Abstract

In this Online Appendix, we elaborate on the material in section 5 of the main text. In particular, we present two extensions of our baseline model in which we analyze the optimal remedy and remedial period in trade and investment agreements and prove Propositions 4 and 5.
In this Online Appendix, we elaborate on the material in section 5 of the main text. In particular, we present two extensions of our baseline model in which we analyze the optimal remedy and remedial period in trade and investment agreements and prove Propositions 4 and 5.

1. Nature of the Remedy

Here we prove Proposition 4. To this end and as noted in the body of the paper, rather than simply assuming that the Home or Host government switches to $\tau = FT$ or $\iota = FT$ when convicted by the DSB as in our baseline models of sections 3 and 4 in the main text, we now allow the government to choose instead to maintain $\tau = P$ or $\iota = T$ and make damage payments.

We compare two forms of damage payment: one in which the court allows the injured party to engage in reciprocal retaliation, and another in which the injured party is awarded cash damages by the court. The key trade-off featured by our extended models is that retaliation is less efficient but that cash damages can be difficult to assess. Which remedy is optimal in a given setting then depends on which force is stronger in that setting.

To make our points on the nature of the remedy as clearly as possible, we build on our analysis of standing and now take as given that trade agreements limit standing to governments while investment agreements also afford standing to investors, and we adopt two further simplifying assumptions. First, when retaliation is the remedy we assume that retaliation is sufficiently inefficient and costly that the Home or Host government switches to $\tau = FT$ or $\iota = FT$ when convicted by the DSB to avoid retaliation - this means that the analysis of trade and investment agreements with retaliation is exactly the same as in our baseline analysis of these agreements in sections 3 and 4 in the main text (with the trade agreement featuring SSDS and the investment treaty also including ISDS). Second, when cash payments are the remedy, we assume that cash payments are perfectly efficient so that surplus can be costlessly transferred internationally.

With these two simplifying assumptions we adopt an extreme position on the inefficiency of retaliation relative to cash, so that we can focus our analysis of the optimal remedy in trade and investment agreements on the degree of difficulty faced by the court in assessing damages in each setting.
1.1. Trade Agreements

We first consider the choice of retaliation versus cash payments in the context of trade agreements. Our task is to introduce cash damages into our model of trade agreements with SSDS, and to compare the outcomes from our earlier model (which we now refer to as the outcomes under the $V_R$ institution) to the outcomes under cash damages (which we refer to as the outcomes under the $V_C$ institution).

To capture the notion that the court may struggle to accurately assess cash damages, we assume that the court-assessed damages are realizations of a random variable. Denoting the damages awarded to the injured foreign government by $d^*(s)$, we assume $\Pr [d^*(s) = |\gamma^*_G(s)|] = 1 - 2m(s)$, $\Pr [d^*(s) > |\gamma^*_G(s)|] = m(s)$, and $\Pr [d^*(s) < |\gamma^*_G(s)|] = m(s)$, with $m(s) \in [0, \frac{1}{2}]$. Hence, the court awards the correct damages with probability $1 - 2m(s)$ and overestimates or underestimates the damages with symmetric probabilities $m(s)$. For simplicity, we also assume that court mistakes are sufficiently severe in the sense that $d^*(s) > \gamma_G(s)$ if $d^*(s) > |\gamma^*_G(s)|$ and $d^*(s) < \gamma_G(s)$ if $d^*(s) < |\gamma^*_G(s)|$.

Under these assumptions, Home’s reaction to a DSB ruling of $\tau^{DSB} = FT$ depends on the level of damages. If the DSB underestimates the damages, Home always continues the violation $\tau = P$ since then $d^*(s) < \gamma_G(s)$. Conversely, if the DSB overestimates the damages, Home always switches to $\tau = FT$ since then $d^*(s) > \gamma_G(s)$. And if the DSB makes the correct damage assessment, Home switches to $\tau = FT$ in states $s \in \sigma^{FT}$ (since then $\gamma_G(s) < |\gamma^*_G(s)| = d^*(s)$) and continues the violation $\tau = P$ in states $s \in \sigma^P$ (since then $\gamma_G(s) > |\gamma^*_G(s)| = d^*(s)$). Essentially, a correct damage assessment makes Home internalize the effects of its policy choice on Foreign and therefore choose the efficient policy.\(^1\)

Now that we have understood how cash damages change Home’s behavior following a DSB ruling of $\tau^{DSB} = FT$, we can characterize Foreign’s filing choice and Home’s policy choice under the $V_C$ institution. Foreign files a complaint if $\tau = P$ and the expected benefits from litigation exceed the costs of litigation. As in the baseline model, Foreign only benefits from litigation if the court rules $\tau^{DSB} = FT$ because otherwise Home simply continues with

\(^1\)Notice also that the accuracy with which the court assesses damages ($d^*(s)$), as parameterized by $m(s)$, is distinct from the accuracy of the court ruling ($FT$ or $P$), as parameterized by $q_k(s)$. Thus, for example, the court might be good at determining whether or not the imposition of protection was warranted in a particular state of the world (e.g., Does protection preserve more jobs in the Home country than it destroys in the Foreign country?) but bad at assessing the value of the harm done to the foreign government (e.g., What is the monetary value of a job?).
\( \tau = P \). However, there is now a positive probability \( \Pr (\tau = P \mid \text{ruling is } FT, s) \) that Home defies the DSB’s ruling and instead chooses to pay damages \( d^* (s) \). Introducing the shorthands \( \Pr (\text{defy}) \equiv \Pr (\tau = P \mid \text{ruling is } FT, s) \), \( \Pr (\text{comply}) \equiv \Pr (\tau = FT \mid \text{ruling is } FT, s) \), and \( \tilde{d}^* (s) \equiv E [d^* (s) \mid \tau = P, \tau^{DSB} = FT, s] \), we can therefore write Foreign’s expected benefits from litigation as \( \Pr (\text{ruling is } FT \mid s) \times \left[ \Pr (\text{comply}) \cdot |\gamma_{G^*} (s)| + \Pr (\text{defy}) \cdot \tilde{d}^* (s) \right] \). Defining the ratio \( \mu_C^* (s) \equiv \frac{c^* (s)}{\Pr (\text{comply}) \cdot |\gamma_{G^*} (s)| + \Pr (\text{defy}) \cdot \tilde{d}^* (s)} \), we can thus summarize that Foreign files a complaint if and only if \( \tau = P \) and

\[
\Pr (\text{ruling is } FT \mid s) > \mu_C^* (s).
\] (OA 1.1)

As in the baseline model, Home chooses \( \tau = P \) if either condition (OA 1.1) fails or the expected benefits from litigation outweigh the costs of litigation. However, we now need to account for the possibility that Home can choose to defy a DSB ruling of \( \tau^{DSB} = FT \) when calculating the expected benefits from litigation; with this accounting, the expected benefits for Home of choosing \( \tau = P \) when (OA 1.1) holds are given by \( \Pr (\text{ruling is } P \mid s) \cdot \gamma_G (s) + \Pr (\text{ruling is } FT \mid s) \cdot \Pr (\text{defy}) \cdot \left[ \gamma_G (s) - \tilde{d}^* (s) \right] \). Defining \( \mu_C (s) \equiv \frac{c(s) - \Pr (\text{defy}) \cdot |\gamma_G (s) - \tilde{d}^* (s)|}{\gamma_G (s) - \Pr (\text{defy}) \cdot |\gamma_G (s) - \tilde{d}^* (s)|} \), we can therefore summarize that Home chooses \( \tau = P \) if either condition (OA 1.1) fails or if (OA 1.1) holds and:

\[
\Pr (\text{ruling is } P \mid s) > \mu_C (s).
\] (OA 1.2)

Notice that conditions (OA 1.1) and (OA 1.2) are identical to the corresponding conditions (3.1) and (3.2) in the main text, up to the definition of \( \mu_C (s) \) and \( \mu_C^* (s) \). Assuming again that dispute costs are low relative to dispute stakes in the sense that \( \mu_C (s) + \mu_C^* (s) < 1 \) for all \( s \), we can therefore characterize equilibrium actions just as in Lemma 1 in the main text:

**Lemma OA 1.** **Equilibrium actions under the \( V_C \) institution are as follows:**

1. **In states** \( s \in \sigma^{FT} \): If DSB quality is high in the sense that \( q_k (s) \leq \mu_C (s) \), we have \( \tau = FT \) and no dispute; if DSB quality is intermediate in the sense that \( q_k (s) \in (\mu_C (s), 1 - \mu_C^* (s)) \), we have \( \tau = P \) and a dispute; if DSB quality is low in the sense that \( q_k (s) \geq 1 - \mu_C^* (s) \), we have \( \tau = P \) and no dispute.

---

\(^2\)Note that these expressions have to be evaluated according to the considered state. In states \( s \in \sigma^{FT} \), Home only defies a DSB ruling of \( \tau^{DSB} = FT \) if the DSB also underestimates the damages so that \( \Pr (\text{defy}) = m(s) \) and \( \tilde{d}^* (s) = E [d^* (s) \mid d^* (s) < |\gamma_{G^*} (s)|] \equiv \tilde{d}_{\text{low}} (s) \). In states \( s \in \sigma^P \), Home instead also defies a DSB ruling of \( \tau^{DSB} = FT \) if the DSB correctly assesses the damages so that \( \Pr (\text{defy}) = 1 - m(s) \) and \( \tilde{d}^* (s) = \frac{m(s)}{1 - m(s)} \tilde{d}_{\text{low}} (s) + \frac{1 - 2m(s)}{1 - m(s)} |\gamma_{G^*} (s)| \).
2. In states \( s \in \sigma^P \): If DSB quality is high in the sense that \( q_k (s) \leq \mu_C^* (s) \), we have \( \tau = P \) and no dispute; if DSB quality is intermediate in the sense that \( q_k (s) \in (\mu_C^* (s), 1 - \mu_C (s)) \), we have \( \tau = P \) and a dispute; if DSB quality is low in the sense that \( q_k (s) \geq 1 - \mu_C (s) \), we have \( \tau = FT \) and no dispute.

For future reference, we denote the different action sets by

\[
\begin{align*}
\sigma_{1,C}^{FT} & \equiv \{ s \in \sigma^{FT} \mid q_k (s) \leq \mu_C (s) \}, \\
\sigma_{2,C}^{FT} & \equiv \{ s \in \sigma^{FT} \mid q_k (s) \in (\mu_C (s), 1 - \mu_C^* (s)) \}, \text{ and} \\
\sigma_{3,C}^{FT} & \equiv \{ s \in \sigma^{FT} \mid q_k (s) \geq 1 - \mu_C^* (s) \},
\end{align*}
\]

as well as

\[
\begin{align*}
\sigma_{1,C}^{P} & \equiv \{ s \in \sigma^P \mid q_k (s) \leq \mu_C^* (s) \}, \\
\sigma_{2,C}^{P} & \equiv \{ s \in \sigma^P \mid q_k (s) \in (\mu_C^* (s), 1 - \mu_C (s)) \}, \text{ and} \\
\sigma_{3,C}^{P} & \equiv \{ s \in \sigma^P \mid q_k (s) \geq 1 - \mu_C (s) \}.
\end{align*}
\]

We can now write down the efficiency loss associated with the \( V_C \) institution, \( L (V_C) \), relative to the first-best outcome:

\[
L (V_C) = \sum_{s \in \sigma_{2,C}^{FT} \cup \sigma_{1,C}^{P}} p (s) q_k (s) \mid \Gamma (s) \mid + \sum_{s \in \sigma_{2,C}^{FT} \cup \sigma_{1,C}^{P}} p (s) [c (s) + c^* (s)] + \sum_{s \in \sigma_{2,C}^{FT} \cup \sigma_{3,C}^{P}} p (s) \mid \Gamma (s) \mid + \sum_{s \in \sigma_{1,C}^{FT}} p (s) [1 - q_k (s)] m (s) \mid \Gamma (s) \mid - \sum_{s \in \sigma_{2,C}^{P}} p (s) q_k (s) \mid 1 - m (s) \mid \Gamma (s) \mid.
\]

There are two main changes relative to our baseline analysis (and thus two new terms in equation (OA 1.3) relative to equation (3.3) in the main text). First, if the court correctly rules \( FT \) there is now still a probability \( m (s) \) that it awards excessively low damages in which case Home will continue to choose \( \tau = P \), leading to an additional efficiency loss (the first term on the last line in equation (OA 1.3)). Second, if the court incorrectly rules \( FT \) there is still a probability \( 1 - m (s) \) that it does not award excessively high damages in which case Home will
continue to choose \( \tau = P \), leading to an additional efficiency gain (the second term on the last line in equation (OA 1.3)).

Finally, as noted above the equilibrium actions under the \( V_R \) institution are exactly the same as in the \( V_{G^*} \) institution of our baseline model, and hence the efficiency loss associated with the \( V_R \) institution, \( L(V_R) \), is exactly the same as in the \( V_{G^*} \) institution of our baseline model. To aid comparison, we use the notation \( \sigma_{2,R}^{FT} \equiv \sigma_{2,G^*}^{FT} \), \( \sigma_{2,R}^P \equiv \sigma_{2,G^*}^P \), \( \sigma_{3,R}^{FT} \equiv \sigma_{3,G^*}^{FT} \) and \( \sigma_{3,R}^P \equiv \sigma_3^P \) where the latter sets are defined in our baseline model. With this we then have

\[
L(V_R) = \sum_{s \in \sigma_{2,R}^{FT} \cup \sigma_{2,R}^P} p(s) qk(s) |\Gamma(s)| + \sum_{s \in \sigma_{3,R}^{FT} \cup \sigma_{3,R}^P} p(s)[c(s) + e^*(s)] + \sum_{s \in \sigma_{3,R}^{FT} \cup \sigma_{3,R}^P} p(s)|\Gamma(s)|.
\]

1.1.1. Cash versus Retaliation

By comparing \( L(V_R) \) and \( L(V_C) \), we can characterize the conditions under which the \( V_R \) institution is preferred to the \( V_C \) institution. Since the \( V_R \) institution mirrors the \( V_{G^*} \) institution from the baseline model, all action thresholds are also exactly the same and we now describe them using the notation \( \mu_R^*(s) \equiv \mu_{G^*}^*(s) \) and \( \mu_R(s) \equiv \mu(s) \) to aid comparison with the action thresholds associated with the \( V_C \) institution.

As is easy to verify, the action thresholds associated with the \( V_C \) institution are below (above) the action thresholds associated with the \( V_R \) institution in states \( s \in \sigma^{FT} \) (states \( s \in \sigma^P \)). Intuitively, the \( V_C \) institution gives the Home government the option to ignore the cease and desist order which makes it more attractive for the Home government to protect and less attractive for the Foreign government to litigate. In states \( s \in \sigma^{FT} \), the Home government therefore starts protecting earlier \( (\mu_C(s) < \mu_R(s)) \) and the Foreign government gives up litigating earlier \( (1 - \mu_C(s) < 1 - \mu_R(s)) \) under the \( V_C \) institution; and in states \( s \in \sigma^P \), the Foreign government starts litigating later \( (\mu_C^*(s) > \mu_R(s)) \) and the Home government gives up protecting later \( (1 - \mu_C(s) > 1 - \mu_R(s)) \) under the \( V_C \) institution.

Defining the sets

\[
\Omega_{1,C}^{FT} \equiv \{ s \in \sigma^{FT} \mid qk(s) \in [\mu_C(s), \mu_R(s)] \}, \\
\Omega_{2,C}^{FT} \equiv \{ s \in \sigma^{FT} \mid qk(s) \in (1 - \mu_C^*(s), 1 - \mu_R^*(s)) \}, \\
\Omega_{1,C}^P \equiv \{ s \in \sigma^P \mid qk(s) \in (\mu_R^*(s), \mu_C^*(s)) \}, \text{ and} \\
\Omega_{2,C}^P \equiv \{ s \in \sigma^P \mid qk(s) \in (1 - \mu_R(s), 1 - \mu_C(s)) \},
\]

we can then write:
\[
L(V_R) - L(V_C) = - \sum_{s \in \sigma_{2,C}^{FT}} p(s) [1 - q_k(s)] m(s) |\Gamma(s)| \\
+ \sum_{s \in \sigma_{2,C}^P} p(s) q_k(s) [1 - m(s)] |\Gamma(s)| \\
- \sum_{s \in \Omega_{1,C}^{FT}} p(s) [q_k(s) |\Gamma(s)| + c(s) + c^*(s)] \\
+ \sum_{s \in \Omega_{1,C}^P} p(s) [q_k(s) |\Gamma(s)| + c(s) + c^*(s)] \\
+ \sum_{s \in \Omega_{2,C}^{FT}} p(s) [q_k(s) |\Gamma(s)| + c(s) + c^*(s) - |\Gamma(s)|] \\
+ \sum_{s \in \Omega_{2,C}^P} p(s) [|\Gamma(s)| - q_k(s) |\Gamma(s)| - c(s) - c^*(s)].
\]

Each line in this expression has an intuitive interpretation. The first line captures a disadvantage of the \(V_C\) institution, which is that correct court rulings are effectively overturned in states \(s \in \sigma_{2,C}^{FT}\) if the court underestimates the damages. The second line then summarizes the corresponding advantage of the \(V_C\) institution, which is that incorrect court rulings are effectively overturned in states \(s \in \sigma_{2,C}^P\) unless the court overestimates the damages. All other lines capture effects associated with movements in the action thresholds. In states \(s \in \sigma^{FT}\), the inefficient policy is introduced earlier (i.e., for lower \(q_k(s)\) and therefore higher court quality) under the \(V_C\) institution, while in states \(s \in \sigma^P\), the efficient policy is challenged later (i.e., for higher \(q_k(s)\) and therefore lower court quality) under the \(V_C\) institution, which is captured in the third and fourth line, respectively. In states \(s \in \sigma^{FT}\), litigation against the inefficient policy is given up earlier under the \(V_C\) institution, while in states \(s \in \sigma^P\), the efficient policy is given up later under the \(V_C\) institution, which is captured in the fifth and sixth line, respectively.

Two observations allow us to describe intuitive conditions under which expression (OA 1.4) can be signed. The first observation is that an improvement in the court’s ability to assess damages tends to favor the \(V_C\) institution. To see this, notice that the sets \(\Omega_{1,C}^{FT}, \Omega_{2,C}^{FT}\), and \(\Omega_{1,C}^P\) become empty as \(m(s) \to 0\) so that expression (OA 1.4) reduces to:

\[
L(V_R) - L(V_C) \big|_{m(s) \to 0} = \sum_{s \in \sigma_{2,C}^P} p(s) q_k(s) |\Gamma(s)| \\
+ \sum_{s \in \Omega_{2,C}^P} p(s) [|\Gamma(s)| - q_k(s) |\Gamma(s)| - c(s) - c^*(s)].
\]
Correctly assessed cash damages make the Home government fully internalize the effect its trade policy choices have on the Foreign government, so that it stops protecting when (rightly) convicted in states $s \in \sigma_{2,C}^P$ and continues protecting when (wrongly) convicted in states $s \in \sigma_{2,C}^P$. As a result, there are no longer any inefficient breaches of the court’s cease and desist order, and the welfare gains associated with the efficient breaches are captured in the first line of the expression above. However, the $V_C$ institution also brings about additional litigation in states $s \in \Omega_{2,C}^P$, which could be socially desirable or not. This is captured in the second line of the above expression, which compares the welfare losses associated with the $V_R$ institution in states $s \in \Omega_{2,C}^P$ (given by $|\Gamma(s)|$) to the expected welfare losses associated with the $V_C$ institution in these states (given by $qk(s)|\Gamma(s)| + c(s) + c^*(s)$). Notice that the second line is positive if the joint litigation costs $c(s) + c^*(s)$ are sufficiently small. Overall, the $V_C$ institution therefore dominates the $V_R$ institution if the court’s ability to assess damages is sufficiently good, assuming that the joint litigation costs are sufficiently small.

The second observation is that an increase in the probability that free trade is the efficient policy tends to favor the $V_R$ institution. To see this, notice that the sets $\sigma_{2,C}^P$, $\Omega_{1,C}^F$, and $\Omega_{2,C}^P$ become empty as $\sum_{s \in \sigma_{FT}} p(s) \to 1$ so that expression (OA 1.4) reduces to:

$$L(V_R) - L(V_C) = \frac{1}{m(s)} \left( \sum_{s \in \sigma_{2,C}^P} - \sum_{s \in \Omega_{1,C}^F} \right) p(s) \left[ 1 - qk(s) \right] m(s) |\Gamma(s)|$$

$$+ \sum_{s \in \Omega_{2,C}^P} p(s) \left[ qk(s) |\Gamma(s)| + c(s) + c^*(s) - |\Gamma(s)| \right].$$

If free trade is the efficient policy, allowing the Home government to ignore the court’s cease and desist ruling by paying cash damages brings about a welfare loss, and this welfare loss is captured in the first line of the expression above. The other two lines summarize the welfare effects associated with the more aggressive protectionism of the Home government in states $s \in \Omega_{1,C}^F$ and $s \in \Omega_{2,C}^P$. While the second term is unambiguously negative (the efficient policy gets revoked earlier under the cash institution), the third term can be positive or negative (the additional litigation under the cash institution may be socially desirable or not), and becomes negative if the joint litigation costs $c(s) + c^*(s)$ are sufficiently small. Overall, the $V_R$ institution

---

3 Notice, however, that even when cash damages are assessed with perfect accuracy, the $V_C$ institution still delivers the inefficient policy choice $\tau = FT$ in states $s \in \sigma_{3,C}^F$ as well as $\tau = P$ in states $s \in \sigma_{3,C}^P$, where court quality is so bad that the court is not even invoked.
therefore dominates the \( V_C \) institution if free trade is sufficiently likely to be the efficient trade policy, assuming that the joint litigation costs are sufficiently small.

Imposing that the joint litigation costs are sufficiently small, we can therefore state:

**Proposition 4 (i).** Allowing for retaliation instead of cash damages in a trade agreement is optimal \((1)\) if the DSB’s ability to assess cash damages is sufficiently bad, and \((2)\) free trade is sufficiently likely to be the efficient policy choice.

### 1.2. Investment Agreements

We now consider the choice of retaliation versus cash payments in the context of investment treaties. Our task is to introduce cash damages into our model of investment treaties with SSDS and ISDS, and compare the outcomes from our earlier model to the outcomes under cash damages derived here.

We parametrize damage payments as \( \tilde{d}^*(I^*, s) = \tilde{d}^*(s) PS(I^*, s) \), in keeping with our assumption in the main text that the costs and benefits of litigation scale with \( PS(I^*, s) \). Following our earlier structure, we then assume that the DSB correctly chooses \( \tilde{d}^*(s) \) with probability \( 1 - 2m(s) \), \( \Pr(\tilde{d}^*(s) = \tilde{\gamma}^*_s (s)) = \Pr(\tilde{d}^*(I^*, s) = \tilde{\gamma}^*_s (I^*, s)) = 1 - 2m(s) \), underestimates \( \tilde{d}^*(s) \) with probability \( m(s) \), \( \Pr(\tilde{d}^*(s) < \tilde{\gamma}^*_s (s)) = \Pr(\tilde{d}^*(I^*, s) < \tilde{\gamma}^*_s (I^*, s)) = m(s) \), and overestimates \( \tilde{d}^*(s) \) with probability \( m(s) \), \( \Pr(\tilde{d}^*(s) > \tilde{\gamma}^*_s (s)) = \Pr(\tilde{d}^*(I^*, s) > \tilde{\gamma}^*_s (I^*, s)) = m(s) \). We also assume that court mistakes are sufficiently severe in the sense that \( \tilde{d}^*(I^*, s) > \tilde{\gamma}^*_G (I^*, s) \) if \( \tilde{d}^*(I^*, s) > \tilde{\gamma}^*_G (I^*, s) \) and \( \tilde{d}^*(I^*, s) < \tilde{\gamma}^*_G (I^*, s) \) if \( \tilde{d}^*(I^*, s) < \tilde{\gamma}^*_G (I^*, s) \).

Keeping in mind that we consider the case with SSDS and ISDS, it is now helpful to define the thresholds \( \bar{\mu}_C^*(s) \equiv \min \left\{ \frac{c^*_G(s)}{\Pr(\text{comply}) - \tilde{\gamma}^*_G(s) - \tilde{d}^*(s)} + \frac{\tilde{d}^*(s)}{\Pr(\text{comply}) - \tilde{\gamma}^*_G(s) - \tilde{d}^*(s)} \right\} \) and \( \bar{\mu}_C^*(s) \equiv \frac{\tilde{d}^*(s) - \Pr(\text{comply}) - \tilde{\gamma}^*_G(s) - \tilde{d}^*(s)}{\tilde{\gamma}^*_G(s) - \Pr(\text{comply}) - \tilde{\gamma}^*_G(s) - \tilde{d}^*(s)} \), where \( \Pr(\text{comply}) \equiv \Pr(\tau = T \mid \text{ruling is } FT, s) \), \( \Pr(\text{comply}) \equiv \Pr(\tau = FT \mid \text{ruling is } FT, s) \), and \( \tilde{d}^*(s) \equiv E \left[ \tilde{d}^*(s) \mid \tau = T, \tau^{DSB} = FT, s \right] \). Just as in the previous subsection, we can then summarize that Foreign files a complaint if and only if \( \tau = T \) and

\[
\Pr(\text{ruling is } FT \mid s) > \bar{\mu}_C^*(s). \tag{OA 1.5}
\]

If the Host government is constrained, it chooses \( \tau = \tau_{FB} \). Otherwise, it chooses \( \tau = T \) if either the above condition is violated or

\[
\Pr(\text{ruling is } T \mid s) > \bar{\mu}_C^*(s). \tag{OA 1.6}
\]
Conditions (OA 1.5) and (OA 1.6) are identical to conditions (4.9) and (4.10) from the main text, up to the definition of $\bar{\mu}_C(s)$ and $\bar{\mu}_C^*(s)$. Assuming again that dispute costs are low relative to dispute stakes in the sense that $\bar{\mu}_C(s) + \bar{\mu}_C^*(s) < 1$ for all $s$, Lemma 3 in the main text essentially carries over to this extended setting:

**Lemma OA 2.** Equilibrium actions under cash damages are as follows:

1. **In states** $s \in \sigma^{FT}$:

   1. *If the Host government is constrained:* We have $\iota = FT$ and no dispute.

   2. *If the Host government is unconstrained:* If DSB quality is high in the sense that $q_k(s) \leq \bar{\mu}_C(s)$, we have $\iota = FT$ and no dispute; if DSB quality is intermediate in the sense that $q_k(s) \in (\bar{\mu}_C(s), 1 - \bar{\mu}_C^*(s))$, we have $\iota = T$ and a dispute; if DSB quality is low in the sense that $q_k(s) \geq 1 - \bar{\mu}_C^*(s)$, we have $\iota = T$ and no dispute.

2. **In states** $s \in \sigma^{T}$:

   1. *If the Host government is constrained:* We have $\iota = T$, no dispute if $q_k(s) \leq \bar{\mu}_C^*(s)$, and a dispute if $q_k(s) > \bar{\mu}_C^*(s)$.

   2. *If the Host government is unconstrained:* If DSB quality is high in the sense that $q_k(s) \leq \bar{\mu}_C^*(s)$, we have $\iota = T$ and no dispute; if DSB quality is intermediate in the sense that $q_k(s) \in (\bar{\mu}_C^*(s), 1 - \bar{\mu}_C(s))$, we have $\iota = T$ and a dispute; if DSB quality is low in the sense that $q_k(s) \geq 1 - \bar{\mu}_C(s)$, we have $\iota = FT$ and no dispute.

For future reference, we denote the different action sets by

$$\bar{\sigma}_{1,C}^{FT} \equiv \{ s \in \bar{\sigma}^{FT} \mid q_k(s) \leq \bar{\mu}_C(s) \},$$

$$\bar{\sigma}_{2,C}^{FT} \equiv \{ s \in \bar{\sigma}^{FT} \mid q_k(s) \in (\bar{\mu}_C(s), 1 - \bar{\mu}_C^*(s)) \},$$

$$\bar{\sigma}_{3,C}^{FT} \equiv \{ s \in \bar{\sigma}^{FT} \mid q_k(s) \geq 1 - \bar{\mu}_C^*(s) \},$$

as well as

$$\bar{\sigma}_{1,C}^{T} \equiv \{ s \in \bar{\sigma}^{T} \mid q_k(s) \leq \bar{\mu}_C^*(s) \},$$

$$\bar{\sigma}_{2,C}^{T} \equiv \{ s \in \bar{\sigma}^{T} \mid q_k(s) \in (\bar{\mu}_C^*(s), 1 - \bar{\mu}_C(s)) \},$$

$$\bar{\sigma}_{3,C}^{T} \equiv \{ s \in \bar{\sigma}^{T} \mid q_k(s) \geq 1 - \bar{\mu}_C(s) \}.$$
With this lemma, we now characterize the expected welfare of the Host government under an investment treaty with cash damages, \( E_s \left[ \bar{\omega}_{G^*,I^*} (I^*, s) \right] \), using the fact that the expected return to the Foreign investor on an investment of \( I^* \) under cash damages must equal the world interest rate \( r^* \) net of the investment incentive. Following the same steps as in the Appendix of the main text, this can be shown to yield:

\[
E_s \left[ \bar{\omega}_{G^*,I^*} (I^*, s) \right] = (1 - \bar{p}) \sum_{s \in \bar{\sigma}_{2,C}^{FT}} p(s) \left[ CS (I^*, s) + PS (I^*, s) \right] \\
+ (1 - \bar{p}) \sum_{s \in \bar{\sigma}_{2,C}^{FT}} p(s) \left\{ CS (I^*, s) + PS (I^*, s) - [q_k (s) + (1 - q_k (s))m(s)]\tilde{\Gamma} (I^*, s) - \bar{c} (I^*, s) - \bar{c}^* (I^*, s) \right\} \\
+ (1 - \bar{p}) \sum_{s \in \bar{\sigma}_{2,C}^{FT}} p(s) \left[ CS (I^*, s) + PS (I^*, s) - \tilde{\Gamma} (I^*, s) \right] \\
+ (1 - \bar{p}) \sum_{s \in \bar{\sigma}_{2,C}^{FT}} p(s) \left\{ q_k (s) m(s) \left[ -\tilde{\Gamma} (I^*, s) \right] - \bar{c} (I^*, s) - \bar{c}^* (I^*, s) \right\} \\
+ (1 - \bar{p}) \sum_{s \in \bar{\sigma}_{2,C}^{FT}} p(s) \left[ -\tilde{\Gamma} (I^*, s) \right] \\
+ \bar{p} \sum_{s \in \bar{\sigma}_{2-C}^{FT}} p(s) \left[ CS (I^*, s) + PS (I^*, s) \right] \\
- \bar{p} \sum_{s \in \bar{\sigma}_{2-C}^{FT} \cup \bar{\sigma}_{3,C}^{FT}} p(s) \left[ \bar{c} (I^*, s) + \bar{c}^* (I^*, s) \right] \\
r^* I^*.
\]

There are three substantive differences between the expression in (OA 1.7) and its analog from the baseline model in (4.11). First, in states \( s \in \bar{\sigma}_{2,C}^{FT} \) there is now a higher probability of losing \( \tilde{\Gamma} (I^*, s) = (1 - \kappa) \) \( PS (I^*, s) \) in social surplus since the Host government now chooses \( \tau = T \) even if it is rightly convicted by the DSB if the DSB awards excessively low damages (the third line of both expressions). Second, in states \( s \in \bar{\sigma}_{2-C}^{FT} \), there is now a lower probability of losing \( \tilde{\Gamma} (I^*, s) = e (I^*, s) - CS (I^*, s) - PS (I^*, s) \) in social surplus since the Host government only follows an incorrect cease and desist order if the assessed damage payments are excessively high (the fourth line of both expressions). And finally, in states \( s \in \bar{\sigma}_{2-C}^{FT} \cap \bar{\sigma}_{3-C}^{FT} \) there is now a zero probability of losing \( \tilde{\Gamma} (I^*, s) = e (I^*, s) - CS (I^*, s) - PS (I^*, s) \) in social surplus if the Host government is constrained in its policy choices simply because it is now possible to continue implementing the first-best policy in case of an erroneous conviction (the seventh line of both expressions).
1.2.1. Cash versus Retaliation

We now provide sufficient conditions under which the Host government’s ex-ante expected welfare under an investment treaty with cash damages exceeds its ex-ante expected welfare under an investment treaty that relies on retaliation, with the latter given by (4.11) from the main text and which we now denote by $E_s \left[ \tilde{\omega}_{G^*, R}^R \left( \bar{I}^R, s \right) \right]$. Comparing Lemma OA 2 to Lemma 2 in the main text, it is easy to verify that the action thresholds with cash damages are lower (higher) than the action thresholds under retaliation in states $s \in \sigma_T \ (s \in \sigma_T)$ so that we can again define sets $\tilde{\Omega}^{FT}_{1, C}, \tilde{\Omega}^{FT}_{2, C}, \tilde{\Omega}^T_{1, C},$ and $\tilde{\Omega}^T_{2, C}$, using the same notation convention as above. Denoting by $\bar{I}^s$ the optimal level of investment under an ISDS with retaliation, we may then use (OA 1.7) to write

$$E_s \left[ \tilde{\omega}_{G^*, \bar{I}^*}^C \left( \bar{I}^s, s \right) \right] - E_s \left[ \tilde{\omega}_{G^*, \bar{I}^*}^R \left( \bar{I}^s, s \right) \right] = - (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^{FT}_{1, C}} p(s) [1 - q_k(s)] \cdot (1 - \kappa) \cdot PS \left( \bar{I}^s, s \right) + (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^{FT}_{2, C}} p(s) \cdot q_k(s) \cdot (1 - m(s)) \cdot [e \left( \bar{I}^s, s \right) - CS \left( \bar{I}^s, s \right) - PS \left( \bar{I}^s, s \right)] - (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^T_{1, C}} p(s) \cdot q_k(s) \cdot (1 - \kappa) \cdot PS \left( \bar{I}^s, s \right) + \bar{c} \left( \bar{I}^s, s \right) + \bar{c}^* \left( \bar{I}^s, s \right) + (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^T_{2, C}} p(s) \cdot \{1 - q_k(s) \} \cdot (1 - \kappa) \cdot PS \left( \bar{I}^s, s \right) - \bar{c} \left( \bar{I}^s, s \right) - \bar{c}^* \left( \bar{I}^s, s \right) + (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^T_{2, C}} p(s) \cdot \{1 - q_k(s) \} \cdot [e \left( \bar{I}^s, s \right) - CS \left( \bar{I}^s, s \right) - PS \left( \bar{I}^s, s \right)] - \bar{c} \left( \bar{I}^s, s \right) - \bar{c}^* \left( \bar{I}^s, s \right) + \bar{p} \sum_{s \in \tilde{\Omega}^T_{1, C}} p(s) \cdot \bar{c} \left( \bar{I}^s, s \right) + \bar{c}^* \left( \bar{I}^s, s \right) + \bar{p} \sum_{s \in \tilde{\Omega}^T_{2, C}} p(s) \cdot q_k(s) \cdot [e \left( \bar{I}^s, s \right) - CS \left( \bar{I}^s, s \right) - PS \left( \bar{I}^s, s \right)].$$

Clearly, if $E_s \left[ \tilde{\omega}_{G^*, \bar{I}^*}^C \left( \bar{I}^s, s \right) \right] - E_s \left[ \tilde{\omega}_{G^*, \bar{I}^*}^R \left( \bar{I}^s, s \right) \right] \geq 0$, it must also be that $E_s \left[ \tilde{\omega}_{G^*, \bar{I}^*}^C \left( \bar{I}^C, s \right) \right] - E_s \left[ \tilde{\omega}_{G^*, \bar{I}^*}^R \left( \bar{I}^s, s \right) \right] \geq 0$, where $\bar{I}^C$ is the level of investment that is optimal under an investment treaty with cash damages and an ISDS. For this reason (OA 1.8) allows us to establish sufficient conditions under which an investment treaty with cash damages is preferred to an investment treaty with retaliation, similar to how our earlier equation (OA 1.4) allowed us
to establish conditions under which a trade agreement with retaliation is preferred to a trade agreement with cash damages. The individual lines of both equations have a very similar interpretation, keeping in mind of course that the Host government is now constrained to choose the first-best policies with probability $p$ (the last two lines).

Just as we saw in the case of a trade agreement, an improvement in the court’s ability to assess damages tends to favor the cash institution also in the case of investment agreements. To see this, notice that the sets $\tilde{\Omega}^{FT}_{1,C}$, $\tilde{\Omega}^{FT}_{2,C}$, and $\tilde{\Omega}^{T}_{1,C}$ become empty as $m(s) \to 0$ so that equation (OA 1.8) becomes:

$$E_s [\tilde{\omega}^{C}_{G^* \& I^*} (\tilde{I}^R, s)] - E_s [\tilde{\omega}^{R}_{G^* \& I^*} (\tilde{I}^R, s)]$$

$$= (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^{T}_{2,C}} p(s) qk (e (\tilde{I}^R, s) - CS (\tilde{I}^R, s) - PS (\tilde{I}^R, s))$$

$$+ (1 - \bar{p}) \sum_{s \in \tilde{\Omega}^{T}_{1,C}} p(s) \{[1 - qk (s)] [e (\tilde{I}^R, s) - CS (\tilde{I}^R, s) - PS (\tilde{I}^R)] - \bar{c} (\tilde{I}^R, s) - \bar{c}^* (\tilde{I}^R, s)\}$$

$$+ \bar{p} \sum_{s \in \tilde{\Omega}^{T}_{2,R}, \tilde{\Omega}^{T}_{3,R}} p(s) qk (s) [e (\tilde{I}^R, s) - CS (\tilde{I}^R, s) - PS (\tilde{I}^R, s)] .$$

The term on the second line is unambiguously positive (capturing the social benefits of efficient breach under cash damages), while the term on the third line can be positive or negative in principle but becomes positive if the joint litigation costs are sufficiently small (capturing the social net benefits of the additional litigation under cash damages). The term on the fourth line is new relative to our earlier discussion of trade agreements and refers to the case in which the Host government can commit to implementing the first-best policies. It is unambiguously positive since the Host government is always able to implement the first best policy under cash damages by paying damages even if the court makes a mistake. Overall, an investment treaty with cash damages therefore dominates an investment treaty with retaliation if the court’s ability to assess damages is sufficiently good, assuming that the joint litigation costs are sufficiently small.

Also just as in the case of a trade agreement, an increase in the probability that free trade is the efficient policy tends to favor an investment treaty with retaliation. To see this, notice that the sets $\sigma^{T}_{2,C}$, $\Omega^{T}_{1,C}$, and $\tilde{\Omega}^{T}_{2,C}$ become empty as $\sum_{s \in \sigma^{FT}} p(s) \to 1$ so that expression (OA
1.8) reduces to:

\[
E_s [\tilde{\omega}_{C, k, I}^*(\tilde{I}^R, s)] - E_s [\tilde{\omega}_{C, k, I}^* (\tilde{I}^R, s)] \\
= -(1 - \bar{p}) \sum_{s \in \mathcal{S}_T^F} p(s) [1 - qk(s)] m(s) (1 - \kappa) PS (\tilde{I}^R, s) \\
- (1 - \bar{p}) \sum_{s \in \mathcal{S}_T^F} p(s) [qk(s) (1 - \kappa) PS (\tilde{I}^R, s) + \bar{c} (\tilde{I}^R, s) + \bar{c}^* (\tilde{I}^R, s)] \\
- (1 - \bar{p}) \sum_{s \in \mathcal{S}_T^F} p(s) \{[1 - qk(s)] (1 - \kappa) PS (\tilde{I}^R, s) - \bar{c} (\tilde{I}^R, s) - \bar{c}^* (\tilde{I}^R, s)\}.
\]

The terms on the second and third lines are unambiguously negative (capturing the costs of inefficient breach and more aggressive protectionism under cash damages, respectively). The term on the last line can be positive or negative in principle but is negative if the joint litigation costs are sufficiently small (capturing the social net benefits of the additional litigation under cash damages in states \(s \in \tilde{\Omega}_2^{FT}\)). Overall, an investment treaty with retaliation therefore dominates an investment treaty with cash damages if free trade is sufficiently likely to be the efficient investment policy, assuming that the joint litigation costs are sufficiently small.

Imposing that the joint litigation costs are sufficiently small, we can therefore state:

**Proposition 4 (ii).** Allowing for cash damages instead of retaliation in an investment treaty is optimal if (1) the DSB’s ability to assess cash damages is sufficiently good, and (2) there is a non-trivial probability that a taking is the efficient policy.

2. The Remedial Period

Here we prove Proposition 5. As noted in the body of the paper, we let \(\delta \in [0, 1]\) parameterize the fraction of the harm from the policy action at issue that occurs retrospectively, that is, prior to the court ruling. If \(\delta = 0\), there is no pre-ruling harm, as in sections 3 and 4; at the other extreme, if \(\delta = 1\), the harm has all occurred and is a bygone by the time of the ruling. We build on our analysis in sections 3 and 4 and assume that the trade agreement has adopted SSDS while the investment treaty also includes ISDS. And building on our analysis of the nature of the remedy as summarized in Proposition 4, for the retrospective damages we assume that the trade agreement relies on retaliation for damage payments while the investment treaty employs cash, and we now assume that the court can perfectly assess the level of damages so that we
can focus on the inefficiency of retaliation relative to cash as a form of damage payments. And finally, to keep the comparison clean we continue to assume that the prospective remedy for both trade agreements and investment treaties is a cease and desist order, just as in the models of sections 3 and 4: this means that in the case where $\delta = 0$ and there is no pre-ruling harm, the augmented models that we develop below collapse to the original models of sections 3 and 4, a feature that makes our comparisons easier but is not necessary for our results.

2.1. Trade Agreements

We consider first the case of a trade agreement, and look for conditions under which prospective remedies would be optimal. To this end, we consider the implications of adopting retrospective remedies in a trade agreement. Under retrospective remedies, if the DSB rules for $FT$, the Home government must both cease and desist its $P$ policy and revert to $FT$ henceforth (prospective damages), and it must make damage payments to the Foreign government in the amount of the harm $\delta|\gamma^*_G(s)|$ already suffered (retrospective damages).  

A key question is the form that such retrospective damage payments take. As indicated in the text and consistent with GATT/WTO practice and our results from Proposition 4, we assume that in the context of a trade dispute these damage payments take the form of additional tariff adjustments in other sectors, made either by the importer government (who would reduce these other tariffs) or the exporter government (who would raise tariffs and hence engage in reciprocal retaliation), that amount to a costly transfer to the exporter government. We capture the cost of such ex-post transfers in this setting with the parameter $\beta \in (0, 1]$ representing the fraction of each dollar given up by the Home government that reaches the Foreign government. Hence, a damage payment of $\delta \times |\gamma^*_G(s)|$ received by the Foreign government costs the Home government $\frac{1}{\beta}[\delta \times |\gamma^*_G(s)|]$ in lost surplus. In the context of trade agreements we will highlight outcomes that arise in the absence of cash transfers where $\beta$ is small.

---

4As we note, our assumption that for the retrospective damages the trade agreement relies on retaliation while the BIT relies on cash can be rationalized by our findings in Proposition 4, but there are also other arguments that can provide support for this assumption (see, for example, Sykes, 2005, Limao and Saggi, 2008, and Bagwell and Staiger, 2010, note 10).

5We do not include the litigation costs $c^*(s)$ borne by the Foreign government in this damage payment, but this could be considered as well.

6Our implicit assumption is that trade agreements have resulted in the importer government’s tariffs in other sectors being set at efficient levels, so that if the importer government were to make an adjustment (downward) in these tariffs to pay damages to the exporter government, these adjustments would have negative efficiency consequences, just as would be the case if the exporter government were to collect damage payments by raising its tariffs against the importer government.
Adopting retrospective remedies will have important implications for the conditions describing the equilibrium behavior of the two governments. Consider first the Foreign government’s filing behavior under a retrospective remedy. The Foreign government files a complaint if and only if $\tau = P$ and its expected benefit of filing exceeds its cost of filing. Using the shorthand $\mu_R^* (s) \equiv \frac{c^*(s)}{\gamma_{G*}(s)}$, this can be written as

$$\Pr(\text{DSB ruling is } FT \mid s) > \mu_R^* (s). \quad (\text{OA 2.1})$$

Condition (OA 2.1) is the “filing” condition for the Foreign government to invoke the DSB in response to a policy choice by the importer government of $\tau = P$. Notice that it is the same as condition (3.1) in the main text, except that we have already imposed SSDS. This is because the Foreign government still receives $|\gamma_{G*}(s)|$ overall if its complaint is successful, $\delta \times |\gamma_{G*}(s)|$ for the pre-ruling harm it suffered and $(1 - \delta) \times |\gamma_{G*}(s)|$ from the subsequent switch to $FT$.

Next consider the Home government’s policy choice under a retrospective remedy. The Home government chooses $\tau = P$ if either (OA 2.1) fails – because then it can set $\tau = P$ without triggering a dispute – or if (OA 2.1) holds and the expected benefit to the Home government from trade protection exceeds the cost to the Home government of a dispute, which is the case if

$$\Pr(\text{DSB ruling is } P \mid s) \times \gamma_G(s) + \Pr(\text{DSB ruling is } FT \mid s) \times \left\{ \delta \times |\gamma_{G*}(s)| - \delta \times \frac{|\gamma_{G*}(s)|}{\beta} \right\} > c(s).$$

Defining $\mu_R (s) \equiv \frac{c(s) - \delta \times |\gamma_{G*}(s)|}{(1 - \delta) \gamma_G(s) + \delta \times |\gamma_{G*}(s)|}$, this can be expressed more compactly as:

$$\Pr(\text{DSB ruling is } P \mid s) > \mu_R (s). \quad (\text{OA 2.2})$$

The important novel element of (OA 2.2) is that if the Home government chooses $P$ and the DSB rules in favor of $FT$, the Home government will be responsible for compensating the Foreign government for the harm done prior to the ruling (with the retroactive damage payment $\delta \times \frac{|\gamma_{G*}(s)|}{\beta}$), and there is no action that the Home government can take to avoid making these damage payments once they are assessed (i.e., while the harm to the Foreign government going forward can be removed by reverting to a policy of $FT$, this does nothing to address the retrospective harm).

We can now derive the equilibrium actions for each state $s$ in the presence of retrospective remedies. Conditions (OA 2.1) and (OA 2.2) are identical to conditions (3.1) and (3.2) in the main text, up to the definition of $\mu_R (s)$ and $\mu_R^* (s)$. Therefore, Lemma 1 again carries over to this extended setting, assuming that dispute costs are low relative to dispute stakes in the sense that $\mu_R (s) + \mu_R^* (s) < 1$ for all $s$: 
Lemma OA 3. Equilibrium actions under retrospective damages are as follows:

1. In states $s \in \sigma^{FT}$: If DSB quality is high in the sense that $q_k(s) \leq \mu_R(s)$, we have $\tau = FT$ and no dispute; if DSB quality is intermediate in the sense that $q_k(s) \in (\mu_R(s), 1 - \mu_R^*(s))$, we have $\tau = P$ and a dispute; if DSB quality is low in the sense that $q_k(s) \geq 1 - \mu_R(s)$, we have $\tau = P$ and no dispute.

2. In states $s \in \sigma^P$: If DSB quality is high in the sense that $q_k(s) \leq \mu_R^*(s)$, we have $\tau = P$ and no dispute; if DSB quality is intermediate in the sense that $q_k(s) \in (\mu_R^*(s), 1 - \mu_R(s))$, we have $\tau = P$ and a dispute; if DSB quality is low in the sense that $q_k(s) \geq 1 - \mu_R(s)$, we have $\tau = FT$ and no dispute.

Lemma OA 3 reveals a key point: for any $\delta > 0$ and as $\beta$ approaches zero, $\mu_R(s)$ approaches one and $1 - \mu_R(s)$ approaches zero, and the importer government will always choose $FT$ to avoid any possibility of having to make costly transfer payments to the foreign exporter government for retroactive damages (because then $q_k(s) < 1 = \mu_R(s)$ for all $s \in \sigma^{FT}$ and $q_k(s) > 0 = 1 - \mu_R(s)$ for all $s \in \sigma^P$).

Put differently, for any fixed court quality $q$, the joint surplus under a trade agreement with retrospective remedies will approach the joint surplus associated with $FT$ in all states as $\beta$ approaches zero and the cost of transfers becomes prohibitive. We record this in:

**Remark OA 1.** If transfers in the context of a trade dispute are sufficiently costly ($\beta$ small), then for any $\delta > 0$ the joint surplus under a trade agreement with retrospective remedies will approach the joint surplus associated with $FT$ in all states, no matter how accurate the court may be (for any $q > 0$).

Remark OA 1 implies that for any $\delta > 0$ and $q > 0$, if $\beta$ is sufficiently small then a trade agreement with retrospective remedies will be dominated by a trade agreement with prospective remedies provided that the trade agreement with prospective remedies delivers a level of joint surplus higher than that associated with $FT$ in all states. But if the quality of the court $q$ is fixed at a sufficiently high level, then for any level of $\beta$, as $\delta$ approaches zero so that litigation delay becomes sufficiently short, the joint surplus under a trade agreement with prospective remedies

---

7 Notice too that as we allow $\beta$ to approach zero we must also have $c^*(s)$ approaching zero in order to maintain our “relatively small litigation cost” focus and ensure that $\mu_R(s) < 1 - \mu_R^*(s)$.
can be brought arbitrarily close to the first best level, which exceeds the joint surplus associated with $FT$ in all states and therefore beats a trade agreement with retrospective remedies under these conditions. We may therefore state:

**Proposition 5 (i).** Provided that the quality of the DSB is above a threshold level, a prospective remedy is optimal for a trade agreement if the degree of litigation delay is sufficiently short and transfers in the context of a trade dispute are sufficiently costly.

### 2.2. Investment Agreements

We now turn to the case of an investment treaty, and look for conditions under which retrospective remedies would be optimal. To this end, we consider the implications of adopting prospective remedies in an investment treaty, where the foreign government and the foreign investor have standing. Relative to our earlier analysis of investment treaties (i.e., relative to the case where $\delta = 0$), for $\delta > 0$ the conditions describing equilibrium behavior of the Host government and the Foreign complainant (either the Foreign government or the Foreign investor) are altered.

Consider first Foreign’s filing behavior under a prospective remedy. Each agent $a \in \{G^*, I^*\}$ then has an incentive to file if and only if $\Pr(\text{DSB ruling is } FT \mid s) > \bar{c}^*(I^*, s)$. Defining $\bar{\mu}_R^*(s) \equiv \min \left\{ \frac{\bar{c}^*(I^*, s)}{(1-\delta)\tilde{c}^*(I^*, s)}, \frac{\bar{c}^*(I^*, s)}{(1-\delta)\tilde{c}^*(I^*, s)} \right\}$, we can therefore write that Foreign files a complaint if and only if $\iota = T$ and

$$\Pr(\text{DSB ruling is } FT \mid s) > \bar{\mu}_R^*(s). \quad \text{(OA 2.3)}$$

Condition (OA 2.3) is Foreign’s “filing” condition to invoke the DSB in response to a policy choice by the Host government of $\iota = T$.

Next consider the Host government’s policy choice. When it has discretion to do so, this government chooses $\iota = T$ if either (OA 2.3) fails – because then the Host government can set $\iota = T$ without triggering a dispute – or if (OA 2.3) holds and the expected benefit to the Host government from a taking exceeds the cost to the Host government of a dispute, which is the case if $\Pr(\text{DSB ruling is } FT \mid s) \times \delta \bar{\gamma}_G(I^*, s) + \Pr(\text{DSB ruling is } T \mid s) \times \bar{\gamma}_G(I^*, s) > \bar{c}(I^*, s)$. Defining $\bar{\mu}_R(s) \equiv \frac{\bar{c}(I^*, s) - \delta \bar{\gamma}_G(I^*, s)}{(1-\delta)\bar{c}(I^*, s)}$, we can write this as:

$$\Pr(\text{DSB ruling is } T \mid s) > \bar{\mu}_R(s). \quad \text{(OA 2.4)}$$

The important novel element of (OA 2.4) is that if the Host government chooses $T$ and the
DSB rules in favor of FT, the Host government still enjoys the benefits of T for the pre-ruling period \((\delta \tilde{\tau}_G(I^*, s))\).

Conditions (OA 2.3) and (OA 2.4) are the same as conditions (4.9) and (4.10) in the main text, up to the definition of \(\tilde{\mu}_R(s)\) and \(\tilde{\mu}_R^*(s)\). Assuming again that dispute costs are low relative to dispute stakes in the sense that \(\tilde{\mu}_R(s) + \tilde{\mu}_R^*(s) < 1\) for all \(s\), Lemma 3 from the main text carries over to this setting:

**Lemma OA 4.** Equilibrium actions under prospective damages are as follows:

1. In states \(s \in \sigma^{FT} \):
   1. If the Host government is constrained: We have \(\iota = FT\) and no dispute.
   2. If the Host government is unconstrained: If DSB quality is high in the sense that \(q_k(s) \leq \tilde{\mu}_R(s)\), we have \(\iota = FT\) and no dispute; if DSB quality is intermediate in the sense that \(q_k(s) \in (\tilde{\mu}_R(s), 1 - \tilde{\mu}_R^*(s))\), we have \(\iota = T\) and a dispute; if DSB quality is low in the sense that \(q_k(s) \geq 1 - \tilde{\mu}_R^*(s)\), we have \(\iota = T\) and no dispute.

2. In states \(s \in \sigma^T \):
   1. If the Host government is constrained: We have \(\iota = T\), no dispute if \(q_k(s) \leq \tilde{\mu}_R^*(s)\), and a dispute if \(q_k(s) > \tilde{\mu}_R^*(s)\).
   2. If the Host government is unconstrained: If DSB quality is high in the sense that \(q_k(s) \leq \tilde{\mu}_R^*(s)\), we have \(\iota = T\) and no dispute; if DSB quality is intermediate in the sense that \(q_k(s) \in (\tilde{\mu}_R^*(s), 1 - \tilde{\mu}_R(s))\), we have \(\iota = T\) and a dispute; if DSB quality is low in the sense that \(q_k(s) \geq 1 - \tilde{\mu}_R(s)\), we have \(\iota = FT\) and no dispute.

Note that if \(\delta = 0\) and there is hence no litigation delay, the above characterization of equilibrium behavior collapses to our earlier analysis of investment treaties under SSDS and ISDS. On the other hand, if \(\delta\) is sufficiently close to one, (OA 2.3) and (OA 2.4) together with our focus on the relatively-low-dispute-cost case imply that the Host government will always choose T when it has the discretion to do so and the foreign investor will never invoke the DSB, and hence for \(\delta\) in this range and conditional on any level of investment, the investment treaty with prospective remedies would be valueless, as it would deliver the noncooperative outcome in which the Host government always expropriates in \(\sigma^T\) and expropriates in \(\sigma^{FT}\) with probability.
1 − \bar{p}. Formally, the critical level of \delta beyond which an investment treaty with prospective remedies would be valueless, which we denote by \delta, is defined by

\bar{\delta} = 1 - \min_{a,s} \left[ \frac{c^*(s)}{\overline{\alpha}_a(s)} \right] > 0,

where the inequality follows from our focus on the relatively-low-dispute-cost case. For \delta \in [\delta, 1], it follows that in all states s \in \sigma^{FT} with probability 1 − \bar{p} the Host government chooses T and the foreign investor will not file (because qk(s) > 0 ≥ 1 − \bar{\mu}^*_R(s) for all s), and in all states s \in \sigma^{T} the Host government chooses T and the foreign investor does not file (because qk(s) < 1 ≤ \bar{\mu}^*_R(s) for all s). We record this in:

**Remark OA 2.** If litigation delay is sufficiently high (for \delta ≥ \bar{\delta}), the Host government cannot improve upon a stand-alone program of up-front investment incentives to foreign investors by introducing an investment treaty with prospective remedies, no matter how accurate the court may be (for any q > 0).

Remark OA 2 implies that for any q > 0, if \delta is sufficiently high then an investment treaty with prospective remedies will be dominated by an investment treaty with retrospective remedies provided that the investment treaty with retrospective remedies can improve upon a stand-alone program of up-front investment incentives to foreign investors. But if the quality of the court q is fixed at a sufficiently high level, then for any level of \delta as \beta approaches one so that transfers in the context of an investment treaty are sufficiently efficient, the Host government surplus under an investment treaty with retrospective remedies can be brought arbitrarily close to the first best level, which exceeds the Host government surplus under a stand-alone program of up-front investment incentives and therefore beats an investment treaty with prospective remedies under these conditions. We may therefore state:

**Proposition 5 (ii).** Provided that the quality of the DSB is above a threshold level, a retrospective remedy is optimal for an investment treaty if the degree of litigation delay is sufficiently long and transfers in the context of an investment treaty are sufficiently efficient.