victim starting from period 2 is larger than the one period gain of successful coordinated resistance in period 1.

Recall that to support the separating equilibrium for the repeated game without communication, we require that

$$\frac{6^\alpha\left(1-\delta\right)}{8^\alpha - \delta 6^\alpha} < p < \min\left(\frac{7^\alpha - 2^\alpha}{7^\alpha - 1}, \frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(1-\delta\right)\left(1+\theta 8^\alpha\right) - \delta\left(2^\alpha + \theta 8^\alpha\right)\right]}\right).$$

With communication, however, to support separation we only require that

$$\frac{6^\alpha\left(1-\delta\right)}{8^\alpha - \delta 6^\alpha} < p < \frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(2^\alpha + \theta 8^\alpha\right)\right]}.$$

When deciding whether to practice DAC, the leader will do so if he believes there is a high enough probability that the beneficiary is an *S*-type so that he has a high enough probability to succeed. This is given by the condition $p > \frac{6^\alpha\left(1-\delta\right)}{8^\alpha - \delta 6^\alpha}$ and is required to ensure that the leader will attempt DAC in equilibrium both in the presence and the absence of communication.

Interestingly, while supporting a separating equilibrium between the *S*-type and *SP*-type beneficiary requires that the probability that an agent is the *S*-type to be sufficiently small both with and without communication, communication increases this critical probability. That is, communication makes it easier for separation to be supported as equilibrium because, compared to the case without communication, the *SP*-type beneficiary will be more willing to take the risk to challenge with communication even facing a slightly higher risk that the other responder is the *S*-type.

To see this, note that in the separating equilibrium with communication described in Result 2 in this Appendix, communication in period 1 does not help B learn A's type because both types of

A will challenge in period 1 in equilibrium. Since B still faces uncertainty about A's type, p still needs to be "sufficiently small" to support this informative equilibrium even in the presence of communication. What communication buys for the *SP*-type B in the repeated game when making decision as the beneficiary in period 1 is that if she challenges in period 1, communication will help her learn about A's type when she makes her decision *in period 2*.

In the absence of communication, an *SP*-type B who has challenged in period 1 as the beneficiary will challenge as a victim, and will get an expected utility $\frac{1}{\alpha}\left\{p\left(1+\theta 8^{\alpha}\right)+(1-p)7^{\alpha}\right\}$ in period 2 (evaluated in the beginning of period 2 before any action takes place). This reflects the fact that if A turns out to be the *S*-type (which happens with probability p), the *SP*-type B will get a payoff of $\frac{1}{\alpha}\left(1+\theta 8^{\alpha}\right)$ in period 2 because she will challenge while the *S*-type will acquiesce. On the other hand, if A turns out to be the *SP*-type (which happens with probability $(1-p)$), the *SP*-type B will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2 because both she and the *SP*-type A will challenge. With communication, an *SP*-type B will learn the type of A after communication in period 2, and will only challenge if she believes that A is an *SP*-type. Hence, an *SP*-type B will get an expected utility $\frac{1}{\alpha}\left\{p\left(2^{\alpha}+\theta 8^{\alpha}\right)+(1-p)7^{\alpha}\right\}$ in period 2 (evaluated in the beginning of period 2 before any action takes place) because she will acquiesce (and be able to save the cost of challenge) in the event that nature chooses A to be the *S*-type. This fact will make challenge in period 1 slightly more attractive to an *SP*-type B when making a decision as the beneficiary in period 1. This will therefore slightly increases the critical probability of A being an *S*-type that will support an *SP*-type B challenging TA in period 1 from $\frac{(1-\delta)(1+\delta)7^{\alpha}+\delta^{2}8^{\alpha}-9^{\alpha}-\theta 8^{\alpha}}{\delta\left[(1-\delta)7^{\alpha}+\delta 8^{\alpha}-(1-\delta)\left(1+\theta 8^{\alpha}\right)-\delta\left(2^{\alpha}+\theta 8^{\alpha}\right)\right]}$ to $\frac{(1-\delta)(1+\delta)7^{\alpha}+\delta^{2}8^{\alpha}-9^{\alpha}-\theta 8^{\alpha}}{\delta\left[(1-\delta)7^{\alpha}+\delta 8^{\alpha}-\left(2^{\alpha}+\theta 8^{\alpha}\right)\right]}$.

It is also interesting to observe that, as indicated in Result F1 above, in the one-shot CR game with social preferences so long as there is a positive probability that the beneficiary is the *SP*-type (which means that $p\neq 1$), then the condition $-\theta>\frac{9^{\alpha}-7^{\alpha}}{8^{\alpha}}$ alone is sufficient to ensure that cheap talk is informative and can support separation between the *SP*-type and the *S*-type beneficiary. Revealing that she is an *SP*-type does not expose the beneficiary to be picked as the victim in the future, and can help coordinate successful resistance with the victim to achieve the *SP*-type beneficiary's preferred outcome. *The result that cheap talk can be informative and can support separation between the SP-type and the S-type beneficiary in the one-shot CR game with social preferences does not require that* p *be smaller than a critical value between zero and one (although obviously we require* $p\neq 1$ *so that the beneficiary can be the SP-type), while such a threshold condition is required to support separation in the repeated CR game with social* *preferences both with and without and communication.* This observation provides a formal illustration of the implications of our key observation that repeated interaction is a two-edged sword in facilitating resistance in settings like this with endogenous role asymmetries. Repetition enables the victims and beneficiaries to use history-dependent strategies to facilitate their cooperation and coordination, but it also allows the leader to use history-dependent strategies to punish any beneficiary who refuses to cooperate with him.

Finally, recall that in the repeated CR game with social preferences and without communication, supporting the separating equilibrium that involves both victims challenging DAC in early periods also require that $p < \frac{7^{\alpha} - 2^{\alpha}}{7^{\alpha} - 1}$. When deciding whether to challenge TA as a victim in period 1, an *S*-type A will do so if and only if she expects that there is a high enough probability that B is an *SP*-type who will challenge as the beneficiary. This condition is no longer needed in the repeated CR game with social preferences and with communication. This is because A (as the victim in period 1 when TA occurs in period 1) can actually learn B's type after the communication stage in period 1.

We now proceed to the proof.

Proof: When we prove Result 2 in Appendix E for the indefinitely repeated CR game with social preferences and without communication, we present a lengthy proof in which we carefully specify the players' equilibrium strategies and posterior beliefs, including the players' actions and posterior beliefs at different information sets off the equilibrium path. The combined Appendices up to this point are already lengthy and we would like to keep this Appendix F at a reasonable length. Therefore, we do not provide a complete formal proof for Result F2 that carefully describes what happens at the large set of off equilibrium information sets (though we shall discuss some of these information sets for the responders in the first period in the proof below). Instead, we focus on explaining the main reasons why the responders will not deviate from their equilibrium strategies in the first period. The fact that the responders will not deviate from their equilibrium strategies in other periods, and that the leader will not deviate from their equilibrium strategy, can be proved in essentially the same manner by adapting the proof for Result 2 in Appendix E to take into account the presence of communication.

2.1 The no deviation conditions for the responders when the leader practices DAC against A in the first period

2.1.1 Consider first the no deviation conditions at the responder action stage in the first period. The most important point is to show that when the probability that the responders are the *SP*-type is sufficiently high, then an *SP*-type beneficiary facing a DAC transgression will be willing to take the risk to challenge the leader. We illustrate this by examining the decision of the *SP*-type B when the leader chooses TA in period 1. In particular, we first consider the case when the observed message profile is $(m_A, m_B) = (CH, CH)$ and investigates whether the *SP*-type B will follow through on his message of challenging in period 1. Considering this message profile

$\left(m_A, m_B\right) = \left(CH, CH\right)$ also enables us to show that communication also reduces the uncertainty regarding the type of the beneficiary faced by the victim when making decision in period 1.

2.1.1.1 Suppose that the observed message profile is $\left(m_A, m_B\right) = \left(CH, CH\right)$.

We need to check that both types of victim and the *SP*-type beneficiary will challenge, and that the *S*-type beneficiary will acquiesce.

We first consider the no deviation conditions of a beneficiary.

Because both types of A (as a victim) will send a message of CH in the first period, observing a message of CH sent by the victim in the first period does not give the beneficiary new information about the victim's type. The beneficiary's posterior belief will be the same as her prior belief, which assigns probability p to the event that the victim is an *S*-type.

2.1.1.1.1 No Deviation Condition for the *SP*-type B (Who is the Beneficiary)

Expecting that both types of victim will challenge in the first period, if the *SP*-type beneficiary challenges, she gets

$$\frac{1}{\alpha}\left\{7^\alpha + p\delta\left[\left(2^\alpha + \theta 8^\alpha\right) + \left(2^\alpha + \theta 8^\alpha\right)\left(\frac{\delta}{1-\delta}\right)\right] + \left(1-p\right)\delta\left[7^\alpha + 8^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\} \quad (14)$$

If she acquiesces, she gets

$$\frac{1}{\alpha}\left[9^\alpha + \theta 8^\alpha + \left(9^\alpha + \theta 8^\alpha\right)\left(\frac{\delta}{1-\delta}\right)\right] \quad (15)$$

The *SP*-type B will not deviate iff $(14) > (15)$. This requires that

$$p < \frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(2^\alpha + \theta 8^\alpha\right)\right]} \quad (16)$$

It is straightforward to establish that this critical probability on the R.H.S. of (16) is between zero and one by applying the Intermediate Value Theorem as we did in Appendix E. Furthermore, since $\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(2^\alpha + \theta 8^\alpha\right)\right] < \delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(1-\delta\right)\left(1+\theta 8^\alpha\right) - \delta\left(2^\alpha + \theta 8^\alpha\right)\right]$, we have:

$$\frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(1-\delta\right)\left(1+\theta 8^\alpha\right) - \delta\left(2^\alpha + \theta 8^\alpha\right)\right]} < \frac{\left(1-\delta\right)\left(1+\delta\right)7^\alpha + \delta^2 8^\alpha - 9^\alpha - \theta 8^\alpha}{\delta\left[\left(1-\delta\right)7^\alpha + \delta 8^\alpha - \left(2^\alpha + \theta 8^\alpha\right)\right]} \quad (17)$$

As indicated above, communication increases the critical value of p that will support separation between the *SP*-type and the *S*-type beneficiary and thus makes it easier to support separation.

Explanations for (14): If the *SP*-type B challenges TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}7^{\alpha}$ in the current period, and reveals that she is the *SP*-type. The leader will switch to playing TB in period 2, and the *SP*-type B will send a message of CH as the victim.

If A is the *S*-type (which occurs with probability p), A will a send a message of AC in period 2 as the beneficiary. Observing that the message profile is $(m_A, m_B) = (AC, CH)$, the *SP*-type B will acquiesce as a victim in period 2, and will get a payoff of $(2^{\alpha} + \theta 8^{\alpha})$ in period 2. Furthermore, both the leader and B will now conclude that A is an *S*-type. At the beginning of period 3, the leader now believes that A is the *S*-type and B is the *SP*-type, and he will continue to practice TB every period from period 3 onward. Starting from period 3, both A and B will send a message of AC and will acquiesce. The players will be playing (TB, AC, AC) every period from period 3 onward. The *SP*-type B will therefore be "trapped" in the outcome of (TB, AC, AC) with the payoff of $(2^{\alpha} + \theta 8^{\alpha})$ every period starting from period 3.

On the other hand, if A is the *SP*-type (which occurs with probability $(1-p)$), A will send a message of CH in period 2 as the beneficairy and will challenge in period 2 (so that B will also challenge upon seeing the message profile $(m_A, m_B) = (CH, CH)$ and will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2), which will lead both the leader and B conclude that A is an *SP*-type. At the beginning of period 3, the leader now believes that both A and B are the *SP*-type, and he will switch to playing NT in period 3. Starting from period 3, both A and B will send a message of AC and will acquiesce, the players will be playing (NT, AC, AC) thereafter, and the *SP*-type B will get a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period starting from period 3.

Explanations for (15): If the *SP*-type B deviates and acquiesces toward TA as a beneficiary in period 1, she gets a payoff of $\frac{1}{\alpha}(9^{\alpha} + \theta 8^{\alpha})$ in period 1, and both the leader and responder A now believe that B is the *S*-type. The leader will continue to practice TA in period 2, both A and B will send a message of AC and will acquiesce. From period 2 onward the players will be playing (TA, AC, AC) every period. By deviating to acquiesce toward TA in period 1, the *SP*-type B induces the leader and responder A to believe that she is an *S*-type, and the *SP*-type B will receive a payoff of $\frac{1}{\alpha}(9^{\alpha} + \theta 8^{\alpha})$ in every period thereafter.

Summary: The act of challenging TA as a beneficiary in period 1 is a *risky* action for the *SP*-type B. The lower the probability that A is an *S*-type, the less risky—and hence the more attractive—it is for the *SP*-type B to challenge TA in period 1. Note that when $-\theta > (9^{\alpha} - 7^{\alpha}) / 8^{\alpha}$, the *SP*-

type B's social preferences are sufficiently strong so that she prefers successful coordinated resistance against TA with a utility of $\frac{1}{\alpha}7^{\alpha}$ to getting the highest material payoff of 9 with a utility of $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$. Since A will always challenge TA as a victim regardless of her type in period 1, comparing to acquiescing, challenging TA in period 1 brings an increase in utility from $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$ to $\frac{1}{\alpha}7^{\alpha}$ in the current period.

If A is the *S*-type, however, challenging TA in period 1 will cause the *SP*-type B to get a utility of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ in period 2, which is lower than $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$–the payoff that the *SP*-type B can guarantee herself in period 2 by acquiescing in period 1. In addition, challenging TA in period 1 induces the leader to switch to TB in period 2 and will cause the *SP*-type B to be trapped in the outcome of (TB, AC, AC) with a utility of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)<\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$ every period starting from period 2. Thus, compared to acquiescing in period 1, challenging TA in period 1 will give the *SP*-type B a lower continuation payoff starting from period 2 if A turns out to be the *S*-type.

On the other hand, if A is an *SP*-type, challenging TA in period 1 will ensure that in period 2, the *SP*-type B gets a utility of $\frac{1}{\alpha}7^{\alpha}$ --which is higher than $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$. Furthermore, the successful coordinated resistance with the *SP*-type A in both periods 1 and 2--by revealing that both A and B are the *SP*-type--will deter the leader from practicing any transgression starting from period 3, and will give the *SP*-type B a utility of $\frac{1}{\alpha}8^{\alpha}>\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$ every period starting from period 3. Thus, compared to acquiescing in period 1, challenging TA in period 1 will give the *SP*-type B a higher continuation payoff starting from period 2 if A turns out to be the *SP*-type.

Equation (16) defines the critical probability p that ensures that the gain from successful coordinated resistance in period 1, together with the prospect of getting a higher continuation payoff starting from period 2 when A is the *SP*-type, are sufficient to offset the risk of getting a lower continuation payoff starting from period 2 if A is the *S*-type, so that the *SP*-type B is willing to take the risk to challenge in period 1.

Observation 1: As indicated earlier, in the presence of communication, when A is an *S*-type, challenging TA in period 1 will cause the *SP*-type B to get a utility of $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$ in the second period. This is because the *SP*-type B will learn from the message sent by A that A is an *S*-type and hence can avoid incurring the cost of unsuccessful challenge in period 2. On the other hand, in the absence of communication, when A is an *S*-type, challenging TA in period 1 will cause the *SP*-type B to get a utility of $\frac{1}{\alpha}\left(1+\theta 8^{\alpha}\right)$ as the *SP*-type B will engage in unsuccessful challenge in period 2. Communication increases the critical value of p that will support separation between the *SP*-type and the *S*-type beneficiary and makes it easier to support separation.

2.1.1.1.2 No Deviation Condition for the *S*-type B (Who is the Beneficiary)

An *S*-type beneficiary will only send a message of AC in equilibrium, so the event that nature has chosen the beneficiary to be an *S*-type and this *S*-type beneficiary observing a message profile $(m_A, m_B) = (CH, CH)$ in period 1 is a probability zero event that should not happen in equilibrium. If the *S*-type beneficiary finds himself at this off-equilibrium information set, the *S*-type beneficiary is supposed to acquiesce in period 1.

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha} + 9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \qquad (18)$$

Deviation - (CH)

$$\frac{1}{\alpha}\left\{7^{\alpha} + p\delta\left[2^{\alpha} + 2^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] + (1-p)\delta\left[7^{\alpha} + 8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\} \qquad (19)$$

The *S*-type B will not deviate because $(18) > (19)$.

<u>Explanations for (18)</u>: If the *S*-type B acquiesces toward TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}9^{\alpha}$ in period 1, and reveals that she is the *S*-type. The leader will continue to practice TA in period 2, both A and B will send a message of AC and will acquiesce, and from period 2 onward the players will be playing (TA, AC, AC) every period, and B will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ in every period starting from period 2. Note that (18) is the highest possible payoff that the *S*-type B can get. Hence, she cannot gain from deviating. Nevertheless, we now explain how (19) is determined.

<u>Explanations for (19)</u>: If the *S*-type B deviates to challenge TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 1, and both the leader and A will believe that she is the *SP*-type. The leader will switch to practice TB in period 2, and the *S*-type B will send a message of CH as a victm. If A is the *S*-type, A will send a message of AC in period 2 as a beneficiary. Observing that $(m_A, m_B) = (AC, CH)$, both A and B will acquiesce and the *S*-type B will get a payoff of $\frac{1}{\alpha}2^{\alpha}$ in period 2. Both the leader and B conclude that A is an *S*-type. The leader now believes that A is the *S*-type and B is the *SP*-type at the beginning of period 3, and will continue to practice TB

every period from period 3 onward. Starting from period 3, both A and B will send a message of AC and will acquiesce. and the players will be playing (TB, AC, AC) thereafter. This implies that starting from period 3, the *S*-type B will get a payoff of $\frac{1}{\alpha}2^{\alpha}$ every period.

On the other hand, if A is the *SP*-type, A will challenge in period 2 (so that B will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2), which will lead both the leader and B conclude that A is an *SP*-type. The leader now believes that both A and B are the *SP*-type at the beginning of period 3, and will switch to play NT every period starting from period 3. Starting from period 3, both A and B will send a message of AC and will acquiesce, the players will be playing (NT, AC, AC) thereafter, and the *S*-type B will get a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period.

Summary: Acquiescence in period 1 will ensure that the *S*-type B gets the highest payoff $\frac{1}{\alpha}9^{\alpha}$ in period 1, as well as in every period thereafter. So the *S*-type B cannot gain from deviating. In fact, deviating to challenging TA in period 1 is costly in in period 1 for the *S*-type B, and will cause the *S*-type B to always get a payoff lower than $\frac{1}{\alpha}9^{\alpha}$ in period 2 regardless of the type of A. Furthermore, if A is the *S*-type, challenging TA in period 1 can cause B to suffer from perpetual DAC transgression against her starting from period 2. On the other hand, if A is the *SP*-type, challenging TA in period 1 can cause the leader to play NT in period 3 and get a payoff lower than $\frac{1}{\alpha}9^{\alpha}$ thereafter. Since deviating to challenging in the first period involves a current loss, and will always lead to future losses, the *S*-type B will never want to deviate.

We now consider the no deviation conditions of a victim.

Because only an *SP*-type beneficiary will send a message of CH, having observed a message of CH by the beneficiary, the victim's posterior belief is that the beneficiary is an *SP*-type.

2.1.1.1.3 No Deviation Condition for the *S*-type A (Who is the Victim)

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{20}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \tag{21}$$

It is obvious that the *S*-type A will not deviate.

Interpretation: In this equilibrium, when TA occurs in the first period, the leader's decision regarding whether to continue to practice TA or switch to TB in the second period depends on whether B--the beneficiary of TA in the first period--challenges. In this repeated game equilibrium, given the leader's strategy, B's strategy, and A's own strategy, the *S*-type A's continuation payoff starting from period 2 is essentially determined by B's type, and is not affected by the *S*-type A's choice between acquiescing or challenging in period 1. In particular, given that B has revealed that she is an *SP*-type by sending a message CH in period 1, A expects that the leader will switch to practicing TB in period 2. Hence, the *S*-type A expects that her action in period 2 will cause B to be trapped in the (TB, AC, AC) outcome forever starting from period 2. Thus, the *S*-type A expects that regardless of her action in period 1, she will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ every period from period 2 onward. Therefore, in deciding whether to challenge as a victim in the first period, the *S*-type A is comparing the payoff of challenging (which is $\frac{1}{\alpha}7^{\alpha}$) in the first period with the payoff of acquiescing (which is $\frac{1}{\alpha}2^{\alpha}$) in the first period. Hence, she will challenge.

2.1.1.1.4 No Deviation Condition for the *SP*-type A (Who is the Victim)

No deviation - (CH)

$$\frac{1}{\alpha}\left[7^{\alpha}+\delta 7^{\alpha}+8^{\alpha}\left(\frac{\delta^{2}}{1-\delta}\right)\right] \tag{22}$$

Deviation - (AC)

$$\frac{1}{\alpha}\left[2^{\alpha}+\theta 8^{\alpha}+\delta 7^{\alpha}+8^{\alpha}\left(\frac{\delta^{2}}{1-\delta}\right)\right] \tag{23}$$

It is obvious that the *SP*-type A will not deviate.

Interpretation: Given that B has revealed that she is an *SP*-type by sending a message CH in period 1, A expects that the leader will switch to practicing TB in period 2. Hence, the *SP*-type A expects that she and B will both send a message of CH in period 2 and will both challenge, and she will get a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2. Furthermore, successful joint resistance in both periods

1 and 2 will induce the leader to switch to NT in period 3, and the players will be playing (NT, AC, AC) thereafter, and the *SP*-type B will get a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period from period 3 onward.

Therefore, in deciding whether to challenge as a victim in period 1, the *SP*-type A is comparing the payoff of challenging (which is $\frac{1}{\alpha}7^{\alpha}$) in period 1with the payoff of acquiescing (which is $\frac{1}{\alpha}\left(2^{\alpha}+\theta 8^{\alpha}\right)$) in period 1. Hence, she will challenge.

Observation 2: Note that in the repeated CR game with communication, when the leader practices TA in period 1, A can infer B's type from B's message sent as a beneficiary in period 1. Hence, when deciding whether to challenge as a victim in period 1, A does not face uncertainty about the beneficiary's type. Hence, there is no need for the probability that the responders are an *S*-type to be small enough to guarantee that an *S*-type A will take the risk to challenge divide-and-conquer in the first period as a victim.

On the other hand, in the repeated CR game without communication, in deciding whether to challenge as a victim in period 1, A faces uncertainty about the beneficiary's type. Hence, the probability that the responders are an *S*-type needs to be smaller than $\frac{7^{\alpha}-2^{\alpha}}{\left(7^{\alpha}-1\right)}$ to guarantee that an *S*-type A will take the risk to challenge divide-and-conquer in the first period as a victim. Similarly, the probability that the responders are an *S*-type needs to be smaller than $\frac{7^{\alpha}-2^{\alpha}-\theta 8^{\alpha}}{\left(7^{\alpha}-1\right)-\theta 8^{\alpha}}$ to guarantee that an *SP* -type A will take the risk to challenge divide-and-conquer in the first period as a victim.

In the equilibrium considered, an *S*-type B will send the message of AC in period 1 as the beneficiary. We thus need to consider whether an *S*-type B will follow through on a message of AC. For this purpose, we shall consider the case when the observed message profile is $\left(m_A, m_B\right)=\left(CH, AC\right)$. For brevity, we shall just show that the *S*-type B will follow through on his message in this case and do not report the no deviation conditions of the *SP*-type B and both types of A.

2.1.1.2 Suppose that the observed message profile is $\left(m_A, m_B\right)=\left(CH, AC\right)$.

Because both types of A (as a victim) will send a message of CH in the first period, observing a message of CH sent by the victim in the first period does not give the beneficiary new information about the victim's type. The beneficiary's posterior belief will be the same as her prior belief, which assigns probability p to the event that the victim is an *S*-type. The *S*-type B is supposed to acquiesce when observing this message profile:

No deviation - (AC)

$$\frac{1}{\alpha}\left[9^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right] \qquad (24)$$

If the *S*-type B acquiesces toward TA as a beneficiary, she gets a payoff of $\frac{1}{\alpha}9^{\alpha}$ in period 1, and reveals that she is the *S*-type to the leader. The leader will continue to practice TA in period 2, both A and B will send a message of AC and will acquiesce, and from period 2 onward the players will be playing (TA, AC, AC), and B will get a payoff of $\frac{1}{\alpha}9^{\alpha}$ in every period starting from period 2. Hence, acquiescing will give the *S*-type B the highest payoff every period starting in period 1, and she cannot gain from deviating.

A complete proof of Result F2 requires a complete specification of the players' equilibrium strategies and posterior beliefs, including the players' actions and posterior beliefs at different information sets that are off the equilibrium path. Such a specification and the corresponding complete proof can be done along the same lines as the proof in Appendix E, taking into account that the availability of communication expands the number of information sets (which further lengthens the complete proof compared to the case without communication in the already extremely lengthy Appendix E). Instead of presenting a complete proof, we offer some "informal" arguments that convey the key insight regarding why both types of B will in fact send the equilibrium messages in period 1, which, together with the proof above that shows that the responders have the incentive to follow through on their messages in period 1, are the key to Result F2.

For illustration, suppose that the leader will behave according to the following equilibrium strategies: being indifferent between TA and TB, he chooses TA in period 1. He will continue to practice TA unless TA was challenged by B in the previous period. In that case, he will switch to TB, and will continue to practice TB unless TB was challenged by A in the previous period. If both TA and TB have been challenged by the corresponding beneficiaries (which in equilibrium will imply that both TA and TB have been defeated given the equilibrium behavior of the victims), then he will switch to NT thereafter.

Given the leader (and victim) behavior, we first explain why when facing TA in period 1, the *SP*-type B will in fact send the message CH. We shall focus on the information set that is on the equilibrium path in which the victim A sends the message CH. We have already shown above that when the message profile is $(m_A, m_B) = (CH, CH)$, the *SP*-type B will follow through on her message and will challenge, and will get an intertemporal payoff given by equation (14) reproduced below:

$$\frac{1}{\alpha}\left\{7^{\alpha}+p\delta\left[\left(2^{\alpha}+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\delta\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\} (14)$$

This is also the payoff that the *SP*-type B will get be sending the message CH as the beneficiary when facing TA in period 1. That is, by sending the message CH in period 1, the *SP*-type B induces successful coordinated resistance in period 1 and gets a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 1. She also induces the leader to switch to TB in period 2, which gives her a continuation payoff of $\frac{1}{\alpha}\left\{p\left[\left(2^{\alpha}+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}$ in period 2 (evaluated at period 2). We now argue that given the behavior of A and the leader, the *SP*-type B cannot do any better be deviating to sending the message AC.

To see this, first note that given that the leader has chosen TA, the highest payoff that the *SP*-type B can get in period 1 is when successful joint resistance occurs (that is, $\left(A_A, A_B\right)=\left(CH, CH\right)$), which will give her a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 1. By deviating and sending AC, the *SP*-type B will induce the outcome of $\left(A_A, A_B\right)=\left(AC, AC\right)$ and get the lower payoff of $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$ in period 1. Since deviating to sending the message AC in period 1 involves a current loss that equals $\frac{1}{\alpha}\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]$ in period 1, for such a deviation to be profitable, it must give the *SP*-type B a continuation payoff at period 2 that is sufficiently higher than $\frac{1}{\alpha}\left\{p\left[\left(2^{\alpha}+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}$ to compensate for the loss in period 1 due to the deviation.

Suppose that following the deviation of sending the message AC, the *SP*-type B chooses AC in period 1. This will induce the leader to choose TA in period 2 and lead to the players getting the outcome (TA, AC, AC) every period starting from period 2. In this case, deviating to sending the message AC in period 1 will give the *SP*-type B a payoff of $\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$ in period 1 and a continuation payoff of $\frac{1}{\alpha}\left[\frac{1}{1-\delta}\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]$ in period 2 (evaluated at period 2). That is, in this case, deviating to sending the message of AC will give the *SP*-type B an intertermporal payoff given by (15):

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \qquad (15)$$

However, we already showed earlier that when $\frac{6^{\alpha}(1-\delta)}{8^{\alpha}-\delta 6^{\alpha}}<p<\frac{(1-\delta)(1+\delta)7^{\alpha}+\delta^{2}8^{\alpha}-9^{\alpha}-\theta 8^{\alpha}}{\delta\left[(1-\delta)7^{\alpha}+\delta 8^{\alpha}-\left(2^{\alpha}+\theta 8^{\alpha}\right)\right]}$, the *SP*-type B prefers to challenge in period 1 and get the payoff given by (14) to getting the payoff given by (15). Therefore, the *SP*-

type B cannot gain by deviating to sending the message AC if she chooses the action AC in period 1 after sending this deviating message. This is the case we focused on in our discussion above, which will be the case, for example, if both A and B always acquiesce starting from period 1 upon seeing the message $(m_A, m_B) = (CH, AC)$ in period 1. That is, a message of AC by B as a beneficiary in period 1 will induce the responders to play the "pessimistic equilibrium" of both acquiescing thereafter.

However, we can show that even if we consider the possibility that coordinated resistance may still be possible after a message of AC by B as a beneficiary in period 1, the *SP*-type B again cannot gain by deviating to sending the message AC and then choosing the action CH in period 1 after sending this deviating message either. By choosing the action CH in period 1, the *SP*-type B will induce the leader to choose TB in period 2 (note that the leader observes only the action chosen by B in period 1, while A observes both the message and the action chosen by B in period 1. Hence, it is possible that A and the leader may have different posterior belief regarding the type of B in the beginning of period 2). Importantly, in this case, the *SP*-type B cannot expect to get a continuation payoff in period 2 that is higher than $\frac{1}{\alpha}\left\{p\left[\left(2^\alpha+\theta 8^\alpha\right)+\left(2^\alpha+\theta 8^\alpha\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[7^\alpha+8^\alpha\left(\frac{\delta}{1-\delta}\right)\right]\right\}$, the original equilibrium continuation payoff at period 2 she can get by sending the equilibrium message CH in period 1.

To see this, first note that by sending the equilibrium message CH in period 1, when A turns out to be the *S*-type, the *SP*-type B gets a continuation payoff of $\frac{1}{\alpha}\left[\left(2^\alpha+\theta 8^\alpha\right)+\left(2^\alpha+\theta 8^\alpha\right)\left(\frac{\delta}{1-\delta}\right)\right]$ in period 2. This is because the *S*-type A will send the message AC and choose her dominant action AC in period 2, which will give the *SP*-type B a payoff of $\frac{1}{\alpha}\left(2^\alpha+\theta 8^\alpha\right)$ in period 2, and also causes the players to get the outcome (TB, AC, AC) that will give the *SP*-type B a payoff of $\frac{1}{\alpha}\left(2^\alpha+\theta 8^\alpha\right)$ every period from period 3 onward. However, if the *SP*-type B deviates to sending the message AC and then choose the action CH in period 1 to induce the leader to choose TB in period 2, when the A turns out to be the *S*-type, the *SP*-type B cannot get a continuation payoff higher than $\frac{1}{\alpha}\left[\left(2^\alpha+\theta 8^\alpha\right)+\left(2^\alpha+\theta 8^\alpha\right)\left(\frac{\delta}{1-\delta}\right)\right]$ in period 2. This is because of the fact that the *S*-type A will again choose her dominant action AC in period 2, which will give the *SP*-type B a payoff less than or equal to $\frac{1}{\alpha}\left(2^\alpha+\theta 8^\alpha\right)$ in period 2, and also causes the players to get the outcome (TB, AC, AC) that will give the *SP*-type B a payoff of $\frac{1}{\alpha}\left(2^\alpha+\theta 8^\alpha\right)$ every period from period 3 onward.

On the other hand, by sending the equilibrium message CH in period 1, when A turns out to be the *SP*-type, the *SP*-type B gets a continuation payoff of $\frac{1}{\alpha}\left\{\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}$ in period 2. This is because the *SP*-type A will send the message CH and choose CH in period 2, which will give the *SP*-type B a payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2, and also causes the players to get the outcome (NT, AC, AC) that will give the *SP*-type B a payoff of $\frac{1}{\alpha}8^{\alpha}$ every period from period 3 onward. If the *SP*-type B deviates to sending the message AC in period 1 and then choose the action CH in period 1 to induce the leader to choose TB in period 2, when A turns out to be the *S*-type, the *SP*-type B cannot get a continuation payoff higher than $\frac{1}{\alpha}\left\{\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}$ in period 2. In period 2, given that the leader chooses TB, the original equilibrium already gives the *SP*-type B the highest possible payoff of $\frac{1}{\alpha}7^{\alpha}$ in period 2 so she cannot gain anything in period 2 from the contemplated deviation. Furthermore, the original equilibrium actually gives the *SP*-type B the payoff of $\frac{1}{\alpha}8^{\alpha}$. This is the highest payoff that the *SP*-type B can ever get in this game (note that $\frac{1}{\alpha}8^{\alpha}>\frac{1}{\alpha}7^{\alpha}>\frac{1}{\alpha}\left(9^{\alpha}+\theta 8^{\alpha}\right)$) every period starting from period 3, so she cannot get a higher continuation payoff in period 3 from the contemplated deviation. These observations imply that the *SP*-type B cannot get a continuation payoff in period 2 higher than $\frac{1}{\alpha}\left\{p\left[\left(2^{\alpha}+\theta 8^{\alpha}\right)+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right]+(1-p)\left[7^{\alpha}+8^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]\right\}$ by sending the message AC in period 1 and then choosing the action CH in period 1. Since the contemplated deviation will involve a current loss of $\frac{1}{\alpha}\left[7^{\alpha}-\left(9^{\alpha}+\theta 8^{\alpha}\right)\right]$ in period 1 and cannot lead to a higher continuation payoff in period 2, it is unprofitable.

Finally, in equilibrium, by sending the message AC and then choosing the action AC in period 1, the *S*-type B gets an intertemporal payoff of $\frac{1}{\alpha}\left[9^{\alpha}+9^{\alpha}\left(\frac{\delta}{1-\delta}\right)\right]$. Under the original equilibrium, the *S*-type B gets a payoff of $\frac{1}{\alpha}9^{\alpha}$ --which is the highest payoff she can get in this game-- in every period starting from period 2, so she cannot gain from deviating.

This discussion provides intuition regarding why both types of beneficiary cannot gain by deviating from their equilibirum messages in period 1. This, together with the earlier discussion that shows that both types of beneficiary will follow through on her sent message in period 1, are the key to establishing Result F2 since the key challenge of sustaining the equilibrium is to ensure that the beneficiary does not deviate. To complete the arguments, one will also need to show that a beneficiary in other periods cannot gain from deviating to sending the non-equilibrium messages, that the leader and the victim will not deviate from the equilibrium strategies and that the postulated belief updating rules are consistent with the equilibrium strategies. These steps can be completed by adapting the proof in Appendix E to take into account the presence of communication as we have done here.

Appendix G

A model of CR Game with Social Preferences and Complete Information

In the text, we discuss informally a one-shot CR game with social preferences and complete information, in which an agent has social preferences for certain and this fact is common knowledge. In this setting, social preferences transform the DAC subgame into a stag hunt game with complete information. This observation can be stated formally as follows:

Result G 1: If $|\theta| > \left(9^\alpha - 7^\alpha\right)/8^\alpha$, then each of the three following strategy profiles constitutes a Nash Equilibrium in the DAC subgame with social preferences and complete information:

(i) Both the victim and the beneficiary acquiesce.
(ii) Both the victim and the beneficiary challenges.
(iii) The victim challenges with a probability $\tau = \frac{9^\alpha - 8^\alpha}{(7^\alpha - 8^\alpha - \theta 8^\alpha)} \in (0,1)$, and the beneficiary challenges with a probability $\gamma = \frac{2^\alpha - 1}{\left(7^\alpha - 1 - \theta 8^\alpha\right)} \in (0,1)$.

This result can be obtained by applying the results described in Figure 1 in Appendix C by setting $p = 0$, as the one-shot CR game with social preferences and complete information is a special case of the one-shot CR game with social preferences and incomplete information. Note that (iii) in Result G1 corresponds to case (iii)' in Appendix C.

We now discuss the repeated CR game with social preferences and complete information. To keep this already lengthy appendix at a reasonable length, we do not present detailed proofs of the no-deviating conditions for every player at every possible information set, as we did in Appendix E. Such proofs are straightforward for both the Divide-and-Conquer equilibrium with Persistent Beneficiary and the Divide-and-Conquer equilibrium with Punishment against the Challenging Beneficiary in this model with complete information about social preferences. Instead, we focus on discussing the no-deviation condition for the beneficiary when DAC takes place.

Consider first the Divide-and-Conquer equilibrium with Persistent Beneficiary, in which the leader practices DAC against responder A in the first period and will continue to do so in every period, while both responders always acquiesce when facing DAC, but always challenge if the leader transgresses against both of them.

If the beneficiary does not deviate and acquiesces, she gets

$$\frac{1}{\alpha}\left[9^\alpha + \theta 8^\alpha + \left(9^\alpha + \theta 8^\alpha\right)\left(\frac{\delta}{1-\delta}\right)\right] \qquad (1)$$

If she deviates and challenges, she gets

$$\frac{1}{\alpha}\left\{8^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right\} \tag{2}$$

So, she will not deviate.

In this equilibrium, the leader always chooses B as the beneficiary of DAC, and B does not want to deviate to challenge. Challenging is costly and will not only cause her to suffer a loss in the current period, but is futile and will not lead to coordinated resistance either in the current period or in the future. Importantly, the leader always targets A as the victim in this equilibrium. Therefore, when B challenges in the current period but fails to spark off any coordinated resistance, as the persistent beneficiary B will still get her beneficiary payoff in all future periods. She will get the material payoff as a beneficiary, although she suffers a psychological disutility due to the fact that the leader will now succeed in his transgression.

Now consider the Divide-and-Conquer equilibrium with Punishment against the Challenging Beneficiary, in which the leader practices DAC against responder A in the first period and will continue to practice DAC against A except when the current beneficiary challenges him. If the beneficiary in the current period (which is B in this case) challenges him, the leader will punish the challenging beneficiary by switching to DAC against B in the next period, and will continue to do so except when challenged by the beneficiary (now A). The responders will always acquiesce when facing DAC but will always challenge if the leader transgresses against both of them.

If the beneficiary does not deviate and acquiesces, she gets

$$\frac{1}{\alpha}\left[9^{\alpha}+\theta 8^{\alpha}+\left(9^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right] \tag{3}$$

If she deviates and challenges, she gets

$$\frac{1}{\alpha}\left\{8^{\alpha}+\theta 8^{\alpha}+\left(2^{\alpha}+\theta 8^{\alpha}\right)\left(\frac{\delta}{1-\delta}\right)\right\} \tag{4}$$

The beneficiary again does not want to deviate to challenge, because the leader will punish any challenging beneficiary in this equilibrium. If a beneficiary challenges, coordinated resistance will still fail in the current period, and the current beneficiary will also then be trapped as a victim of DAC in all future periods.

References

Cho, In-Koo, and David Kreps, "Signalling Games and Stable Equibria," *Quarterly Journal of Economics,* 102: 79-221, 1987.

Cox, James, Daniel Friedman, and Steven Gjerstad, "A Tractable Model of Reciprocity and Fairness," *Games and Economic Behavior,* 59: 17-45, 2007.

Govindan, Srihari, and Robert Wilson, "On Forward Induction," *Econometrica,* 77: 1-28, 2009.