

# Depth versus Rigidity in the Design of International Trade Agreements

Leslie Johns

Supplemental Appendix – September 3, 2012

## 1 Alternative Punishment Mechanisms

The one-period utility functions of the home and foreign government— $W$  and  $W^*$ , respectively—are as follows:

$$\begin{aligned} W(t, \tau, a) &= a u(t) - t - u(\tau) \\ W^*(t, \tau, \alpha) &= \alpha u(\tau) - \tau - u(t) \end{aligned}$$

Losses are:

$$\begin{aligned} L(\tau) &= W(t, t_B, a) - W(t, \tau, a) = u(\tau) - u(t_B) \\ L^*(t) &= W^*(t_B, \tau, \alpha) - W^*(t, \tau, \alpha) = u(t) - u(t_B) \end{aligned}$$

Let  $\chi_P$  denote the continuation payoff from the punishment that occurs if at least one player defects. Assume that  $\chi_P$  is not a function of the specific value of the defection tariff.

Let  $\chi_C$  denote the continuation payoff if the treaty remains in effect (neither player defects).

Recall that  $a, \alpha \sim_{iid} U[1, A]$  for large  $A$ ,  $u' > 0$ , and  $u'' < 0$ .

### 1.1 Optimal Tariffs

The home country's expected utility from violating the binding and not paying the fine (defection) is:

$$EU(D|t, a) = a u(t) - t - \int_1^A u(\tau(\alpha)) dH(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dH(\alpha) + \delta \chi_P$$

So the optimal defection tariff solves:

$$\begin{aligned} \frac{\partial EU(D|t, a)}{\partial t} &= a u'(t) - 1 = 0 \\ \Leftrightarrow u'(t) &= \frac{1}{a} \Leftrightarrow t_D(a) = u'^{-1}\left(\frac{1}{a}\right) \end{aligned}$$

This violates the binding iff:

$$t_D(a) = u'^{-1}\left(\frac{1}{a}\right) > t_B \Leftrightarrow \frac{1}{a} < u'(t_B) \Leftrightarrow a > \frac{1}{u'(t_B)} \equiv a_B$$

The home country's expected utility from violating the binding and paying the fine (settlement) is:

$$\begin{aligned} EU(S|t, a) &= a u(t) - t - \sigma L^*(t) - \int_1^A u(\tau(\alpha)) dH(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dH(\alpha) \\ &\quad + H(\alpha_D) \delta \beta \chi_C + [1 - H(\alpha_D)] \delta \chi_P \end{aligned}$$

So the optimal settlement tariff solves:

$$\begin{aligned} \frac{\partial EU(S|t, a)}{\partial t} &= a u'(t) - 1 - \sigma u'(t) = 0 \\ \Leftrightarrow u'(t) &= \frac{1}{a - \sigma} \Leftrightarrow t_S(a) = u'^{-1} \left( \frac{1}{a - \sigma} \right) \end{aligned}$$

This violates the binding iff:

$$\begin{aligned} t_S(a) &= u'^{-1} \left( \frac{1}{a - \sigma} \right) > t_B \Leftrightarrow \frac{1}{a - \sigma} < u'(t_B) \\ \Leftrightarrow a &> \frac{1}{u'(t_B)} + \sigma \equiv a_S \end{aligned}$$

Note that:  $t_S(a) < t_D(a)$  for all  $a$ .

The optimal cooperative tariff is:

$$t_B(a) = \begin{cases} t_D(a) & \text{if } a < a_B \\ t_B & \text{if } a_B \leq a \end{cases}$$

## 1.2 Equilibrium Regions

The home country's expected utility from actions  $C$ ,  $S$ , and  $D$  given optimal tariff levels are:

$$\begin{aligned} EU(C|t_B(a), a) &= a u(t_B(a)) - t_B(a) - \int_1^A u(\tau(\alpha)) dH(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dH(\alpha) \\ &\quad + H(\alpha_D) \delta \beta \chi_C + [1 - H(\alpha_D)] \delta \chi_P \\ EU(S|t_S(a), a) &= a u(t_S(a)) - t_S(a) - \sigma L^*(t_S(a)) - \int_1^A u(\tau(\alpha)) dH(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dH(\alpha) \\ &\quad + H(\alpha_D) \delta \beta \chi_C + [1 - H(\alpha_D)] \delta \chi_P \\ EU(D|t_D(a), a) &= a u(t_D(a)) - t_D(a) - \int_1^A u(\tau(\alpha)) dH(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dH(\alpha) + \delta \chi_P \end{aligned}$$

To compare utility from actions  $C$  and  $S$ , define for  $a_S \leq a$ :

$$\begin{aligned} \hat{\Delta}(a) &= EU(C|t_B(a), a) - EU(S|t_S(a), a) \\ &= a u(t_B) - t_B - a u(t_S(a)) + t_S(a) + \sigma L^*(t_S(a)) \end{aligned}$$

Note that  $t_S(a_S) = t_B$ , so  $\hat{\Delta}(a_S) = 0$ . Also:

$$\begin{aligned} \frac{\partial \hat{\Delta}}{\partial a} &= u(t_B) - u(t_S(a)) - (a - \sigma) u'(t_S(a)) \frac{\partial t_S(a)}{\partial a} + \frac{\partial t_S(a)}{\partial a} \\ &= u(t_B) - u(t_S(a)) < 0 \end{aligned}$$

So  $S$  strictly dominates  $C$  for all  $a_S < a$ .

To compare utility from actions  $S$  and  $D$ , define for  $a_S \leq a$ :

$$\begin{aligned}
\bar{\Delta}(a) &= EU(S|t_S(a), a) - EU(D|t_D(a), a) \\
&= a u(t_S(a)) - t_S(a) - \sigma L^*(t_S(a)) \\
&\quad - a u(t_D(a)) + t_D(a) + \delta H(\alpha_D) (\beta \chi_C - \chi_P) \\
\text{So: } \frac{\partial \bar{\Delta}}{\partial a} &= (a - \sigma) u'(t_S(a)) \frac{\partial t_S(a)}{\partial a} - \frac{\partial t_S(a)}{\partial a} + u(t_S(a)) \\
&\quad + \frac{\partial t_D(a)}{\partial a} - a u'(t_D(a)) \frac{\partial t_D(a)}{\partial a} - u(t_D(a)) \\
&= u(t_S(a)) - u(t_D(a)) < 0
\end{aligned}$$

So  $D$  strictly dominates  $S$  for sufficiently large values of  $a$ . By symmetry, indifference point  $a_D$  is implicitly defined by:

$$\begin{aligned}
\lambda &= a_D [u(t_S(a_D)) - u(t_D(a_D))] + t_D(a_D) - t_S(a_D) \\
&\quad - \sigma L^*(t_S(a_D)) + \delta H(a_D) (\beta \chi_C - \chi_P) = 0
\end{aligned}$$

The equilibrium exists iff:  $\bar{\Delta}(a_S) > 0$ .

### 1.3 Continuation Values

Let  $t_E(a)$  denote equilibrium tariffs when the institution is in place.

The continuation payoff for home from the treaty being in effect is:

$$\begin{aligned}
\chi_C &= \int_0^A [a u(t_E(a)) - t_E(a)] dH(a) - \sigma \int_{a_S}^{a_D} L^*(t_E(a)) dH(a) - \int_1^A u(\tau_E(\alpha)) dH(\alpha) \\
&\quad + \sigma \int_{\alpha_S}^{\alpha_D} L^*(\tau_E(\alpha)) dH(\alpha) + \delta H(a_D)^2 \beta \chi_C + \delta [1 - H(a_D)^2] \chi_P \\
&= \frac{\Psi}{1 - \delta \beta H(a_D)^2} \\
\text{where } \Psi &= \int_1^A [(a - 1) u(t_E(a)) - t_E(a)] dH(a) + \delta [1 - H(a_D)^2] \chi_P
\end{aligned}$$

### 1.4 Comparative Statics

#### Full Compliance

Recall that the binding is not violated if  $a < a_S = \frac{1}{u'(t_B)} + \sigma$ . So the probability that the binding is not violated is  $H(a_S)$ .

$$\frac{\partial a_S}{\partial t_B} = \frac{-u''(t_B)}{[u'(t_B)]^2} > 0 \quad \text{and} \quad \frac{\partial a_S}{\partial \sigma} = 1 > 0$$

#### Stability

The institution is stable if  $a < a_D$ . By the implicit function theorem:

$$\frac{\partial a_D}{\partial t_B} = -\frac{\lambda_{t_B}}{\lambda_{a_D}} \quad \text{and} \quad \frac{\partial a_D}{\partial \sigma} = -\frac{\lambda_\sigma}{\lambda_{a_D}}$$

Then for large  $A$ :

$$\begin{aligned} \lambda_{a_D} &= (a_D - \sigma) u'(t_S(a_D)) \frac{\partial t_S(a_D)}{\partial a_D} - \frac{\partial t_S(a_D)}{\partial a_D} - a_D u'(t_D(a_D)) \frac{\partial t_D(a_D)}{\partial a_D} + \frac{\partial t_D(a_D)}{\partial a_D} \\ &\quad + u(t_S(a_D)) - u(t_D(a_D)) + \delta H(a_D) \frac{\partial (\beta \chi_C - \chi_P)}{\partial a_D} + \delta h(a_D) (\beta \chi_C - \chi_P) \\ &= u(t_S(a_D)) - u(t_D(a_D)) + \delta H(a_D) \frac{\partial (\beta \chi_C - \chi_P)}{\partial a_D} + \delta h(a_D) (\beta \chi_C - \chi_P) < 0 \end{aligned}$$

$$\lambda_{t_B} = \sigma u'(t_B) + \delta H(a_D) \frac{\partial (\beta \chi_C - \chi_P)}{\partial t_B} > 0$$

$$\begin{aligned} \lambda_\sigma &= (a_D - \sigma) u'(t_S(a_D)) \frac{\partial t_S(a_D)}{\partial \sigma} - \frac{\partial t_S(a_D)}{\partial \sigma} \\ &\quad - [u(t_S(a_D)) - u(t_B)] + \delta H(a_D) \frac{\partial (\beta \chi_C - \chi_P)}{\partial \sigma} \\ &= -[u(t_S(a_D)) - u(t_B)] + \delta H(a_D) \frac{\partial (\beta \chi_C - \chi_P)}{\partial \sigma} < 0 \end{aligned}$$

$$\text{So:} \quad \frac{\partial a_D}{\partial t_B} > 0 \quad \text{and} \quad \frac{\partial a_D}{\partial \sigma} < 0$$

### Depth versus Rigidity

Recall that  $\chi_C$  is the expected utility of a state from being a member of the cooperative regime. In equilibrium,  $\lambda = 0$ . So for any pair  $(t_B, \sigma)$ :

$$\chi_C = \frac{a_D [u(t_D(a_D)) - u(t_S(a_D))] - t_D(a_D) + t_S(a_D) + \sigma L^*(t_S(a_D))}{\delta \beta H(a_D)} + \frac{\chi_P}{\beta}$$

The two first-order conditions on the optimal pair  $(t_B, \sigma)$  are:

$$\begin{aligned} \frac{d\chi_C}{dt_B} &= \frac{\partial \chi_C}{\partial t_B} + \frac{\partial \chi_C}{\partial a_D} \frac{\partial a_D}{\partial t_B} = 0 \\ \frac{d\chi_C}{d\sigma} &= \frac{\partial \chi_C}{\partial \sigma} + \frac{\partial \chi_C}{\partial a_D} \frac{\partial a_D}{\partial \sigma} = 0 \end{aligned}$$

This implies that:

$$\frac{\frac{\partial \chi_C}{\partial t_B}}{\frac{\partial \chi_C}{\partial \sigma}} = \frac{\frac{\partial a_D}{\partial t_B}}{\frac{\partial a_D}{\partial \sigma}}$$

$$\frac{\left( \frac{\partial \chi_C}{\partial t_B} \right) \left( \frac{\partial a_D}{\partial \sigma} \right)}{\frac{\partial \chi_C}{\partial \sigma}} = \frac{\partial a_D}{\partial t_B}$$

So for any pair  $(t_B, \sigma)$  that generates  $\chi_C(t_B, \sigma) = \chi^*$ :

$$\begin{aligned}
\frac{dt_B}{d\sigma} &= \frac{-\left(\frac{d\chi_C}{d\sigma}\right)}{\frac{d\chi_C}{dt_B}} = \frac{-\left(\frac{\partial\chi_C}{\partial\sigma} + \frac{\partial\chi_C}{\partial a_D} \frac{\partial a_D}{\partial\sigma}\right)}{\frac{\partial\chi_C}{\partial t_B} + \frac{\partial\chi_C}{\partial a_D} \frac{\partial a_D}{\partial t_B}} \\
&= \frac{-\left(\frac{\partial\chi_C}{\partial\sigma} + \frac{\partial\chi_C}{\partial a_D} \frac{\partial a_D}{\partial\sigma}\right)}{\frac{\partial\chi_C}{\partial t_B} + \left(\frac{\partial\chi_C}{\partial a_D}\right) \frac{\left(\frac{\partial\chi_C}{\partial t_B}\right) \left(\frac{\partial a_D}{\partial\sigma}\right)}{\frac{\partial\chi_C}{\partial\sigma}}} \\
&= \frac{-\frac{\partial\chi_C}{\partial\sigma} \left(\frac{\partial\chi_C}{\partial\sigma} + \frac{\partial\chi_C}{\partial a_D} \frac{\partial a_D}{\partial\sigma}\right)}{\left(\frac{\partial\chi_C}{\partial t_B}\right) \left(\frac{\partial\chi_C}{\partial\sigma}\right) + \left(\frac{\partial\chi_C}{\partial a_D}\right) \left(\frac{\partial\chi_C}{\partial t_B}\right) \left(\frac{\partial a_D}{\partial\sigma}\right)} \\
&= \frac{-\frac{\partial\chi_C}{\partial\sigma} \left(\frac{\partial\chi_C}{\partial\sigma} + \frac{\partial\chi_C}{\partial a_D} \frac{\partial a_D}{\partial\sigma}\right)}{\frac{\partial\chi_C}{\partial t_B} \left[\frac{\partial\chi_C}{\partial\sigma} + \left(\frac{\partial\chi_C}{\partial a_D}\right) \left(\frac{\partial a_D}{\partial\sigma}\right)\right]} = \frac{-\frac{\partial\chi_C}{\partial\sigma}}{\frac{\partial\chi_C}{\partial t_B}}
\end{aligned}$$

where:

$$\begin{aligned}
\frac{\partial\chi_C}{\partial\sigma} &= \frac{1}{\delta\beta H(a_D)} \left[ -a_D u'(t_S(a_D)) \frac{\partial t_S(a_D)}{\partial\sigma} + \frac{\partial t_S(a_D)}{\partial\sigma} + \sigma u'(t_S(a_D)) \frac{\partial t_S(a_D)}{\partial\sigma} \right] \\
&\quad + \frac{1}{\delta\beta H(a_D)} [L^*(t_S(a_D))] + \left(\frac{1}{\beta}\right) \frac{\partial\chi_P}{\partial\sigma} \\
&= \frac{L^*(t_S(a_D))}{\delta\beta H(a_D)} + \left(\frac{1}{\beta}\right) \frac{\partial\chi_P}{\partial\sigma} \\
\frac{\partial\chi_C}{\partial t_B} &= \frac{-\sigma u'(t_B)}{\delta\beta H(a_D)} + \left(\frac{1}{\beta}\right) \frac{\partial\chi_P}{\partial t_B}
\end{aligned}$$

So:

$$\begin{aligned}
\frac{dt_B}{d\sigma} &= \frac{-\frac{\partial\chi_C}{\partial\sigma}}{\frac{\partial\chi_C}{\partial t_B}} = \frac{-\left[\frac{L^*(t_S(a_D))}{\delta\beta H(a_D)} + \left(\frac{1}{\beta}\right) \frac{\partial\chi_P}{\partial\sigma}\right]}{\frac{-\sigma u'(t_B)}{\delta\beta H(a_D)} + \left(\frac{1}{\beta}\right) \frac{\partial\chi_P}{\partial t_B}} \\
&= \frac{L^*(t_S(a_D)) + \delta H(a_D) \frac{\partial\chi_P}{\partial\sigma}}{\sigma u'(t_B) - \delta H(a_D) \frac{\partial\chi_P}{\partial t_B}} > 0 \quad \text{for small } A
\end{aligned}$$

## 2 Asymmetric Type Distributions

Assume that home country type,  $a$ , is distributed according to distribution function  $H(a)$ . Denote home continuation payoffs by  $\chi_N$  and  $\chi_C$ .

Assume that foreign country type,  $\alpha$ , is distributed according to distribution function  $F(\alpha)$ . Denote foreign continuation payoffs  $\chi_N^*$  and  $\chi_C^*$ .

The one-period utility functions of the home and foreign government— $W$  and  $W^*$ , respectively—are as follows:

$$\begin{aligned}
W(t, \tau, a) &= a u(t) - t - u(\tau) \\
W^*(t, \tau, \alpha) &= \alpha u(\tau) - \tau - u(t)
\end{aligned}$$

Losses are:

$$\begin{aligned} L(\tau) &= W(t, \tau_B, a) - W(t, \tau, a) = u(\tau) - u(\tau_B) \\ L^*(t) &= W^*(t_B, \tau, \alpha) - W^*(t, \tau, \alpha) = u(t) - u(t_B) \end{aligned}$$

## 2.1 Optimal Tariffs

### Home

The home country's expected utility from violating the binding and not paying compensation (defection) is:

$$U(D|t, a) = a u(t) - t - \int u(\tau(\alpha)) dF(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dF(\alpha) + \delta \chi_N$$

So the optimal defection tariff solves:

$$\begin{aligned} \frac{\partial U(D|t, a)}{\partial t} &= a u'(t) - 1 = 0 \\ \Leftrightarrow u'(t) &= \frac{1}{a} \Leftrightarrow t_D(a) = u'^{-1}\left(\frac{1}{a}\right) \end{aligned}$$

This violates the home binding iff:

$$t_D(a) = u'^{-1}\left(\frac{1}{a}\right) > t_B \Leftrightarrow \frac{1}{a} < u'(t_B) \Leftrightarrow a > \frac{1}{u'(t_B)} \equiv a_B$$

The home country's expected utility from violating the binding and paying compensation (settlement) is:

$$\begin{aligned} U(S|t, a) &= a u(t) - t - \sigma L^*(t) - \int u(\tau(\alpha)) dF(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dF(\alpha) \\ &\quad + F(\alpha_D) \delta \beta \chi_C + [1 - F(\alpha_D)] \delta \chi_N \end{aligned}$$

So the optimal settlement tariff solves:

$$\begin{aligned} \frac{\partial U(S|t, a)}{\partial t} &= a u'(t) - 1 - \sigma u'(t) = 0 \\ \Leftrightarrow u'(t) &= \frac{1}{a - \sigma} \Leftrightarrow t_S(a) = u'^{-1}\left(\frac{1}{a - \sigma}\right) \end{aligned}$$

This violates the home binding iff:

$$\begin{aligned} t_S(a) &= u'^{-1}\left(\frac{1}{a - \sigma}\right) > t_B \Leftrightarrow \frac{1}{a - \sigma} < u'(t_B) \\ \Leftrightarrow a &> \frac{1}{u'(t_B)} + \sigma \equiv a_S \end{aligned}$$

Note that:  $t_S(a) < t_D(a)$  for all  $a$ . The optimal cooperative tariff is:

$$t_B(a) = \begin{cases} t_D(a) & \text{if } a < a_B \\ t_B & \text{if } a_B \leq a \end{cases}$$

### Foreign

The foreign country's expected utility from from violating the binding and not paying compensation (defection) is:

$$U^*(D|\tau, \alpha) = \alpha u(\tau) - \tau - \int u(t(a)) dH(a) + \int_{a_S}^{a_D} \sigma L^*(t(a)) dH(a) + \delta \chi_N^*$$

So the optimal defection tariff solves:

$$\begin{aligned} \frac{\partial U^*(D|\tau, \alpha)}{\partial \tau} &= \alpha u'(\tau) - 1 = 0 \\ \Leftrightarrow u'(\tau) &= \frac{1}{\alpha} \Leftrightarrow \tau_D(\alpha) = u'^{-1}\left(\frac{1}{\alpha}\right) \end{aligned}$$

This violates the foreign binding iff:

$$\tau_D(\alpha) = u'^{-1}\left(\frac{1}{\alpha}\right) > \tau_B \Leftrightarrow \frac{1}{\alpha} < u'(\tau_B) \Leftrightarrow \alpha > \frac{1}{u'(\tau_B)} \equiv \alpha_B$$

The foreign country's expected utility from violating the binding and paying compensation (settlement) is:

$$\begin{aligned} U^*(S|\tau, \alpha) &= \alpha u(\tau) - \tau - \sigma L(\tau) - \int u(t(a)) dH(a) + \int_{a_S}^{a_D} \sigma L^*(t(a)) dH(a) \\ &\quad + H(a_D) \delta \beta \chi_C^* + [1 - H(a_D)] \delta \chi_N^* \end{aligned}$$

So the optimal settlement tariff solves:

$$\begin{aligned} \frac{\partial U^*(S|\tau, \alpha)}{\partial \tau} &= \alpha u'(\tau) - 1 - \sigma u'(\tau) = 0 \\ \Leftrightarrow u'(\tau) &= \frac{1}{\alpha - \sigma} \Leftrightarrow \tau_S(\alpha) = u'^{-1}\left(\frac{1}{\alpha - \sigma}\right) \end{aligned}$$

This violates the foreign binding iff:

$$\begin{aligned} \tau_S(\alpha) &= u'^{-1}\left(\frac{1}{\alpha - \sigma}\right) > \tau_B \Leftrightarrow \frac{1}{\alpha - \sigma} < u'(\tau_B) \\ \Leftrightarrow \alpha &> \frac{1}{u'(\tau_B)} + \sigma \equiv \alpha_S \end{aligned}$$

Note that:  $\tau_S(\alpha) < \tau_D(\alpha)$  for all  $\alpha$ . The optimal cooperative tariff is:

$$\tau_B(\alpha) = \begin{cases} \tau_D(\alpha) & \text{if } \alpha < \alpha_B \\ \tau_B & \text{if } \alpha_B \leq \alpha \end{cases}$$

## 2.2 Equilibrium Regions

### Home

The home country's expected utility from actions  $C$ ,  $S$ , and  $D$  given tariff levels above are:

$$\begin{aligned}
 U(C|t_B(a), a) &= a u(t_B(a)) - t_B(a) - \int u(\tau(\alpha)) dF(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dF(\alpha) \\
 &\quad + F(\alpha_D) \delta \beta \chi_C + [1 - F(\alpha_D)] \delta \chi_N \\
 U(S|t_S(a), a) &= a u(t_S(a)) - t_S(a) - \sigma L^*(t_S(a)) - \int u(\tau(\alpha)) dF(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dF(\alpha) \\
 &\quad + F(\alpha_D) \delta \beta \chi_C + [1 - F(\alpha_D)] \delta \chi_N \\
 U(D|t_D(a), a) &= a u(t_D(a)) - t_D(a) - \int u(\tau(\alpha)) dF(\alpha) + \int_{\alpha_S}^{\alpha_D} \sigma L(\tau(\alpha)) dF(\alpha) + \delta \chi_N
 \end{aligned}$$

To compare home utility from actions  $C$  and  $S$ , define for  $a_S \leq a$ :

$$\begin{aligned}
 \hat{\Delta}(a) &= U(C|t_B(a), a) - U(S|t_S(a), a) \\
 &= a u(t_B) - t_B - a u(t_S(a)) + t_S(a) + \sigma L^*(t_S(a))
 \end{aligned}$$

Note that  $t_S(a_S) = t_B$ , so  $\hat{\Delta}(a_S) = 0$ . Also:

$$\begin{aligned}
 \frac{\partial \hat{\Delta}}{\partial a} &= u(t_B) - u(t_S(a)) - (a - \sigma) u'(t_S(a)) \frac{\partial t_S(a)}{\partial a} + \frac{\partial t_S(a)}{\partial a} \\
 &= u(t_B) - u(t_S(a)) < 0
 \end{aligned}$$

So  $S$  strictly dominates  $C$  for all  $a_S < a$ .

To compare home utility from actions  $S$  and  $D$ , define for  $a_S \leq a$ :

$$\begin{aligned}
 \bar{\Delta}(a) &= U(S|t_S(a), a) - U(D|t_D(a), a) \\
 &= a u(t_S(a)) - t_S(a) - \sigma L^*(t_S(a)) \\
 &\quad - a u(t_D(a)) + t_D(a) + \delta F(\alpha_D) (\beta \chi_C - \chi_N) \\
 \text{So: } \frac{\partial \bar{\Delta}}{\partial a} &= (a - \sigma) u'(t_S(a)) \frac{\partial t_S(a)}{\partial a} - \frac{\partial t_S(a)}{\partial a} + u(t_S(a)) \\
 &\quad + \frac{\partial t_D(a)}{\partial a} - a u'(t_D(a)) \frac{\partial t_D(a)}{\partial a} - u(t_D(a)) \\
 &= u(t_S(a)) - u(t_D(a)) < 0
 \end{aligned}$$

So  $D$  strictly dominates  $S$  for sufficiently large values of  $a$ .

Indifference point  $a_D$  is implicitly defined by:

$$\begin{aligned}
 \lambda &= a_D [u(t_S(a_D)) - u(t_D(a_D))] + t_D(a_D) - t_S(a_D) \\
 &\quad - \sigma L^*(t_S(a_D)) + \delta F(\alpha_D) (\beta \chi_C - \chi_N) = 0
 \end{aligned}$$



The equilibrium exists iff:  $\bar{\Delta}(a_S) > 0$ .

### Foreign

The foreign country's expected utility from actions  $C$ ,  $S$ , and  $D$  given tariff levels above are:

$$\begin{aligned}
U^*(C|\tau_B(\alpha), \alpha) &= \alpha u(\tau_B(\alpha)) - \tau_B(\alpha) - \int u(t(a))dH(a) + \int_{a_S}^{a_D} \sigma L^*(t(a))dH(a) \\
&\quad + H(a_D)\delta\beta\chi_C^* + [1 - H(a_D)]\delta\chi_N^* \\
U^*(S|\tau_S(\alpha), \alpha) &= \alpha u(\tau_S(\alpha)) - \tau_S(\alpha) - \sigma L(\tau_S(\alpha)) - \int u(t(a))dH(a) + \int_{a_S}^{a_D} \sigma L^*(t(a))dH(a) \\
&\quad + H(a_D)\delta\beta\chi_C^* + [1 - H(a_D)]\delta\chi_N^* \\
U^*(D|\tau_D(\alpha), \alpha) &= \alpha u(\tau_D(\alpha)) - \tau_D(\alpha) - \int u(t(a))dH(a) + \int_{a_S}^{a_D} \sigma L^*(t(a))dH(a) + \delta\chi_N^*
\end{aligned}$$

To compare foreign utility from actions  $C$  and  $S$ , define for  $\alpha_S \leq \alpha$ :

$$\begin{aligned}
\hat{\Delta}^*(\alpha) &= U^*(C|\tau_B(\alpha), \alpha) - U^*(S|\tau_S(\alpha), \alpha) \\
&= \alpha u(\tau_B) - \tau_B - \alpha u(\tau_S(\alpha)) + \tau_S(\alpha) + \sigma L(\tau_S(\alpha))
\end{aligned}$$

Note that  $\tau_S(\alpha_S) = \tau_B$ , so  $\hat{\Delta}^*(\alpha_S) = 0$ . Also:

$$\begin{aligned}
\frac{\partial \hat{\Delta}^*}{\partial \alpha} &= u(\tau_B) - u(\tau_S(\alpha)) - (\alpha - \sigma)u'(\tau_S(\alpha))\frac{\partial \tau_S(\alpha)}{\partial \alpha} + \frac{\partial \tau_S(\alpha)}{\partial \alpha} \\
&= u(\tau_B) - u(\tau_S(\alpha)) < 0
\end{aligned}$$

So  $S$  strictly dominates  $C$  for all  $\alpha_S < \alpha$ .

To compare foreign utility from actions  $S$  and  $D$ , define for  $\alpha_S \leq \alpha$ :

$$\begin{aligned}
\bar{\Delta}^*(\alpha) &= U^*(S|\tau_S(\alpha), \alpha) - U^*(D|\tau_D(\alpha), \alpha) \\
&= \alpha u(\tau_S(\alpha)) - \tau_S(\alpha) - \sigma L(\tau_S(\alpha)) \\
&\quad - \alpha u(\tau_D(\alpha)) + \tau_D(\alpha) + \delta H(a_D)(\beta\chi_C^* - \chi_N^*) \\
\text{So: } \frac{\partial \bar{\Delta}^*}{\partial \alpha} &= (\alpha - \sigma)u'(\tau_S(\alpha))\frac{\partial \tau_S(\alpha)}{\partial \alpha} - \frac{\partial \tau_S(\alpha)}{\partial \alpha} + u(\tau_S(\alpha)) \\
&\quad + \frac{\partial \tau_D(\alpha)}{\partial \alpha} - \alpha u'(\tau_D(\alpha))\frac{\partial \tau_D(\alpha)}{\partial \alpha} - u(\tau_D(\alpha)) \\
&= u(\tau_S(\alpha)) - u(\tau_D(\alpha)) < 0
\end{aligned}$$

So  $D$  strictly dominates  $S$  for sufficiently large values of  $\alpha$ .

Indifference point  $\alpha_D$  is implicitly defined by:

$$\begin{aligned} \lambda^* &= \alpha_D [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))] + \tau_D(\alpha_D) - \tau_S(\alpha_D) \\ &\quad - \sigma L(\tau_S(\alpha_D)) + \delta H(a_D) (\beta \chi_C^* - \chi_N^*) = 0 \end{aligned}$$

The equilibrium exists iff:  $\bar{\Delta}^*(\alpha_S) > 0$ .

### 2.3 Continuation Values

Let  $t_E(a)$  and  $\tau_E(\alpha)$  denote equilibrium tariffs of the home and foreign country, respectively, when the institution is in place.

If the institution does not exist, then the home and foreign country choose  $t_D(a)$  and  $\tau_D(\alpha)$  in every time period. This yields anarchy continuation payoffs:

$$\begin{aligned} \chi_N &= \frac{1}{1-\delta} \left\{ \int [a u(t_D(a)) - t_D(a)] dH(a) - \int u(\tau_D(\alpha)) dF(\alpha) \right\} \\ \chi_N^* &= \frac{1}{1-\delta} \left\{ \int [\alpha u(\tau_D(\alpha)) - \tau_D(\alpha)] dF(\alpha) - \int u(t_D(a)) dH(a) \right\} \end{aligned}$$

#### Home

The continuation payoff for home from the treaty being in effect is:

$$\begin{aligned} \chi_C &= \int [a u(t_E(a)) - t_E(a)] dH(a) - \sigma \int_{\alpha_S}^{\alpha_D} L^*(t_E(a)) dH(a) - \int u(\tau_E(\alpha)) dF(\alpha) \\ &\quad + \sigma \int_{\alpha_S}^{\alpha_D} L(\tau_E(\alpha)) dF(\alpha) + \delta H(a_D) F(\alpha_D) \beta \chi_C + \delta [1 - H(a_D) F(\alpha_D)] \chi_N \\ &= \frac{\Psi}{1 - \delta \beta H(a_D) F(\alpha_D)} \end{aligned}$$

$$\begin{aligned} \text{where } \Psi &= \int [a u(t_E(a)) - t_E(a)] dH(a) - \sigma \int_{\alpha_S}^{\alpha_D} L^*(t_E(a)) dH(a) - \int u(\tau_E(\alpha)) dF(\alpha) \\ &\quad + \sigma \int_{\alpha_S}^{\alpha_D} L(\tau_E(\alpha)) dF(\alpha) + \delta [1 - H(a_D) F(\alpha_D)] \chi_N \end{aligned}$$

#### Foreign

The continuation payoff for foreign from the treaty being in effect is:

$$\begin{aligned}
\chi_C^* &= \int [\alpha u(\tau_E(\alpha)) - \tau_E(\alpha)] dF(\alpha) - \sigma \int_{\alpha_S}^{\alpha_D} L(\tau_E(\alpha)) dF(\alpha) - \int u(t_E(a)) dH(a) \\
&\quad + \sigma \int_{a_S}^{a_D} L^*(t_E(a)) dH(a) + \delta H(a_D) F(\alpha_D) \beta \chi_C^* + \delta [1 - H(a_D) F(\alpha_D)] \chi_N^* \\
&= \frac{\Psi^*}{1 - \delta \beta H(a_D) F(\alpha_D)} \\
\text{where } \Psi^* &= \int [\alpha u(\tau_E(\alpha)) - \tau_E(\alpha)] dF(\alpha) - \sigma \int_{\alpha_S}^{\alpha_D} L(\tau_E(\alpha)) dF(\alpha) - \int u(t_E(a)) dH(a) \\
&\quad + \sigma \int_{a_S}^{a_D} L^*(t_E(a)) dH(a) + \delta [1 - H(a_D) F(\alpha_D)] \chi_N^*
\end{aligned}$$

## 2.4 Comparative Statics

### Full Compliance

Recall that the home country does not violate its binding if  $a \leq a_S = \frac{1}{u'(t_B)} + \sigma$ .

$$\frac{\partial a_S}{\partial t_B} = \frac{-u''(t_B)}{[u'(t_B)]^2} > 0 \quad \text{and} \quad \frac{\partial a_S}{\partial \sigma} = 1 > 0$$

Recall that the foreign country does not violate its binding if  $\alpha \leq \alpha_S = \frac{1}{u'(\tau_B)} + \sigma$ .

$$\frac{\partial \alpha_S}{\partial \tau_B} = \frac{-u''(\tau_B)}{[u'(\tau_B)]^2} > 0 \quad \text{and} \quad \frac{\partial \alpha_S}{\partial \sigma} = 1 > 0$$

### Stability

The cutpoints  $(a_D, \alpha_D)$  are implicitly defined by the system of equations:

$$\begin{aligned}
\lambda(a_D, \alpha_D) &= 0 \\
\lambda^*(a_D, \alpha_D) &= 0
\end{aligned}$$

By Cramer's Rule:

$$\frac{\partial a_D}{\partial t_B} = \frac{\begin{vmatrix} -\lambda_{t_B} & \lambda_{\alpha_D} \\ -\lambda_{t_B}^* & \lambda_{\alpha_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{a_D} & \lambda_{\alpha_D} \\ \lambda_{a_D}^* & \lambda_{\alpha_D}^* \end{vmatrix}} \quad \text{and} \quad \frac{\partial a_D}{\partial \sigma} = \frac{\begin{vmatrix} -\lambda_{\sigma} & \lambda_{\alpha_D} \\ -\lambda_{\sigma}^* & \lambda_{\alpha_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{a_D} & \lambda_{\alpha_D} \\ \lambda_{a_D}^* & \lambda_{\alpha_D}^* \end{vmatrix}}$$

$$\frac{\partial \alpha_D}{\partial \tau_B} = \frac{\begin{vmatrix} -\lambda_{\tau_B} & \lambda_{a_D} \\ -\lambda_{\tau_B}^* & \lambda_{a_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{\alpha_D} & \lambda_{a_D} \\ \lambda_{\alpha_D}^* & \lambda_{a_D}^* \end{vmatrix}} \quad \text{and} \quad \frac{\partial \alpha_D}{\partial \sigma} = \frac{\begin{vmatrix} -\lambda_{\sigma} & \lambda_{a_D} \\ -\lambda_{\sigma}^* & \lambda_{a_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{\alpha_D} & \lambda_{a_D} \\ \lambda_{\alpha_D}^* & \lambda_{a_D}^* \end{vmatrix}}$$

where:

$$\begin{aligned} \lambda_{a_D} &= u(t_S(a_D)) - u(t_D(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial a_D} \\ \lambda_{\alpha_D} &= \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \alpha_D} \right) + \delta f(\alpha_D) (\beta \chi_C - \chi_N) \\ \lambda_{t_B} &= \sigma u'(t_B) + \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial t_B} \right) \\ \lambda_{\tau_B} &= \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \tau_B} \right) \\ \lambda_{\sigma} &= -L^*(t_S(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial \sigma} \end{aligned}$$

$$\begin{aligned} \lambda_{a_D}^* &= \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial a_D} \right) + \delta h(a_D) (\beta \chi_C^* - \chi_N^*) \\ \lambda_{\alpha_D}^* &= u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \alpha_D} \\ \lambda_{t_B}^* &= \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial t_B} \right) \\ \lambda_{\tau_B}^* &= \sigma u'(\tau_B) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \tau_B} \\ \lambda_{\sigma}^* &= -L(\tau_S(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \sigma} \end{aligned}$$

#### Comparative statics on $a_D$

As  $h(a_D)$ ,  $H(a_D)$ ,  $f(\alpha_D)$ ,  $F(\alpha_D)$  grow small:

$$\begin{aligned} \begin{vmatrix} \lambda_{a_D} & \lambda_{\alpha_D} \\ \lambda_{a_D}^* & \lambda_{\alpha_D}^* \end{vmatrix} &= \lambda_{a_D} \lambda_{\alpha_D}^* - \lambda_{\alpha_D} \lambda_{a_D}^* \\ &= \left[ u(t_S(a_D)) - u(t_D(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial a_D} \right] \times \left[ u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \alpha_D} \right] \\ &\quad - \left[ \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \alpha_D} \right) + \delta f(\alpha_D) (\beta \chi_C - \chi_N) \right] \times \left[ \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial a_D} \right) + \delta h(a_D) (\beta \chi_C^* - \chi_N^*) \right] \\ &\rightarrow [u(t_S(a_D)) - u(t_D(a_D))] \times [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))] > 0 \end{aligned}$$

$$\begin{aligned}
\begin{vmatrix} -\lambda_{t_B} & \lambda_{\alpha_D} \\ -\lambda_{t_B}^* & \lambda_{\alpha_D}^* \end{vmatrix} &= -\lambda_{t_B} \lambda_{\alpha_D}^* + \lambda_{\alpha_D} \lambda_{t_B}^* \\
&= -\left[ \sigma u'(t_B) + \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial t_B} \right) \right] \times \left[ u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \alpha_D} \right] \\
&\quad + \left[ \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \alpha_D} \right) + \delta f(\alpha_D) (\beta \chi_C - \chi_N) \right] \times \left[ \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial t_B} \right) \right] \\
&\rightarrow -\sigma u'(t_B) [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))] > 0
\end{aligned}$$

$$\begin{aligned}
\begin{vmatrix} -\lambda_\sigma & \lambda_{\alpha_D} \\ -\lambda_\sigma^* & \lambda_{\alpha_D}^* \end{vmatrix} &= -\lambda_\sigma \lambda_{\alpha_D}^* + \lambda_{\alpha_D} \lambda_\sigma^* \\
&= -\left[ -L^*(t_S(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial \sigma} \right] \times \left[ u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \alpha_D} \right] \\
&\quad + \left[ \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \alpha_D} \right) + \delta f(\alpha_D) (\beta \chi_C - \chi_N) \right] \times \left[ -L(\tau_S(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \sigma} \right] \\
&\rightarrow L^*(t_S(a_D)) [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))] < 0
\end{aligned}$$

So:

$$\frac{\partial a_D}{\partial t_B} = \frac{\begin{vmatrix} -\lambda_{t_B} & \lambda_{\alpha_D} \\ -\lambda_{t_B}^* & \lambda_{\alpha_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{a_D} & \lambda_{\alpha_D} \\ \lambda_{a_D}^* & \lambda_{\alpha_D}^* \end{vmatrix}} \rightarrow \frac{-\sigma u'(t_B) [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))]}{[u(t_S(a_D)) - u(t_D(a_D))] \times [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))]} > 0$$

$$\frac{\partial a_D}{\partial \sigma} = \frac{\begin{vmatrix} -\lambda_\sigma & \lambda_{\alpha_D} \\ -\lambda_\sigma^* & \lambda_{\alpha_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{a_D} & \lambda_{\alpha_D} \\ \lambda_{a_D}^* & \lambda_{\alpha_D}^* \end{vmatrix}} \rightarrow \frac{L^*(t_S(a_D)) [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))]}{[u(t_S(a_D)) - u(t_D(a_D))] \times [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))]} < 0$$

### Comparative Statics on $\alpha_D$

As  $h(a_D)$ ,  $H(a_D)$ ,  $f(\alpha_D)$ ,  $F(\alpha_D)$  grow small:

$$\begin{aligned}
\begin{vmatrix} \lambda_{\alpha_D} & \lambda_{a_D} \\ \lambda_{\alpha_D}^* & \lambda_{a_D}^* \end{vmatrix} &= \lambda_{\alpha_D} \lambda_{a_D}^* - \lambda_{a_D} \lambda_{\alpha_D}^* \\
&= \left[ \delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \alpha_D} \right) + \delta f(\alpha_D) (\beta \chi_C - \chi_N) \right] \times \left[ \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial a_D} \right) + \delta h(a_D) (\beta \chi_C^* - \chi_N^*) \right] \\
&\quad - \left[ u(t_S(a_D)) - u(t_D(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial a_D} \right] \times \left[ u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \alpha_D} \right] \\
&\rightarrow -[u(t_S(a_D)) - u(t_D(a_D))] \times [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))] < 0
\end{aligned}$$

$$\begin{aligned}
\begin{vmatrix} -\lambda_{\tau_B} & \lambda_{a_D} \\ -\lambda_{\tau_B}^* & \lambda_{a_D}^* \end{vmatrix} &= -\lambda_{\tau_B} \lambda_{a_D}^* + \lambda_{a_D} \lambda_{\tau_B}^* \\
&= -\delta F(\alpha_D) \beta \left( \frac{\partial \chi_C}{\partial \tau_B} \right) \times \left[ \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial a_D} \right) + \delta h(a_D) (\beta \chi_C^* - \chi_N^*) \right] \\
&\quad + \left[ u(t_S(a_D)) - u(t_D(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial a_D} \right] \times \left[ \sigma u'(\tau_B) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \tau_B} \right] \\
&\rightarrow \sigma u'(\tau_B) [u(t_S(a_D)) - u(t_D(a_D))] < 0
\end{aligned}$$

$$\begin{aligned}
\begin{vmatrix} -\lambda_{\sigma} & \lambda_{a_D} \\ -\lambda_{\sigma}^* & \lambda_{a_D}^* \end{vmatrix} &= -\lambda_{\sigma} \lambda_{a_D}^* + \lambda_{a_D} \lambda_{\sigma}^* \\
&= - \left[ -L^*(t_S(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial \sigma} \right] \times \left[ \delta H(a_D) \beta \left( \frac{\partial \chi_C^*}{\partial a_D} \right) + \delta h(a_D) (\beta \chi_C^* - \chi_N^*) \right] \\
&\quad + \left[ u(t_S(a_D)) - u(t_D(a_D)) + \delta F(\alpha_D) \beta \frac{\partial \chi_C}{\partial a_D} \right] \times \left[ -L(\tau_S(\alpha_D)) + \delta H(a_D) \beta \frac{\partial \chi_C^*}{\partial \sigma} \right] \\
&\rightarrow -L(\tau_S(\alpha_D)) [u(t_S(a_D)) - u(t_D(a_D))] > 0
\end{aligned}$$

So:

$$\frac{\partial \alpha_D}{\partial \tau_B} = \frac{\begin{vmatrix} -\lambda_{\tau_B} & \lambda_{a_D} \\ -\lambda_{\tau_B}^* & \lambda_{a_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{\alpha_D} & \lambda_{a_D} \\ \lambda_{\alpha_D}^* & \lambda_{a_D}^* \end{vmatrix}} \rightarrow \frac{\sigma u'(\tau_B) [u(t_S(a_D)) - u(t_D(a_D))]}{-[u(t_S(a_D)) - u(t_D(a_D))] \times [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))]} > 0$$

So:

$$\frac{\partial \alpha_D}{\partial \sigma} = \frac{\begin{vmatrix} -\lambda_{\sigma} & \lambda_{a_D} \\ -\lambda_{\sigma}^* & \lambda_{a_D}^* \end{vmatrix}}{\begin{vmatrix} \lambda_{\alpha_D} & \lambda_{a_D} \\ \lambda_{\alpha_D}^* & \lambda_{a_D}^* \end{vmatrix}} \rightarrow \frac{-L(\tau_S(\alpha_D)) [u(t_S(a_D)) - u(t_D(a_D))]}{-[u(t_S(a_D)) - u(t_D(a_D))] \times [u(\tau_S(\alpha_D)) - u(\tau_D(\alpha_D))]} < 0$$