Disputes in International Investment and Trade*

Ralph Ossa† Robert W. Staiger‡ Alan O. Sykes§

December 2020

Abstract

International investment agreements employ dispute settlement procedures that differ markedly from their counterparts in trade agreements. We develop parallel models of trade and investment agreements and employ them to study these differences. We find that many of the observed differences can be understood as deriving from the fundamentally different problems that trade and investment agreements are designed to solve. The typical dispute settlement process in trade agreements - one that emphasizes state-to-state dispute settlement, tariff retaliation, and prospective damages - can be justified based on the government-to-government nature of trade agreements. The typical dispute settlement process in investment agreements - one that emphasizes investor-state-dispute settlement, cash damages, and retrospective damages - can be justified based on the government-to-investor nature of investment agreements. Our analysis also identifies some important qualifications to the wisdom of these design features, particularly with respect to the optimal design of investment agreements, thereby offering a potential explanation for the strong political controversy they cause.

---

*We thank seminar participants at Boston College, LSE and the University of Zurich, and conference participants at the 2019 NYU-Columbia Trade Conference, the 2020 Villars-sur-Ollon Workshop on Challenges to the Global Trading System, the 2020 NBER conference on International Trade Policy and Institutions, and the 2020 CEPR ERWIT conference for helpful comments. Ralph Ossa gratefully acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 819394). The usual disclaimer applies.

†Department of Economics, University of Zurich; and CEPR.
‡Department of Economics, Dartmouth College; and NBER.
§Stanford Law School.
1. Introduction

Firms that export goods and services to foreign customers are typically protected by trade agreements such as those of the World Trade Organization (WTO), which secure their access to foreign markets. Firms that engage in foreign direct investment are often protected by investment agreements such as those negotiated pursuant to the U.S. bilateral investment treaty (BIT) program, which protect their foreign assets from expropriation and related practices.¹

The dispute settlement procedures in these agreements have become a subject of much political controversy. The predominant remedy for breach of investment treaty commitments - investor-state dispute settlement (ISDS) - has come under attack from both the left and the right. For example, Senator Elizabeth Warren has written: “Conservatives who believe in U.S. sovereignty should be outraged that ISDS would shift power from American courts, whose authority is derived from our Constitution, to unaccountable international tribunals. Libertarians should be offended that ISDS effectively would offer a free taxpayer subsidy to countries with weak legal systems. And progressives should oppose ISDS because it would allow big multinationals to weaken labor and environmental rules.”² An October, 2017 letter to President Trump was signed by 230 law and economics professors, who urged the President to eliminate ISDS from the U.S.-Mexico-Canada Agreement (USMCA), insisting that ISDS “undermines the important roles of our domestic and democratic institutions, threatens domestic sovereignty, and weakens the rule of law.”³ The Trump administration responded by negotiating the phase-out of ISDS for disputes between Canada and the United States, and preserved the prior ISDS mechanism for claims involving Mexico only in a few sectors such as telecommunications and oil and gas.⁴ Interestingly, much of the resistance to eliminating the ISDS mechanism from the USMCA came from the Mexican government itself, which secured its partial retention over the initial objections of the United States.⁵

¹In addition to the agreement establishing the WTO, there are nearly 300 “regional trade agreements” in force. International investment agreements are even more numerous, with over 2000 bilateral investment treaties in force, and roughly 300 additional investment agreements that are part of larger economic arrangements (usually free trade agreements).
⁴Other investors can pursue claims against Mexico on a more limited legal basis and only after litigating in Mexican court for up to 30 months.
⁵On Mexico’s position in these negotiations, see https://www.lexology.com/library/detail.aspx?g=c42ad0d6-3240-4e24-ac21-5cac954962c. The opposition to ISDS is not limited to the United States. In Europe, the Belgian province of Wallonia temporarily blocked the entire Comprehensive Economic and Trade Agreement
The remedy for breach of trade treaty commitments - state-to-state dispute settlement (SSDS) - has also come under attack. On the academic front, critics of the WTO lament the fact that nations can violate treaty obligations yet suffer no formal sanction until the violation is detected, a case is brought and litigated to conclusion, and the violator has had a “reasonable time” to desist from its illegal behavior. This system is said to give violators a “three-year free pass.” More recently, US Trade Representative Robert Lighthizer stated at the 2017 ministerial meeting of the WTO: “The WTO is losing its essential focus on negotiation and becoming a litigation-centered organization. Too often members seem to believe they can gain concessions through lawsuits that they could never get at the negotiating table.” And in July, 2018, Lighthizer referred to a recent WTO dispute panel ruling on U.S. countervailing duty laws as “the latest example of judicial activism at the WTO seeking to undermine those laws.” Dissatisfaction with WTO dispute rulings has led the United States to block the appointment of judges to fill vacancies on the WTO Appellate Body for the past several years, and there are now too few judges to staff a required three-person appellate panel.

All of these controversies arise against a backdrop of intriguing puzzles that the formal economics literature has done little to address. Why do international investment agreements overwhelmingly provide for private rights of action in the form of ISDS rather than simply SSDS? Why do trade agreements uniformly provide for SSDS and not private rights of action in the form of exporter-state dispute settlement (ESDS)? Might private rights of action in the form of ESDS make sense in the international trade arena in light of recent attacks on SSDS? Are the critics of ISDS right that it unwisely burdens national sovereignty and undermines sound regulatory policies? And why did Mexico push to retain the ISDS in the USMCA?

To answer these and related questions, we develop parallel models of trade and investment agreements and employ them to study the relative merits in each setting of the features of dispute settlement that are at the heart of these controversies. We model trade agreements as

---

6See Davey (2010).
9To be sure, some agreements have both trade and investment provisions. In the USMCA, for example, the investor-rights provisions allow for both SSDS and private rights of action in the form of ISDS. But the “trade” provisions of the USMCA, involving commitments on trade in goods and services and intellectual property rights, rely exclusively on SSDS, as does the WTO.

---
focused on the protection of market access commitments negotiated between governments, and thus as government-to-government agreements. We model investment agreements as focused on solving host government commitment problems with respect to foreign investors, and thus as government-to-investor agreements. We show that many of the observed differences in the dispute settlement procedures across trade and investment agreements can be understood as arising from the fundamentally different problems that the two types of agreements are designed to solve. This is a central message of our paper.

Our baseline model of trade agreements mirrors closely that of Maggi and Staiger (2011), where the governments of an importing and an exporting country can make use of a vaguely worded contract written ex ante in the presence of uncertainty about the future state of the world, and where the dispute settlement procedure of the trade agreement involves a court whose mandate is to interpret the contract when invoked ex post in a dispute over whether the trade policy commitments described in the contract have been honored. Our model of investment agreements extends this setting to include an ex-ante investment stage and investment policies of a host-country government that lacks full commitment power, with foreign investors who at the time of their investment decisions form expectations about the ex-post treatment they will receive from the host-country government under the dispute settlement procedures in place.

Our analysis focuses on three distinct features of dispute settlement that form the basis of the controversies described above and that are often conflated: standing, the nature of the remedy, and the remedial period. Standing concerns who has the right to bring cases to a dispute panel. Generally speaking, only member governments have the right to file cases under trade agreements (SSDS). In the vast majority of investment agreements, by contrast, private investors also have standing (ISDS). The nature of the remedy concerns the remedial consequences of an adjudicated violation. In trade agreements, the dispute settlement process may authorize trade sanctions against incompliant violators (as in the WTO) but money awards are not used as a practical matter. Under investment agreements, by contrast, money awards are routinely available to private investors who win in court when the agreements grant them standing. The remedial period refers to the period of time that is “covered” by the remedial measure. In trade agreements, no remedy typically exists for harm done to the complainant prior to the adjudication of a violation, so the focus is on “prospective damages.” Under investment agreements, by contrast, private investors routinely receive “retrospective damages” in the sense that they are compensated for the entirety of the harm suffered as a consequence of the
violation, compensating them for past losses as well as the value of future harms.

The first stage of our analysis concerns standing. We find no support for the inclusion of private standing in trade agreements, while we find qualified - though by no means conclusive - support for the inclusion of private standing in investment agreements. This finding can be understood in two steps. The first step is to observe that moving from SSDS to ESDS in government-to-government trade agreements, and moving from ISDS to SSDS in government-to-investor investment agreements, amounts to transferring the ability to file cases from the “principal” to an “agent,” when private litigants - exporters and investors - are imperfect agents of the government and vice versa. This observation suggests that trade agreements should adopt SSDS, while investment agreements should adopt ISDS, so that the filing authority always remains with the principal. The second step is to observe that foreign complainants do not internalize the costs that filing a dispute imposes on the home/host government, and this negative externality leads to a general tendency to overfile relative to efficient litigation levels. We further argue that private litigants will typically file cases more aggressively than would governments on their behalf. Hence, beginning from the position that trade agreements should adopt SSDS, it would never make sense to introduce the possibility of an even more aggressive filer in the form of ESDS; whereas beginning from the position that investment agreements should adopt ISDS, it might make sense to restrict standing to a less aggressive filer in the form of SSDS.

The second stage of our analysis concerns the nature of the remedy - the choice between retaliatory sanctions and money damages. Building on our analysis of standing, we take as given the fact that trade agreements limit standing to governments while investment agreements afford standing to investors. The key trade-off that we exploit in this stage of our analysis is that trade retaliation creates the deadweight losses of protectionist trade policies, while cash damages, although a costless transfer, can be difficult to assess with accuracy and errors may distort the policy choices of the home/host government. The optimal remedy in a given setting then depends on which of these problems is more acute. Our findings again follow from the observation that trade agreements are government-to-government agreements.

---

10This assumes that free-rider issues facing exporters and private investors are not too severe. But as we explain below, even severe free-rider issues do not overturn our results on standing, because the relevant thought experiment for standing is whether exporters or private investors should be given standing in addition to governments – the agreements on which we focus invariably afford standing to governments even where private actors also have standing.
while investment agreements are government-to-investor agreements. We show that investment agreements should provide for cash payments while trade agreements should not, provided that the cash value of the profits lost by a foreign investor facing adverse host-country investment policies is sufficiently easy to quantify relative to the cash value of the harm suffered by a foreign government whose exporters face trade protection. And we argue that at a broad level this condition is plausibly met given the difficulty of placing a monetary value on the lost jobs and distributional considerations - in addition to the lost revenue of exporters - that typically dominate the trade concerns of real-world governments.

The third stage of our analysis adds consideration of the remedial period in the dispute process - the choice between prospective damages only on the one hand, and the addition of retrospective damages on the other. We find that retrospective remedies are preferable if the retrospective portion of the harm suffered as a consequence of treaty violations is high enough and any inefficiencies associated with the remedy are small enough. We argue that this is likely to be the case for investment disputes, where the pre-ruling harm is usually attributable to policy actions that diminish or destroy the returns to a sunk investment and where, owing to the relative ease of assessing damages in a government-to-investor agreement, the remedy takes the form of monetary transfers. In contrast, we find that prospective remedies alone are preferable when the retrospective portion of the harm suffered as a consequence of treaty violations is low enough and/or the efficiency costs of the remedy are high enough. We argue that this is likely to be the case for trade disputes, where the pre-ruling harm is attributable mostly to the costs of the delay in securing market access and where, owing to the difficulty in assessing damages in a government-to-government agreement, the remedy takes the form of retaliatory trade sanctions. Our analysis thus provides support for the retrospective damages found in investment treaties. But perhaps surprisingly, our analysis also provides some support for the prospective damages approach taken by the WTO and trade agreements more broadly. A “three-year free pass” for violators may be the better option if the alternative involves retrospective damages paid out in the form of highly inefficient retaliatory trade sanctions.

Finally, in deriving our results we allow the host-country government to complement its investment agreement with a domestic program of up-front incentives for foreign investors, and this allows us to ask the following question: \(^{11}\) With such a program in place, is the perceived

\(^{11}\) We have in mind here the kinds of up-front investment incentives that are routinely offered by local, state and federal governments to attract investment, such as tax holidays, government grants and subsidized infrastructure (see, for example, UNCTAD, 1998).
loss of host-country sovereignty that comes with the addition of an (optimally designed) international investment agreement and associated dispute settlement procedure worth it in terms of the enhanced efficiency of investments that the agreement delivers? We find that it can be, but only if the quality of the international court is sufficiently high and the quality of domestic institutions is sufficiently weak; otherwise, countries are better off relying on domestic programs of up-front investment incentives alone to address their commitment problems with foreign investors. This suggests that countries with strong domestic institutions, such as might be expected in many liberal democracies, may well see limited benefits from their own commitments to foreign investors in investment agreements.

Our paper is related to several literatures. A first literature focuses on dispute settlement procedures in trade agreements (see Park, 2016, for a recent survey); within this literature, our paper is most closely related methodologically to Maggi and Staiger (2011), who provide the basic model of dispute settlement in trade agreements on which we build. A second literature focuses on dispute settlement procedures in investment treaties (see, for example Aisbett, Karp and McAusland, 2010a,b, Konrad, 2017, Horn, 2018, Stahler, 2018, Horn and Tangeras, 2019, Janeba, 2019, Kohler and Stahler, 2019 and Sykes, 2019). Our use of the Maggi and Staiger (2011) framework makes us methodologically distinct from these papers (though Janeba, 2019, adopts a model of dispute settlement that shares common elements with ours). Relative to all of this literature, our paper ventures into largely uncharted territory: to the best of our knowledge, we are the first to formally explore the choice of standing, the nature of the remedy, and the remedial period together in either trade or investment agreements and, by considering these features in both trade and investment agreements, the first to offer a formal comparative analysis of the core design features of dispute settlement procedures across these two settings.12

The remainder of the paper proceeds as follows. Section 2 provides additional institutional detail on trade and investment agreements and their dispute settlement systems. Section 3 contains our analysis of standing in trade agreements while section 4 turns to the standing issue in investment agreements. Section 5 summarizes our analysis of the nature of the remedy and the remedial period in trade and investment agreements, with technical details of this analysis included in an online Appendix. Section 6 concludes.

12Horn (2018) considers the implications of switching from ISDS to SSDS in the context of BITs in a model similar to Horn and Tangeras (2019), but he does not optimize other features of the agreement and does not consider trade agreements.
2. Characteristics of Trade and Investment Agreements

Important heterogeneities exist within and among trade and investment agreements and their approaches to dispute settlement, and it is not our intent here to provide a comprehensive survey of these heterogeneities. Instead, we focus on “typical” characteristics of trade and investment agreements with reference to a few illustrative examples.

2.1. Trade Agreements

The predominant economic account of international trade cooperation is the “terms of trade theory” (Bagwell and Staiger 2002), wherein the inefficiency of unilateral trade policy arises because of international price externalities attributable to trade impediments erected by “large” countries. Governments negotiate for reciprocal market access commitments to abate these externalities. The benefits to each government flow not from the gains afforded to foreign exporters, but from the gains to their own exporters. Likewise, the utility to country A derived from a legal commitment to country B lies in maintaining the incentive for country B to honor its own commitments that benefit country A’s exporters. Accordingly, government parties to trade agreements have no interest in permitting enforcement litigation unless foreign governments wish it to go forward – if country B’s government is unconcerned by some measure that violates the commitments of country A, an enforcement action against country A by an exporter from B will impose costs on A without incentivizing any reciprocal beneficial behavior by B. This is the sense in which trade agreements are government-to-government agreements (Sykes 2005).

The key substantive obligations under modern trade agreements are tariff commitments, restrictions on other protective border measures (such as quotas) that can impede market access, and constraints on various domestic policies (e.g., discriminatory taxes) that can undermine market access commitments associated with tariff reduction. A well-known example of such an agreement is the General Agreement on Tariffs and Trade (GATT), first negotiated in 1947.

Under GATT, formal sanctions for violations did not arise in practice, but disputants could agree to allow arbitral panels to adjudicate the merits of a case and issue a report as to whether a violation had occurred. Over time, the membership became dissatisfied with this system and, with the creation of the WTO (which subsumed GATT), established a new system whereby complaining nations can obtain an arbitral panel and secure a final ruling including a right of appeal, accompanied by a formal recommendation to a violator to cease the violation within a
“reasonable time.” If a violator fails to do so and does not negotiate alternative compensation, the complainant may impose retaliatory trade measures (a “suspension of concessions”) in an amount “equivalent” to the level of harm done by the violation, an issue subject to arbitration if requested by the violator. In practice, arbitrators judge “equivalence” in relation to the prospective harm from the ongoing violation (Davey 2010). Scholars have debated whether this system should be viewed as a “property rule” or “liability rule” (Schwartz & Sykes 2002), and we take no position on the issue. But three observations are beyond dispute - only member state governments have standing to file cases (SSDS, but not ESDS), the system relies on negotiated compensation and retaliatory trade measures rather than monetary damage awards as its sole formal “remedy,” and at no time in the history of WTO/GATT has there been any formal remedy for harm done prior to the adjudication of a violation.

The dispute settlement systems under the hundreds of other trade agreements now in force vary somewhat, but most share core features with the WTO system. Under USMCA Chapter 31, for example, only member governments can bring complaints in relation to the trade provisions. A complaining party can seek an arbitral panel, which will rule on the existence of a violation. If one is found, the remedy “if possible” is removal of the offending measure, or negotiated trade compensation. Failing such a solution, the complaining party again has the option to suspend “benefits of equivalent effect” to the violation. There is no monetary remedy, and the concept of “equivalence” is forward looking in relation to the harm done by the ongoing violation until such time as it is eliminated or the case is settled.

2.2. Investment Agreements

Investment agreements primarily serve to protect established investors (often with sizeable sunk costs) against certain measures by host countries that impair the value of their established investments, such as expropriation without adequate compensation, discrimination in favor of host country or third-country investors, and a denial of “fair and equitable treatment” that typically entails governmental deception or denial of due process.

The main inefficiency to be addressed by an investment agreement is a time inconsistency problem, whereby host governments will tend to exploit foreign investors after they incur sunk costs. This problem motivates the host country to seek commitment strategies to assure foreign investors that they will not be exploited. Investment agreements are one such strategy. By assuring the individual investor that its legitimate interests will be respected, the host country
directly benefits through a reduced cost of imported capital. And when capital is elastically supplied, the host country fully internalizes the benefits of its commitments to investors. Accordingly, host countries have an interest in an enforcement regime that effectively protects the legitimate interests of private investors, regardless of whether the investors’ home governments would be inclined to bring enforcement actions. This is the sense in which investment agreements are government-to-investor agreements.

SSDS in some form can be found in virtually all international investment agreements. But in addition, ISDS is included in 95% of investment treaties currently in force according to UNCTAD (2018). Investors who believe that a treaty commitment has been violated, and who have not secured adequate redress through consultations or litigation in the host country, can bring a case to international arbitration before neutral arbitrators (such as before the International Center for the Settlement of Investment Disputes (ICSID) at the World Bank). When a violation is found, the arbitrators proceed to assess damages (the “quantum”), and issue an award directing the host country to compensate the complaining investor for the violation. In cases of expropriation of property, they may offer the host country an opportunity to reduce damages by making restitution of the property. The general principle governing damages is that the investor should be fully compensated for all past and future injury due to the violation.

2.3. Additional Considerations Motivating our Modeling Assumptions

Above we have highlighted the key institutional differences across the dispute settlement procedures of trade and investment agreements. We now motivate a number of features of the economic environment relevant to trade and investment activities that we will attempt to capture in our modeling framework and that we show can give rise to these institutional differences.

We take the position that under both trade and investment agreements, the private beneficiaries of commitments (exporters and investors) typically gain more from the enforcement of commitments than do their home governments. In the trade setting, for example, the enforcement of foreign market access commitments on behalf of exporters will tend to increase prices at home for exported goods and the inputs used to produce them, harming local consumers and other purchasers of those inputs. Governments take these ancillary effects into account while exporters do not. Similarly, in both the trade and investment settings, governments may have a variety of military, diplomatic, security and related reasons for preferring not to bring cases against certain foreign governments when their exporters or investors would nevertheless benefit
(Levy & Srinivasan, 1982). Accordingly, it is reasonable to assume (with some caveats that we will address) that exporters and investors in general, if given standing to bring cases, will be more aggressive litigants than their governments. This observation immediately suggests that cases may be filed too often from the perspective of exporters’ home governments if exporters have standing to sue, while cases will be filed too infrequently from the perspective of investors if investors are not given standing to sue.

As regards the nature of the remedy for a treaty violation, we will consider two possibilities, trade sanctions or cash damages. As we noted in the Introduction, compared to cash damages, trade sanctions carry potentially substantial deadweight costs that diminish their utility, but cash damages can be difficult to assess. We take the position that assessing cash damages is relatively easy and accurate in the context of an investment dispute, but quite difficult and prone to error in the context of a trade dispute. In cases involving violations of an investment agreement, an individual investor, with sunk investments in the host country, brings a claim for lost profits based on the difference between actual profits and profits that would obtain in a counterfactual setting without the treaty violation. To a first approximation, the damages claim is no more difficult to adjudicate than the damages claim in typical breach of contract cases. In cases involving violations of trade agreements, by contrast, even if the lost export revenue were the only relevant measure of harm, vast numbers of exporters may be affected in dozens of countries, and an analysis of how each exporter’s profitability is affected can become immensely complicated, with great uncertainty as to who would have made what quantity of sales at what price, issues as to the ability of exporters to mitigate their damages by selling at home or in third-country markets, and so on. WTO arbitrators make no effort to sort out such issues in cases that proceed to a retaliatory phase, and content themselves simply to estimate how a violation has affected the aggregate volume of trade with the violator. And it is not even clear that the damages suffered by exporters are the right “measure” of damages for a government-to-government agreement: governments, unlike their exporters, may have legitimate concerns about unemployment and other distributional issues across the wider economy that would add further complications to any damage calculation in the context of a trade agreement.

Regarding the remedial period, we take the position that the retrospective harm is generally much more important in the investment context than in the trade context. Investment disputes concern an impairment to the profitability of a sunk investment with a finite lifespan. The impairment may be irreversible (as when property has been taken and put to another use),
and in any case the treaty violation usually greatly reduces the total returns to the investment. The “retrospective harm” due to the treaty violation is thus the bulk of the damage suffered by the investor. A violation of a trade agreement, by contrast, typically involves a protectionist impediment to exports that can be dismantled - indeed, the “preferred” remedy under WTO law (WTO DSU Art. 22.1) is the elimination of the illegal measure. Although exporters no doubt suffer some harm while the illegal measure is in place (perhaps quite difficult to compute as noted), the removal of the measure eliminates the market access impediment and may facilitate a stream of profits to exporters for decades to come.

Finally, to make our analysis pertinent to current controversies over trade and investment disputes, we focus on settings in which disputes arise in equilibrium. In principle, a dispute system might be so inexpensive, accurate, and effective at deterring violations that it never has to be used, or it might be so costly and riddled with error as to be worthless and then fall into disuse. As the Introduction makes clear, however, both the trade and the investment approaches to dispute settlement are generating no shortage of litigated disputes, while both are often criticized as to their costs, accuracy and efficacy.

To capture this reality, we take the position that disputes are costly to litigate but not prohibitively costly. We also focus on scenarios where the stakes for the parties are high enough to justify the expense of litigating. And we will focus on scenarios where the efficiency of challenged conduct is not obvious to a third-party observer (such as a court), creating uncertainty as to the outcome of litigation. Thus, for example, we imagine in the trade context that the challenged measure involves a policy that could appear to be efficient or inefficient (such as a subsidy that might be welfare-enhancing or welfare-reducing, or a violation of a non-discrimination obligation that might or might not have a sufficient regulatory justification). So too in the investment context - the measure might involve a taking of property that might or might not put it to a higher valued use, or a regulatory measure that greatly impairs an existing investment and may or may not yield net social benefits. Our modeling strategy parameterizes these considerations, and then directs attention to parameter regions that yield equilibrium disputes.

3. Standing in Trade Agreements

We begin our formal analysis by focusing exclusively on issues of standing, and thus the question of whether to include ESDS/ISDS, in addition to SSDS, in an optimally designed trade/investment agreement. In this section we consider the standing issue when the underly-
ing inefficiency to be addressed by the agreement relates to market access/terms of trade issues. Consistent with our discussion in section 2, our formal results are derived in the setting of a trade agreement, reflecting the position that these are the central issues of concern in trade agreements, but we also comment on how our results extend to market access issues that are handled in international investment agreements. In the following section we then address the standing issue when the underlying inefficiency to be addressed by the agreement relates to a government’s imperfect ability to make commitments to foreign firms that must make sunk investments to serve the domestic market. There our formal results are derived in the setting of an investment agreement, again consistent with our discussion in section 2 and reflecting the dominance of such issues for investment treaties; and there we comment on how our results extend to commitment issues that are handled in trade agreements.

To isolate the standing issue, we make two simplifying assumptions in this and the next section. First, we assume that damage payments are not part of the court’s ruling - when a case is filed and the court sides with the complainant, the defendant has no choice but to “cease and desist” whatever policy is found to be illegal. Second, we assume that compliance with any such ruling is instantaneous so that there is no pre-compliance harm suffered by exporters or investors for which additional remedies might be desirable. We relax these assumptions in section 5, where we consider the optimal nature of the remedy (cash versus retaliation) and the optimal remedial period (prospective versus retrospective) in trade and investment agreements.

### 3.1. Model Preliminaries

We build on the model of Maggi and Staiger (2011). We focus on a single industry in which Home is an importer and Foreign is an exporter, and where their trade is governed by a trade agreement with a dispute settlement body (DSB) to resolve disputes. Home has to make a binary policy choice \( \tau \in \{FT, P\} \) (free trade or protection) and Foreign has to decide whether to file a complaint with the DSB (there are no export policy instruments). We distinguish between an agreement that includes only SSDS, in which the filing choice is made exclusively by the Foreign government, and an agreement that includes both SSDS and ESDS, where either the Foreign government or the Foreign export industry may file a complaint.

There are \( s \equiv (s_1, s_2, \ldots, s_N) \) binary states of the world, such as “there is/is not an import surge,” and we let \( p(s) \) denote the probability that state \( s \) occurs. We assume that the Home government gains from protection and that both the Foreign government and its exporting
industry suffer from protection in all states of the world. Denoting the Home government’s payoff from choosing policy $\tau$ in state $s$ as $\omega_G(\tau; s)$, we then have Home’s gain from protection given by $\gamma_G(s) \equiv \omega_G(P; s) - \omega_G(FT; s) > 0$ for all $s$; and similarly, denoting the Foreign government and Foreign export industry payoffs as $\omega_{G\tau^*}(\tau; s)$ and $\omega_{E\tau^*}(\tau; s)$ respectively, we have that Foreign agent $a$’s loss from protection for $a \in \{G^*, F^*\}$ is given by $\gamma_a^*(s) \equiv \omega_a^*(P; s) - \omega_a^*(FT; s) < 0$ for all $s$. These assumptions can be given a terms-of-trade interpretation and would hold in any standard trade model. Moreover, $\omega_G(\tau; s)$ and $\omega_{G\tau^*}(\tau; s)$ can also capture distributional/political economy considerations in government objectives such as would be reflected by over-weighting producer surplus.

Free trade is the “first best” policy in all states $s \in \sigma^{FT}$ and protection is the “first-best” policy in all states $s \in \sigma^P$, where we define the first-best policy as the policy that maximizes the governments’ joint payoff. Letting $\Gamma(s) \equiv \gamma_G(s) + \gamma_{G\tau^*}(s)$ denote the governments’ joint (positive or negative) gain from protection, we then have that $s \in \sigma^{FT}$ if $\Gamma(s) \leq 0$ and $s \in \sigma^P$ if $\Gamma(s) > 0$. If the Home government had access to a full set of policy instruments for intervention, including possibly lump sum taxes for redistribution, trade protection would never be first best and the set $\sigma^P$ would be empty; we treat $\sigma^P$ as non-empty, which amounts to an assumption that the Home government lacks this full set of policy instruments.

In the absence of a trade agreement, the Home government would choose $\tau = P$ in all states of the world. This noncooperative policy choice would correspond to the first best for $s \in \sigma^P$, but it would differ from the first best for $s \in \sigma^{FT}$, giving rise to the possibility that the two governments could do better under a trade agreement. We assume that the realized state $s$ is observed by all agents including the DSB, but that the DSB does not observe $\Gamma$ and hence payoff levels are not verifiable. This means that the first-best outcome cannot be trivially achieved with a contract that requires $FT$ if and only if $\Gamma \leq 0$. We assume as well that it is prohibitively costly to describe precisely all the relevant state variables $(s_1, s_2, \ldots, s_N)$ that would be necessary to write a complete contingent contract covering the policy $\tau$. We focus instead on what Maggi and Staiger (2011) call a “vague contract” that takes the form “$\tau = P$ allowed if and only if $\nu$,” where $\nu$ is a vague sentence such as “there is serious injury to the domestic industry due to increased imports.” This off-the-shelf language makes the vague contract essentially costless to write, and it is assumed to specify the first-best policy choice in those states of the world where its meaning is unambiguous. But the meaning of this contract is ambiguous in some states of the world, and it is in such states that a dispute over the setting
of $\tau$ may arise in our model.

Specifically, we assume that governments have given the DSB a mandate to serve an “interpretive” role: if invoked, the DSB observes an unbiased but noisy signal of $\Gamma$, which can be thought of as the outcome of an independent investigation in which the DSB “interprets” the contract. The DSB then issues a “cease-and-desist” ruling $\tau^{DSB} = FT$ if its signal indicates $\Gamma \leq 0$, and it issues the ruling $\tau^{DSB} = P$ if its signal indicates $\Gamma > 0$.\footnote{We follow Maggi and Staiger (2011) and Staiger and Sykes (2017) and abstract from the possibility of negotiated settlements to a dispute (i.e. negotiations between the two governments over the importer government’s policy choice after the state $s$ is realized). See Maggi and Staiger (2018) for an analysis of trade disputes that features the possibility of settlement.} The DSB ruling can therefore be thought of as simply a policy determination that maximizes the expected joint payoff of the governments given the DSB signal.\footnote{See Maggi and Staiger (2011) for a defense of this interpretation of DSB rulings in the context of the GATT/WTO.} We assume that the ruling is automatically enforced, and we denote the probability that the DSB issues the “wrong” ruling by $q_k(s) \in (0, 1/2)$ where $k(s) \in (0, 1/2)$ for all $s$ and $q \in (0, 1)$ parameterizes the (inverse) quality of the court.\footnote{Maggi and Staiger (2011) derive conditions under which it is optimal for the governments to write a vague contract and install a court with a mandate to interpret the contract if invoked. We take these two institutional features as given so that we may focus on other dimensions of the design of dispute settlement procedures.} Invoking the court is costly and we write the litigation costs incurred by Home’s government and Foreign’s agent $a$ as $c(s)$ and $c^*_a(s)$ respectively.

The timing of events is as summarized in Figure 1. First, the state of the world is realized and either $s \in \sigma^{FT}$ or $s \in \sigma^P$. Then, the Home government moves and makes its binary policy choice $\tau \in \{FT, P\}$. If the Home government chooses $\tau = FT$, the Foreign complainant has no incentive to invoke the DSB and free-trade prevails. If the Home government instead chooses $\tau = P$, the Foreign complainant has to weigh its options and may or may not invoke the DSB. If the Foreign complainant does not invoke the DSB, protection prevails. If the Foreign complainant instead invokes the DSB, the DSB moves and issues its ruling $\tau^{DSB} = FT$ or $\tau^{DSB} = P$ which is then adopted instantaneously.

The game is straightforward to solve by backwards induction. Consider first the Foreign complainant’s filing behavior. If the agreement includes only SSDS, then only the Foreign government has standing to file a complaint, and the ratio of its court costs to “court stakes” (the payoff from winning in court) is given by $\frac{c^*_G(s)}{|\Gamma^*_G(s)|} \equiv \mu^*_G(s)$. If the agreement includes both SSDS and ESDS, then both the Foreign government and the Foreign export industry
have standing, and in principal either may file a complaint.\textsuperscript{16} In this latter case, which we will indicate with the subscript \(G^* \& E^*\), we can define the minimum ratio of court costs to court stakes across the Foreign agents with standing:

\[
\frac{c^*_{G^* \& E^*}(s)}{\gamma_{G^* \& E^*}(s)} = \min\left\{ \frac{c^*_{G^*}(s)}{\gamma_{G^*}(s)}, \frac{c^*_{E^*}(s)}{\gamma_{E^*}(s)} \right\} \equiv \mu^*_{G^* \& E^*}(s).
\]

It then follows that for \(f \in \{G^*, G^* \& E^*\}\) and the designation of standing that \(f\) implies, a complaint is filed in state \(s\) if and only if \(\tau = P\) and

\[
\Pr(\text{DSB ruling is } FT \mid s) > \mu^*_f(s). \tag{3.1}
\]

According to (3.1), the Home government’s selection of \(\tau = P\) will be met with a filing in state \(s\) if and only if there is a Foreign agent with standing such that, for that agent, the probability of winning in court is greater than the ratio of court costs to court stakes so that the expected benefit of filing exceeds the cost of filing.\textsuperscript{17}

Next consider the Home government’s policy choice. Defining the ratio of the Home government’s court costs to court stakes by \(\mu(s) \equiv \frac{c_s}{\gamma(s)}\), the Home government chooses \(\tau = P\) in state \(s\) if either (3.1) fails – because then \(\tau = P\) can be set without triggering a dispute – or if (3.1) holds and the expected benefit to the Home government from trade protection exceeds the cost to the Home government of a dispute:

\[
\Pr(\text{DSB ruling is } P \mid s) > \mu(s). \tag{3.2}
\]

We can now derive the equilibrium actions for each state \(s\). For simplicity, in what follows we assume that the states where the vague contract is unambiguous are measure zero, so we can focus only on states where the court if invoked must interpret the contract.\textsuperscript{18} We also assume that dispute costs are low relative to dispute stakes in the specific sense that

\[
\mu(s) + \mu^*_f(s) < 1 \text{ for } f \in \{G^*, G^* \& E^*\} \text{ and all } s. \tag{Assumption 1}
\]

\textsuperscript{16}Where the Foreign export industry has standing, we abstract from possible free-rider problems that could arise and interfere with the ability of the industry to file even when it is in its interests to do so. As will become clear below, our results are robust to the inclusion of free-rider problems provided that they are not so severe as to prevent the Foreign export industry from ever filing.

\textsuperscript{17}In writing (3.1), we are implicitly assuming that if at least one Foreign agent with standing would benefit from filing in response to \(\tau = P\), then in response to \(\tau = P\) a filing will occur (and as will become clear, the identity of the Foreign agent that files is immaterial for our results). We therefore abstract from the possibility of strategic behavior among the potential Foreign complainants that might arise in the case of \(f = G^* \& E^*\) if each potential complainant prefers that a filing occur but each prefers that the other serve as the complainant.

\textsuperscript{18}This is without loss of generality, because under our assumptions in states where the vague contract is unambiguous the Home government would make the first best policy choice and there would be no filing by the Foreign complainant, and hence there would be nothing of consequence for any of the results we emphasize.
With a focus on relatively small dispute costs as embodied in Assumption 1, we direct attention to the parameter regions of the model where equilibrium disputes can arise. In particular, noting that \( \mu(s) < 1 - \mu_f^*(s) \) and \( \mu_f^*(s) < 1 - \mu(s) \) when the dispute costs are low relative to the dispute stakes in the sense of Assumption 1, conditions (3.1) and (3.2) imply the following result:

**Lemma 1.** Equilibrium actions in the presence of a trade agreement are as follows:

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

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

The content of Lemma 1 is depicted in Figure 2a, and follows an intuitive sorting along the dimension of DSB quality: if the DSB quality is high, the Home government makes the efficient policy choice and there is no dispute; if the DSB quality is intermediate, the Home government chooses \( \tau = P \) and there is a dispute; and if the DSB quality is low, the Home government chooses the inefficient policy and there is no dispute. Notice also that the court has its best impact off-equilibrium, when due to its high accuracy (i.e., for \( q_k(s) \leq \mu(s) \) in \( \sigma^{FT} \) and for \( q_k(s) \leq \mu_f^*(s) \) in \( \sigma^P \)) it induces both the Home government and the potential Foreign complainant to behave efficiently in order to avoid a dispute. Where a dispute arises in equilibrium there must be opportunistic behavior, either on the part of the Home government (for \( q_k(s) \in (\mu(s), 1 - \mu_f^*(s)) \) in \( \sigma^{FT} \), where the Home government is exploiting the incompleteness of the contract and the inaccuracy of the DSB and trying to get away with protection when free trade is efficient) or on the part of the Foreign complainant (for \( q_k(s) \in (\mu_f^*(s), 1 - \mu(s)) \)...

---

19If dispute costs are sufficiently high relative to dispute stakes so that Assumption 1 is violated, it is direct to show that equilibrium disputes cannot arise in our model, regardless of court quality. Assumption 1 therefore directs attention to the case emphasized by Shavell (1982) in his classic treatment of the potential inefficiency of the incentives to litigate, namely, the case where disputes can arise in equilibrium and the legal system is not “costless.” See also note 23.
in $\sigma^P$, where the Foreign complainant is exploiting the incompleteness of the contract and the inaccuracy of the DSB and trying to force free trade when protection is efficient). And finally, if the DSB is inaccurate enough (i.e., for $qk(s) \geq 1 - \mu_f^*(s)$ in $\sigma^{FT}$ and for $qk(s) \geq 1 - \mu(s)$ in $\sigma^P$) its beneficial off-equilibrium impact will erode, and such opportunistic behavior occurs while the DSB sits on the sideline. Also implied by Assumption 1 is that at least one of the thresholds $\mu(s)$ and $\mu_f^*(s)$ must be less than $1/2$. Given that we also have $qk(s) < 1/2$, it then follows that: (i) if both $\mu(s) < 1/2$ and $\mu_f^*(s) < 1/2$, the third region in both $\sigma^{FT}$ and $\sigma^P$ is empty; if only $\mu(s) < 1/2$, the third region in $\sigma^P$ is empty while the third region in $\sigma^{FT}$ is non-empty; and if only $\mu_f^*(s) < 1/2$, the third region in $\sigma^{FT}$ is empty while the third region in $\sigma^P$ is non-empty.

We can now write down the expected efficiency loss, relative to the first-best outcome, that is associated with standing choice $f \in \{G^*, G^*\&E^*\}$ in combination with the vague contract and interpretive court mandate, a combination of design features that we denote by $V_f$ and refer to as the $V_f$ institution. Denoting this efficiency loss by $L(V_f)$ and defining the sets $\sigma^{FT}_1 \equiv \{s \in \sigma^{FT} \mid qk(s) \leq \mu(s)\}$, $\sigma^{FT}_2 \equiv \{s \in \sigma^{FT} \mid qk(s) \in (\mu(s), 1 - \mu_f^*(s))\}$, and $\sigma^{FT}_3 \equiv \{s \in \sigma^{FT} \mid qk(s) \geq 1 - \mu_f^*(s)\}$; as well as $\sigma^P_1 \equiv \{s \in \sigma^P \mid qk(s) \leq \mu_f^*(s)\}$, $\sigma^P_2 \equiv \{s \in \sigma^P \mid qk(s) \geq (\mu_f^*(s), 1 - \mu(s))\}$, and, $\sigma^P_3 \equiv \{s \in \sigma^P \mid qk(s) \geq 1 - \mu(s)\}$, we can write:

$$L(V_f) = \sum_{s \in \sigma^{FT}_2 \cup \sigma^P_2} p(s) qk(s) |\Gamma(s)| +$$

$$\sum_{s \in \sigma^{FT}_2 \cup \sigma^P_2} p(s) [c(s) + c_f^*(s)] +$$

$$\sum_{s \in \sigma^{FT}_3 \cup \sigma^P_3} p(s) |\Gamma(s)|.$$ 

Each line of expression (3.3) captures a distinct source of inefficiency arising under the $V_f$ institution. The first line captures the loss associated with DSB error, and it is the product of the probability that state $s$ occurs, $p(s)$, the probability that the DSB makes a mistake, $qk(s)$, and the efficiency loss associated with that mistake, $|\Gamma(s)|$, summed over all states in which the DSB is invoked, $s \in \sigma^{FT}_2 \cup \sigma^P_2$. The second line captures the efficiency loss arising from the cost of a dispute, and it is the product of the probability that state $s$ occurs, $p(s)$, and the joint cost of a dispute in state $s$, $c(s) + c_f^*(s)$, again summed over all states in which the DSB is invoked, $s \in \sigma^{FT}_2 \cup \sigma^P_2$. The third line is the efficiency loss arising from distorted choices made “in the shadow of the court,” and it is the product of the probability that state $s$ occurs,
p(s), and the efficiency loss from getting the inefficient outcome in state s, \(|\Gamma(s)|\), summed over all states in which the DSB quality is so poor that the inefficient policy choice prevails without any dispute, \(s \in \sigma_{3,f}^{FT} \cup \sigma_{3}^{P}\).

3.2. The case for including ESDS in a trade agreement

We now evaluate the desirability of including ESDS in the agreement. When only SSDS is included (the \(V_G\) institution), only the Foreign government has standing and it alone has the right to file a dispute with the DSB. When in addition ESDS is included (the \(V_G\&E\) institution), the Foreign export industry also has standing to file a dispute with the DSB. Defining \(\Delta_{G\&E\&G} \equiv -[L(V_G\&E\&G) - L(V_G)]\) as the gain (if positive) or loss (if negative) in expected joint surplus that occurs when ESDS is added to SSDS, we can evaluate the desirability of including ESDS in the agreement on the basis of whether or not \(\Delta_{G\&E\&G}\) is positive.

According to the expression for \(L(V_f)\) given in (3.3), the sign of \(\Delta_{G\&E\&G}\) can be evaluated once we specify the complainant’s cost of filing \(c_f^*(s)\) and payoff from winning in court \(|\gamma_f^*(s)|\) under each choice of standing. This is because \(c_f^*(s)\) enters the expression for \(L(V_f)\) directly, and both \(c_f^*(s)\) and \(|\gamma_f^*(s)|\) enter \(L(V_f)\) indirectly by their impact on the sets \(\sigma_{2,f}^{FT}, \sigma_{2,f}^{P}\), and \(\sigma_{3,f}^{FT}\) through the filing condition (3.1).

We assume that the cost of filing is independent of the identity of the Foreign complainant. Formally, we assume that

\[
c_G^*(s) = c_E^*(s) = c^*(s) \text{ for all } s.
\]  

(Assumption 2)

Our key assumption, under which the choice of standing is consequential for the efficiency properties of the trade agreement, is that there exists at least one state \(s\) where (i) the payoff from winning in court is greater for the Foreign export industry than it is for the Foreign government and (ii) the Foreign government is at the margin between filing and not filing. Formally, and referring to the states described in (ii) as “marginal filing states,” we assume

\[
|\gamma_E^*(s)| > |\gamma_G^*(s)| \text{ for at least one marginal filing state } s.
\]  

(Assumption 3)

With these assumptions we capture the idea, as described in section 2, that under SSDS alone, the decision to file will always reflect the “public interest” in the Foreign country as embodied in the preferences of the Foreign government, while if ESDS is also included the decision to file will reflect the most aggressive of the interests represented by the two Foreign
agents with standing, namely, the Foreign government and the Foreign export industry, and will sometimes not reflect the public interest. To see this, note that Assumption 3 requires that, for at least one marginal state, the Foreign export industry would have more to gain from a win in court than would the Foreign government.\footnote{Assumption 3 would hold, for example, in any model where the removal of the Home tariff, by increasing the price of the Foreign export good, would hurt Foreign consumers of that good, or would hurt other Foreign producers who compete for the same factors of production and do not also export the same good to Home. The Foreign government would in some fashion (i.e., with non-zero weight) take all of these effects – both the gains to Foreign exporters and the losses to other agents in the Foreign economy – into account when calculating $\gamma^*_G(s)$; but Foreign exporters would ignore the cost imposed on Foreign consumers and other Foreign producers when calculating $\gamma^*_E(s)$. Assumption 3 could also reflect the fact that the Foreign government, unlike the Foreign export industry, might apply a “political filter” when evaluating the benefits of winning in court against Home in light of the broader diplomatic relations between the two countries.} In combination with Assumption 2 and according to (3.1), Assumption 3 therefore ensures that there is at least one state where a filing would not occur under SSDS alone but would occur if ESDS were also included, and hence ensures that when ESDS is included there is at least one state where filing reflects Foreign exporter over public interests. The question is then whether this feature, which embodies a commonly expressed fear associated with international dispute settlement systems that provide standing to private actors, might nevertheless serve the joint interests of the Home and Foreign government.\footnote{Notice from (3.1) that the Foreign export industry would also file more aggressively than the Foreign government if, contrary to Assumption 2, $c^*_G > c^*_E$, because then the industry would face lower dispute costs than its government. But such a dispute cost differential would not give rise to filings that reflect Foreign exporter over public interests, because in this case both Foreign interests would be served by the Foreign export industry’s more efficient (and therefore more aggressive) filing. With Assumption 2 we abstract from possible efficiency differences across potential claimants to focus instead on their different motives.}

To proceed we next define two sets that embody the two key changes that would occur according to Assumption 2 and Assumption 3 if standing were broadened to include not only the Foreign government but also the Foreign export industry. These sets are illustrated in Figure 2b. The first set is

$$\Omega^P_S \equiv \left\{ s \in \sigma^P \mid q_k(s) \in (\mu^*_E(s), \mu^*_G(s)) \right\}.$$

For $s \in \Omega^P_S$ the Foreign government would allow the efficient choice $\tau = P$ to go unchallenged under the $V^*_G$ institution but the Foreign export industry would file in response to $\tau = P$ under $V^*_{G\&E}$. The second set is

$$\Omega^{FT}_S \equiv \left\{ s \in \sigma^{FT} \mid q_k(s) \in [1 - \mu^*_G(s), 1 - \mu^*_E(s)] \right\}.$$

For $s \in \Omega^{FT}_S$ the Home government would choose the inefficient $\tau = P$ with impunity under the
$V_{G^*}$ institution but the Foreign export industry would file in response to $\tau = P$ under $V_{G^*E^*}$. An implication of Assumption 1 in combination with Assumption 2 and Assumption 3 is that at least one of these sets (and possibly both) must be non-empty. Using these sets, we have

$$\Delta_{G^*E^*,G^*} = -\sum_{s \in \Omega_S^F} p(s) [qk(s) \Gamma(s) + c(s) + c^*(s)] - \sum_{s \in \Omega_S^{FT}} p(s) \{[qk(s) |\Gamma(s)| + c(s) + c^*(s)] - |\Gamma(s)|\}. \tag{3.4}$$

Expression (3.4) highlights the costs and benefits associated with the additional litigation that arises when ESDS is added to SSDS. Given that $qk(s) > 0$ for $s \in \sigma^P$, the first line is clearly negative. It captures the fact that the added litigation in states $s \in \Omega_S^P$ is undesirable, since this litigation challenges an efficient policy and thus introduces nothing but court error and litigation costs: $qk(s) \Gamma(s)$ is the expected loss associated with court error, and $c(s) + c^*(s)$ is the joint litigation cost. The second line is also negative, reflecting the fact that the added litigation in states $s \in \Omega_S^{FT}$ is also undesirable, but the sign is less obvious this time. This is because the added litigation now challenges an inefficient policy so that the loss associated with court error and litigation costs $[qk(s) |\Gamma(s)| + c(s) + c^*(s)]$ is counterbalanced by an efficiency gain $|\Gamma(s)|$. However, using the filing condition of the Foreign government (equation (3.1) with $f = G^*$), it is easy to show that the loss associated with court error and litigation costs is greater than the efficiency gain.

Evidently, for $s \in \Omega_S^{FT}$ the Foreign government does not see a filing as worth the dispute cost while the Home government never benefits from a filing, and the fact that Foreign exporters would nevertheless choose to file simply reduces the value of the agreement to the two governments.

We may conclude that $\Delta_{G^*E^*,G^*} < 0$: the two governments would choose not to include ESDS in addition to SSDS in their trade agreement. We summarize with:

**Proposition 1.** Governments, but not their exporters, should have standing to bring disputes in an optimally designed trade agreement. That is, an optimally designed trade agreement should include SSDS, but not ESDS.

Intuitively, the finding contained in Proposition 1 reflects the fact that under SSDS the Foreign government is itself overly litigious at the margin of litigation – in the sense that it

---

22In particular, in states $s \in \Omega_S^{FT}$ the Foreign government would not have filed, which implies that $[1 - qk(s)] [-\gamma_G^*(s)] - c^*(s) < 0$. Using this inequality, it then follows that the term inside the curly brackets in the second line of (3.4) is positive: $[qk(s) |\Gamma(s)| + c(s) + c^*(s)] - |\Gamma(s)| = [1 - qk(s)] [\gamma_G(s) + \gamma_G^*(s)] + c(s) + c^*(s) = [1 - qk(s)] \gamma_G(s) + c(s) - [1 - qk(s)] [-\gamma_G^*(s)] - c^*(s) > [1 - qk(s)] \gamma_G(s) + c(s) > 0$. 

20
chooses to file, when from the perspective of the expected joint surplus of the two governments it should not file – owing to the negative effects on the Home government that the Foreign government does not internalize when making its filing decision.

This can be seen with reference to Figure 2b. For states in $\sigma^{FT}$, the upper panel of Figure 2b displays the margin of litigation for the Foreign government at the point $1 - \mu^{G*} (s)$, where $q_k (s)$ is such that the Foreign government is indifferent between challenging $P$ and letting $P$ go unchallenged. When $q_k (s)$ is at this margin, the Foreign government expects the benefit from possibly winning in court and flipping the Home government policy from $P$ to $FT$ to be equal to its litigation costs, and is indifferent between filing and not filing on this basis. But the Home government would strictly prefer that the Foreign government not file because it can then receive the surplus associated with $P$ and no litigation costs, whereas if the Foreign government files the Home government has to pay the litigation costs and faces the prospect with probability $(1 - q_k (s))$ that it will have to give up $P$ and switch to $FT$. So the expected joint surplus of filing at $q_k (s) = 1 - \mu^{G*} (s)$ is negative, and the Foreign government is therefore overly litigious at the margin of litigation in $\sigma^{FT}$. And it follows from the lower panel of Figure 2b that for states in $\sigma^{P}$ the Foreign government is also overly litigious at the margin of litigation from the perspective of the expected joint surplus of the two governments, because the Foreign government is challenging $P$ in court when $P$ is the efficient policy.

Viewed from this starting point, it is then clear that granting standing also to Foreign exporters must lower the expected joint surplus obtained by the governments under the agreement, because doing so would simply add to this overly litigious behavior. Notice also that the Foreign government’s behavior need not be overly litigious for inframarginal court quality realizations, and indeed for sufficiently high court quality (sufficiently low $q_k (s)$) the incentives of the Foreign government to litigate are efficient as we have observed. But at the realization of court quality that defines the filing margin for the Foreign government ($q_k (s) = 1 - \mu^{G*} (s)$ in $\sigma^{FT}$ and $q_k (s) = \mu^{G*} (s)$ in $\sigma^{P}$), which is what matters for the desirability of including ESDS in the agreement, the filing incentives of the Foreign government are inefficient and the nature of the inefficiency is unambiguous in our model: the Foreign government files when from the perspective of the expected joint surplus of the two governments it should not file.\footnote{Absent from our analysis is the remaining possibility for inefficient incentives to litigate described by Shavell (1982), namely, the possibility that the Foreign government does not file when from the perspective of the expected joint surplus of the two governments it should file. This possibility can only arise in our model when dispute costs are sufficiently high relative to dispute stakes so that Assumption 1 is violated (and even then
3.3. Standing for market access disputes more generally

While we have thus far analyzed market access issues in the context of trade agreements, similar issues could also arise in the context of investment treaties, given the close relationship between exporting and (horizontal) foreign direct investment (FDI). Exporting and FDI are typically viewed as two alternative ways of serving a foreign market, between which firms choose based on a proximity-concentration trade-off (Brainard, 1997; Helpman, Melitz, and Yeaple, 2004). Exporting has the advantage that it allows firms to concentrate production in one location, while FDI has the advantage that it allows firms to avoid trade costs, so that the optimal mode of accessing a foreign market is determined by the relative importance of plant-level economies of scale and trade costs.

In a companion paper (Ossa, Staiger, and Sykes, in progress), we exploit the close relationship between exporting and FDI and show that countries have a terms-of-trade motive for restricting the market access of foreign multinationals just as they do for restricting the market access of foreign exporters: both restrictions reduce the demand for foreign products and thus improve the terms of trade. But this motive is weaker in the context of FDI than in the context of exporting, because local affiliates of foreign multinationals typically also employ local factors which reduces the ability to shift costs to foreign countries. This implies that market access considerations should play less of a role in investment treaties than in trade agreements.24

In light of this, we view the results of this section as applying to market access disputes more generally, whether they arise in trade agreements or in investment agreements: with regard to market access/terms-of-trade issues, only governments should have standing to bring disputes in an optimally designed trade or investment agreement. That is, for the purpose of settling market access disputes, an optimally designed trade agreement should include SSDS, but not ESDS, while an optimally designed investment treaty should include SSDS, but not ISDS.

4. Standing in Investment Agreements

We consider next the issue of standing in investment agreements, and thus the question whether investment treaties should include only SSDS or also ISDS. For simplicity and consistent with only in $\sigma^{FT}$); and as we have noted, when Assumption 1 is violated the legal system is costless, because it will never generate an equilibrium filing and hence dispute. As discussed above, like Shavell we have chosen to de-emphasize the case of a costless legal system, by focusing on the case where Assumption 1 is satisfied; and in our model, that focus is then sufficient to rule out this remaining possibility.

24Sykes (2019) draws a similar conclusion.
our discussion in section 2, we assume that investment treaties are only concerned with helping
the Home – which we now refer to as Host – government make policy commitments to Foreign
investors (i.e., we abstract from any market access issues associated with Foreign investors).

To this end, we make two main modifications to our model of trade agreements in section
3. First, we add an ex-ante investment stage at which Foreign investors make their investment
choices and after which their investments are sunk. At the ex-ante stage, investors anticipate the
treatment they will receive from the Host government ex post, once their investments are sunk,
and this introduces the possibility of an ex-ante inefficiency arising from distorted investment
choices. Accordingly, in addition to an investment agreement focusing on the Host government’s
ex-post policy treatment of Foreign investors, we allow the Host government to offer up-front
investment incentives to address this ex-ante inefficiency. And second, once the investments
are sunk, the Host government implements an ex-post investment policy $\iota = \{FT, T\}$ (free
trade or “taking”), where $\iota = T$ is a stand-in for a variety of investment policies (e.g., tax,
regulatory, nationalization) that could amount to a taking broadly defined. This gives rise to
the possibility of an ex-post conditional-on-investment inefficiency analogous to the inefficiency
arising in our model of trade agreements, which the Host government can now address through
the investment treaty.

Otherwise, the model essentially follows our earlier structure. In particular, as with trade
agreements, real-world investment agreements include vague language which is subject to in-
terpretation, and we will therefore model the investment treaty as we did the trade agreement,
that is, as a vague contract combined with a DSB whose mandate is to interpret the contract
when invoked. For example, we could think of the vague contract as stating the following
(roughly corresponding to provisions in Article 6 of the US Model BIT):

“Foreign investments may be expropriated for a public purpose, provided that ade-
quate and effective compensation is promptly paid.”

Here, what constitutes “a public purpose,” and what suffices for “adequate and effective com-
pensation” paid “promptly,” are matters of interpretation. If invoked, the job of the court is
then to interpret such phrases as best it can and determine whether the contract has been
violated.25

25See also Janeba (2019) for a discussion of the vague language used in investment agreements and its impli-
cations for the role of the court in settling investment disputes.
Finally, while our simple reduced-form trade model is sufficient for capturing the market access issues highlighted in section 3, the commitment problem on which we focus here warrants a more explicit modeling of the investment setting and the ex-ante and ex-post inefficiencies that can arise. We begin by describing the underlying structure of our investment model in some detail.

4.1. Model Preliminaries

As in the model of section 3, we assume that \( s \equiv (s_1, s_2, \ldots, s_N) \) is a vector of state variables with each \( s_i \) corresponding to a binary event. To fix ideas, we consider a specific FDI opportunity in the Host country that, to exploit, requires a sunk capital investment by risk-neutral Foreign investors before the state of the world is revealed. We abstract from market access issues by assuming that the Host country is small in world capital markets, so that it faces an infinitely elastic ex-ante supply of Foreign capital at the world rate of return \( r^* \). Notice that under the small-country assumption foreign investors are assured an expected return of \( r^* \) regardless of the investment policies that the Host country is expected to ultimately adopt; so it makes sense that the investment agreement is not about (ex ante) market access into the Host country, since foreign investors have no reason to care about such issues there. It follows that only the Host government’s expected surplus can be impacted by its investment policies, and we can therefore evaluate the efficiency properties of the investment agreement according to the impact that the agreement has on expected Host-government surplus.

For simplicity we assume that a single Foreign investor makes a sunk investment \( I^* \), which is converted one-for-one into output through the production function \( Q = I^* \); and we assume that Host-country demand \( D(P, s) \) for this output in state \( s \) is elastic, \( \frac{D'(P, s)P}{Q} < -1 \), so that the investor always finds it optimal to sell all output ex-post. We denote the resulting market clearing price by \( \bar{P} = D^{-1}(I^*, s) \equiv \bar{P}(I^*, s) \) and note that \( \bar{P}(I^*, s) \) is decreasing in \( I^* \) for \( \bar{P} > 0 \). Host-country consumer surplus, conditional on a level of FDI \( I^* \), is then given by \( CS(I^*, s) = \int_{\bar{P}(I^*, s)}^{\infty} D(P, s) dP \) and is increasing in \( I^* \). Similarly, the ex-post (conditional on sunk investment \( I^* \)) Foreign operating profits or producer surplus is given by \( PS(I^*, s) = \int_{0}^{\bar{P}(I^*, s)} I^* dP = \bar{P}(I^*, s) I^* \) and is increasing in \( I^* \) given our assumption that demand is elastic. As a result, the sum of consumer and producer surplus in the market is also increasing in \( I^* \).

Production or consumption of this output may generate a negative (local) externality that is ignored by investors and individual consumers, and the realization of the state variables
determines the magnitude of the utility cost of the negative externality from the investment, which we denote by $e(I^*, s)$. We assume that $e(I^*, s) = e(s) I^*$ so that the utility cost is proportional to the size of the investment $I^*$ with $e(s) \geq 0$ for all $s$. Moreover, we assume that in any state for which a negative externality is present (i.e., any $s$ for which $e(s) > 0$), the externality $e(I^*, s)$ is large enough to turn the social value of the investment negative – that is, to ensure $PS(I^*, s) + CS(I^*, s) - e(I^*, s) < 0$ – for any positive investment level $I^*$.

This assumption simplifies the ensuing analysis by preserving the binary-policy-choice structure of the previous subsection, but it is not necessary for any of the results we emphasize. Under this assumption, the states of the world may be partitioned into those states $s \in \sigma^T$ where a negative externality turns the social value of the investment negative, and those states $s \in \sigma^{FT}$ where there is no externality ($e(s) = 0$) and the social value of the investment is positive and given by $PS(I^*, s) + CS(I^*, s)$.

Conditional on a given level of FDI $I^*$, the Foreign investor’s ex-post payoff is determined by the operating profits $\pi$ that it collects. If the investor is subject to a taking, the investor earns operating profits $\pi(I^*, T, s) = 0$ from the investment (regardless of the state of the world) implying an ex-post return on FDI of zero in that case. If the investor is not subject to a taking, the investor collects the market-clearing price $\hat{P}(I^*, s)$ for the output from the investment and therefore earns operating profits $\pi(I^*, FT, s) = \hat{P}(I^*, s) I^* = PS(I^*, s)$ implying an ex-post return on FDI of $\hat{P}(I^*, s)$ in that case.

The Host government’s ex-post payoff is more involved. For $s \in \sigma^T$, the Host government receives $\overline{\omega}(I^*, T, s) = 0$ if there is a taking and the output from the investment is destroyed, and if there is not a taking the Host government suffers the negative payoff $\overline{\omega}(I^*, FT, s) = CS(I^*, s) - e(I^*, s) < 0$. For $s \in \sigma^{FT}$, the Host government receives $\overline{\omega}(I^*, T, s) = CS(I^*, s) + \kappa PS(I^*, s)$ if there is a taking and receives $\overline{\omega}(I^*, FT, s) = CS(I^*, s)$ if there is not a taking, where $\kappa \in (0,1)$ parameterizes the degree of ex-post inefficiency associated with the Host-government’s taking in $\sigma^{FT}$. In particular, if $\kappa$ were equal to one then the taking in $\sigma^{FT}$ would amount to a zero-sum transfer of surplus from the Foreign investor to the Host government and would have no ex-post efficiency consequences (though, of course, possibly still ex-ante efficiency.

---

26 This amounts to an assumption that $e(s) > \hat{P}(0, s)$ whenever $e(s) > 0$, where $\hat{P}(0, s)$ is the “choke” price in state $s$ at which demand drops to zero.

27 Here and throughout, we use “over-bars” to distinguish notationally between functions in our analysis of investment agreements that also appear in our analysis of trade agreements.

28 There is no role for the parameter $\kappa$ when a taking occurs in $\sigma^T$, because by definition such takings are ex-post efficient.
consequences for the level of investment), whereas when \( \kappa \in (0, 1) \) the Host government gains less from the taking in \( \sigma^{FT} \) than the Foreign investor loses and the taking therefore amounts to a negative-sum transfer between the parties. As we will show, the role of \( \kappa \) is central to the nature of the inefficiency facing the Host government and to the potential role of an investment treaty in helping to address that inefficiency.

We can now define the ex-post gain that the Host government enjoys from a taking in state \( s \) as

\[
\bar{\gamma}_G (I^*, s) \equiv \bar{\omega} (I^*, T, s) - \bar{\omega} (I^*, FT, s) = \begin{cases} 
  e (I^*, s) - CS (I^*, s) & \text{for } s \in \sigma^T \\
  \kappa PS (I^*, s) & \text{for } s \in \sigma^{FT}
\end{cases}
\]

and the lost rents suffered by the Foreign investor in a taking as

\[
\bar{\gamma}_I^* (I^*, s) \equiv -PS (I^*, s) \quad \text{for all } s.
\]

The joint ex-post gain from a taking for the Host government and Foreign investor is then given by

\[
\bar{\Gamma} (I^*, s) \equiv \bar{\gamma}_G (I^*, s) + \bar{\gamma}_I^* (I^*, s),
\]

and we have

\[
\bar{\Gamma} (I^*, s) = \begin{cases} 
  -[PS (I^*, s) + CS (I^*, s) - e (I^*, s)] > 0 & \text{for } s \in \sigma^T \\
  - (1 - \kappa) PS (I^*, s) < 0 & \text{for } s \in \sigma^{FT}.
\end{cases}
\]

Hence, in states of the world \( s \in \sigma^T \) the policy that maximizes the joint ex-post surplus for the Host government and Foreign investor, which we refer to as the “first-best” ex-post policy, is complete expropriation (a taking) and destruction of the output from the investment (\( T \)); and in states of the world \( s \in \sigma^{FT} \) the first-best ex-post policy is no expropriation, amounting to a government policy that allows the sale of the output from the investment to proceed unhindered (\( FT \)).

Formally, and denoting by \( \iota_{FB} \) the first-best ex-post policy, we have \( \iota_{FB} = FT \) for \( s \in \sigma^{FT} \) and \( \iota_{FB} = T \) for \( s \in \sigma^T \). And if it anticipated this first-best ex-post treatment, the Foreign investor would receive an expected return on investment \( I^* \) of

\[
E_s [\rho_{FB}^* (I^*, s)] = \sum_{s \in \sigma^{FT}} p (s) \bar{P} (I^*, s),
\]

leading to a first-best investment level \( I^*_{FB} \) implicitly defined by

\[
\sum_{s \in \sigma^{FT}} p (s) \bar{P} (I^*_{FB}, s) = r^*
\]

and a first-best level for expected Host government surplus of

\[
E_s [\bar{\omega} (I^*_{FB}, \iota_{FB}, s)] = \sum_{s \in \sigma^{FT}} p (s) CS (I^*_{FB}, s).
\]
As (4.5) indicates, when the Host government adopts the first-best ex-post policy toward Foreign investment $t_{FB}$, it can expect to receive the consumer surplus generated in states $s \in \sigma^{FT}$ by the first-best level of Foreign investment.

We emphasize that our stylized description of a taking can be given a broad interpretation. The essential element is that states of the world exist in which the government action is socially efficient, and states of the world exist in which the government action is inefficient. For example, nationalization of an investor’s property may transfer it to a higher valued use, or it may lead to a lower valued use. New regulations applicable to an investment may impose costs that are less than the social benefits, or may impose costs that exceed the social benefits. And so on.

### 4.2. Limited Commitment and Up-Front Investment Incentives

If the Host government did not have access to any commitment technology, it would choose $t = T$ in all states of the world, expropriating the Foreign investment in good states ($\sigma^{FT}$) and bad ($\sigma^T$), owing to the sunk nature of the FDI at the time that the Host government makes its taking decision. This ex-post policy choice would correspond to the first best for $s \in \sigma^T$, but it would differ from the first best for $s \in \sigma^{FT}$. Moreover, anticipating this ex-post treatment and hence a zero ex-post return on FDI, in the absence of any further policy interventions no Foreign investment would be forthcoming, and the Host government’s surplus would be driven to zero in this market.

We assume that this stark commitment problem and the impact that it has on Foreign investment is potentially mitigated by two additional features of the Host country that exist whether or not the Host country also signs an investment agreement. First, we allow for the presence of domestic institutions which limit the Host government’s ability to expropriate Foreign investors. And second, we allow the Host government to offer an up-front investment incentive to compensate the Foreign investor for the ex-post treatment to follow.

Regarding the first feature, we have in mind domestic institutions that already protect the property rights of Foreign investors in the Host country, such as the domestic property law itself and the courts that enforce it. We capture this in a reduced-form fashion by assuming that the Host government is forced (e.g., by the domestic court) to implement the first-best ex-post policy with probability $\bar{p} \in [0, 1]$ and can act at its own discretion with probability $1 - \bar{p}$. We think of $\bar{p}$ as a parameter that varies across countries capturing differences in institutional quality, with $\bar{p} < 1$ signifying a lack of full commitment on the part of the Host government.
Denoting by $\iota_C$ the ex-post policies implemented under this regime of limited commitment, we thus have $\iota_C = FT$ with probability $\bar{p}$ and $\iota_C = T$ with probability $1 - \bar{p}$ for states $s \in \sigma^{FT}$ and $\iota_C = T$ in states $s \in \sigma^T$. Notice that these policies correspond to the first-best ex-post policies for $\bar{p} = 1$ and to the no-commitment policies for $\bar{p} = 0$. Under limited commitment and absent any further policy interventions, the Foreign investor therefore receives an expected return on investment $I^*$ of $E_s [\rho^*_C (I^*, s)] = \bar{p} \sum_{s \in \sigma^{FT}} p(s) \tilde{P} (I^*, s)$, since the Host government expropriates the Foreign investor unless a state $s \in \sigma^{FT}$ is realized and the Host government is constrained by domestic institutions to implement the first-best ex-post policy (which happens with probability $\bar{p}$).

Regarding the second feature, we allow the Host government to respond to its lack of full commitment to ex-post policies by offering an up-front payment of $\{r^* - E_s [\rho^*_C (I^*, s)]\} I^*$ to Foreign investors that is conditional on investing $I^*$. Since Foreign investors expect a return on their investment of $E_s [\rho^*_C (I^*, s)]$ given the Host government’s ex-post policy choices $\iota_C$, the additional payment $r^* - E_s [\rho^*_C (I^*, s)]$ makes the overall expected return from investing $I^*$ in the Host country exactly equal to the outside option $r^*$, and with these up-front payments the Host government can then induce any desired level of Foreign investment $I^*$ along the infinitely elastic Foreign investment supply curve, despite its lack of full commitment with regard to ex-post policies.

Before introducing the possibility of an investment agreement, we first consider how close the Host country can come to the first best without such an agreement when it exhibits these two features. With limited commitment and a program of up-front investment incentives, the Host government’s expected payoff for any level of $I^*$ is equal to

$$E_s [\bar{\omega} (I^*, \iota_C, s)] = \sum_{s \in \sigma^{FT}} p(s) [CS (I^*, s) + (1 - \bar{p}) \kappa PS (I^*, s)] - \{r^* - E_s [\rho^*_C (I^*, s)]\} I^* \quad (4.6)$$

The first line of (4.6) states the Host government’s payoff consists of this expected surplus minus the costs of the up-front investment incentive scheme $\{r^* - E_s [\rho^*_C (I^*, s)]\} I^*$ necessary to induce the investment level $I^*$. The second line of (4.6)
then follows using $E_s[p^*_C(I^*, s)] = \bar{\rho} \sum_{s \in \sigma^{FT}} p(s) \tilde{P}(I^*, s)$.

To determine the optimal level of investment to induce under an up-front investment incentive program, $\hat{I}^*$, the Host government solves $\max_{\hat{I}^*} E_s[\omega(I^*, \iota_C, s)]$. This yields the first-order condition

$$\sum_{s \in \sigma^{FT}} p(s) [\tilde{P}(I^*, s) - (1 - \bar{\rho})(1 - \kappa) \frac{\partial P_s(\hat{I}^*, s)}{\partial I^*}] = r^*,$$

which implicitly defines $\hat{I}^*$. As long as $\bar{\rho} < 1$ so that the Host government is limited in its ability to make commitments to Foreign investors, it is straightforward to show that $\hat{I}^*$ is larger than the level of Foreign investment that would prevail in the absence of the Host government’s up-front investment incentive program, which implies that the Host government’s expected surplus is higher with this program than without it.\footnote{Absent the up-front investment incentives program, the level of Foreign investment would be determined by the condition $\bar{\rho} \sum_{s \in \sigma^{FT}} p(s) \tilde{P}(I^*, s) = r^*$ which using (4.7) implies that this level of investment is smaller than $\hat{I}^*$ as long as $\bar{\rho} < 1$.} On the other hand, it is immediate from (4.4) and (4.7) that $\hat{I}^*$ is smaller than the first-best level of investment $I^*_{FB}$ as long as $\bar{\rho} < 1$ since $\frac{\partial P_s(\hat{I}^*, s)}{\partial I^*} > 0$ and $\frac{\partial P(\hat{I}^*, s)}{\partial I^*} < 0$.

Finally, with this optimal level of FDI secured by the appropriate up-front investment incentives program, the expected surplus enjoyed by the Host government is then given by

$$E_s[\tilde{\omega}(\hat{I}^*, \iota_C, s)] = \sum_{s \in \sigma^{FT}} p(s) \left[ CS(\hat{I}^*, s) + PS(\hat{I}^*, s) - (1 - \bar{\rho})(1 - \kappa) PS(\hat{I}^*, s) \right] - r^* \hat{I}^*$$

(Equation 4.8)

$$= \sum_{s \in \sigma^{FT}} p(s) \left[ CS(\hat{I}^*, s) + (1 - \bar{\rho})(1 - \kappa) \frac{\partial P(\hat{I}^*, s)}{\partial I^*} (\hat{I}^*)^2 \right].$$

Equations (4.4)-(4.8) imply that the up-front investment program cannot achieve the first-best level of expected surplus for the Host government as long as $\bar{\rho} < 1$ and the Host government is limited in its ability to make commitments to Foreign investors, since then:

$$E_s[\tilde{\omega}(\hat{I}^*, \iota_C, s)] < \sum_{s \in \sigma^{FT}} p(s) CS(\hat{I}^*, s)$$

$$< \sum_{s \in \sigma^{FT}} p(s) CS(I^*_{FB}, s)$$

$$= E_s[\tilde{\omega}(I^*_{FB}, \iota_{FB}, s)],$$

where the first inequality follows from $\bar{\rho} < 1$ and the second inequality follows from $I^*_{FB} > \hat{I}^*$. We summarize with:
Lemma 2. An up-front investment incentive program can help solve a Host-government’s commitment problem with regard to Foreign investors, but it cannot achieve the first best.

According to Lemma 2, up-front investment incentive programs fall short of fully solving the Host-government’s commitment problem. Intuitively, this is because such programs can address the ex-ante inefficiency due to under-investment, but they cannot address the ex-post inefficiency resulting from takings in states $s \in \sigma^{FT}$. The possibility of correcting these latter inefficiencies provides a natural role for an investment treaty. We next turn to assessing this role, and evaluate whether it is best served by an investment treaty that features only SSDS or also ISDS.

4.3. Investment Agreements

We are now ready to dive into our full analysis, allowing for both up-front investment incentives and an investment agreement. The timing of events is described in Figure 3. The investment agreement determines what happens once the investment is sunk and works exactly like the trade agreement in section 3. In particular, after the Foreign investment choice has been made, the Host government chooses its ex-post investment policy $\iota \in \{FT, T\}$; if it chooses a taking, the Foreign claimant (the Foreign government if only SSDS is included in the agreement, and either the Foreign government or the Foreign investor if both SSDS and ISDS are included in the agreement) then decides whether to file a complaint with the DSB; and if a complaint is filed, the DSB then issues a ruling based on a noisy signal of which ex-post policy is first-best in the realized state of the world. Together with the investment agreement, the up-front investment incentive determines the ex-ante investment level, and the Host government chooses the level of the investment incentive to maximize its expected payoff keeping the implications of the investment agreement in mind. We proceed by backwards induction, considering first the Foreign claimant’s filing choice, then the Host government’s ex-post investment policy choice, and finally the Host government’s up-front investment incentive choice.

Recall that for the Host government, the ex-post gain from a taking, $\tilde{\gamma}_G (I^*, s)$ as defined in (4.1), is positive in all states of the world, while the loss to the Foreign investor as defined in (4.2)

\[\text{In fact, our model overstates the likely effectiveness of up-front investment incentives in reality to solve the Host government’s commitment problem, because we abstract from incentive issues on the side of the investor (e.g., to deliver worthless investment once the up-front payment has been received). Nevertheless, as we will show, this serves a useful pedagogical purpose, in that it allows for a clean targeting of the up-front investment incentive to the ex ante investment inefficiency and permits targeting the design of the investment treaty to addressing the ex-post inefficiency.}\]
is given by \( \overline{\gamma}^* (I^*, s) = -PS (I^*, s) \). Similarly to our analysis of trade agreements in section 3, here we allow the Foreign government’s loss from a taking (and hence payoff from winning in court) to differ from that of the Foreign investor, and we capture the loss from a taking for each of these Foreign agents with the simple parameterization \( \gamma^*_I (s) \) for \( a \in \{G^*, I^*\} \), where the parameter \( \gamma^*_I (s) \equiv 1 \) by (4.2) and where we will later place restrictions on the parameter \( \gamma^*_G (s) \) but for now only assume that it is positive. We denote by \( c_a (\cdot) \) the cost incurred by Foreign agent \( a \) whenever it invokes the DSB, and we allow this cost to be a function of the level of investment \( I^* \), and in particular we assume that this cost rises in proportion to the magnitude of the producer surplus (operating profits) that is at stake in the taking. Formally, we assume that the cost incurred by Foreign agent \( a \) if it invokes the DSB in state \( s \) is given by \( c_a (I^*, s) \equiv \gamma^*_I (s) PS (I^*, s) \). We make the analogous assumption for the Host government: if the Host government is taken to court in state \( s \), it incurs a cost \( \bar{c} (I^*, s) \equiv \bar{c} (s) \gamma^*_G (I^*, s) \) to defend the taking.\(^{31}\)

As in section 3, we assume that the realized state \( s \) is observed by all agents including the DSB, and that \( \bar{I} \) is observed by the agents but not by the DSB.\(^{32}\) And as before, we will think of the DSB as issuing a policy ruling, in the present context denoted by \( i^{DSB} \) and corresponding either to \( FT \) or \( T \), to maximize the expected ex-post (once-the-foreign-investment-is-sunk) joint payoff of the Host government and the Foreign investor given its noisy signal of \( \bar{I} \).

It is worth pausing here to elaborate on what we have in mind more broadly regarding the DSB ruling. Under the interpretation that \( T \) represents a “regulatory taking,” the DSB ruling could be seen in a richer model as corresponding to a determination of whether the regulation complies with some explicit but vaguely worded commitment included in the contract, such as national treatment or the MFN clause, which itself can be interpreted as an attribute of internationally efficient policy intervention. Alternatively, under the interpretation that \( T \) represents an outright expropriation, the decision to expropriate could be left in the hands of the Host government subsequent to the DSB ruling under the interpretation that the DSB rules on a level of compensation to be paid by the Host government to the Foreign investor in the event of expropriation, with the ruling \( FT \) then corresponding to a level of compensation

\(^{31}\)As will become clear below, allowing dispute costs to rise with the level of foreign investment in this way ensures that the Foreign filing decision and Host policy choice are independent of the level of investment \( I^* \), simplifying the analysis to follow.

\(^{32}\)And we are assuming implicitly that the DSB cannot observe what the Host government does with the production facility if it expropriates it, i.e., whether or not the facility is shut down.
sufficiently high to prevent the Host government from following through with the expropriation and the ruling \( T \) corresponding to a level of compensation (which could be set arbitrarily to zero) under which the Host government would go through with expropriation.

Court behavior looks much like that under our trade agreement analysis of section 3. Consider first the Foreign complainant’s filing behavior. If the agreement includes only SSDS, then only the Foreign government has standing to file a complaint, and the ratio of its court costs to court stakes is given by \( \frac{c_{G^*}(I^*,s)}{\gamma_{G^*}(I^*,s)} = \frac{\bar{c}_{G^*}(s)}{\bar{\gamma}_{G^*}(s)} \equiv \bar{\mu}_{G^*}^* (s) \). If the agreement includes both SSDS and ISDS, then both the Foreign government and the Foreign investor have standing, and in principal either may file a complaint.\(^{33}\) We denote this case with the subscript \( G^* \& I^* \), and define the minimum ratio of court costs to court stakes across the Foreign agents with standing:

\[
\frac{c_{G^* \& I^*}(I^*,s)}{\gamma_{G^* \& I^*}(I^*,s)} \equiv \min\{\frac{c_{G^*}(I^*,s)}{\gamma_{G^*}(I^*,s)}, \frac{c_{I^*}(I^*,s)}{\gamma_{I^*}(I^*,s)}\} = \min\{\frac{\bar{c}_{G^*}(s)}{\bar{\gamma}_{G^*}(s)}, \frac{\bar{c}_{I^*}(s)}{\bar{\gamma}_{I^*}(s)}\} \equiv \bar{\mu}_{G^* \& I^*}^* (s). \]

It then follows that for \( f \in \{G^*, G^* \& I^*\} \) and the designation of standing that \( f \) implies, a complaint is filed in state \( s \) if and only if \( \mu = T \) and

\[
\Pr(\text{DSB ruling is } FT \mid s) > \bar{\mu}_{f}^* (s). \tag{4.9}
\]

According to (4.9), the Home government’s selection of \( \mu = T \) will be met with a filing in state \( s \) if and only if there is a Foreign agent with standing such that, for that agent, the expected benefit of filing exceeds the cost of filing.

Next consider the Host government’s ex-post policy choice, keeping in mind now that the Host government is constrained to implement the first-best policies with probability \( \bar{p} \). If the Host government is constrained, it chooses \( \mu = \mu_{FB} \). Otherwise, and defining the ratio of the Host government’s court costs to court stakes by \( \bar{\mu} (s) \equiv \frac{c_{G^*}(I^*,s)}{\gamma_{G^*}(I^*,s)} = \bar{c}(s) \), the Host government chooses \( \mu = T \) if either (4.9) fails – because then \( \mu = T \) can be set without triggering a dispute – or if (4.9) holds and the expected benefit to the Host government from a taking exceeds the cost to the Host government of a dispute:

\[
\Pr(\text{DSB ruling is } T \mid s) > \bar{\mu} (s). \tag{4.10}
\]

We can now derive the equilibrium actions, conditional on investment level \( I^* \), for each state \( s \). For simplicity and as before, in what follows we assume that the states where the

\(^{33}\)As with our analysis of trade agreements, we abstract here from a potential free rider problem that could arise when Foreign investors have standing and which could interfere with their ability to file even when it is in their interests as a group to do so, this time in the form of the possibility that several firms might be threatened by the same Host-government policy. But as we discussed in note 16, our results are robust to the inclusion of free-rider problems provided that they are not so severe as to prevent the Foreign investor from ever filing.
vague contract is unambiguous are measure zero, so we can focus only on states where the
court if invoked must interpret the contract. And as before, we also assume that dispute costs
are low relative to dispute stakes in the specific sense that

\[ \bar{\mu}(s) + \bar{\mu}_f^*(s) < 1 \quad \text{for } f \in \{G^*, G^* & I^*\} \]  

and all \( s \).  

(Assumption 1\(^\prime\))

Noting that \( \bar{\mu}(s) < 1 - \bar{\mu}_f^*(s) \) and \( \bar{\mu}_f^*(s) < 1 - \bar{\mu}(s) \) when the dispute costs are low relative to the dispute stakes in the sense of Assumption 1\(^\prime\), conditions (4.9) and (4.10) imply:

**Lemma 3.** Equilibrium actions in the presence of an investment agreement 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
      \( q_k(s) \leq \bar{\mu}(s) \), we have \( \iota = FT \) and no dispute; if DSB quality is intermediate in the
      sense that \( q_k(s) \in (\bar{\mu}(s), 1 - \bar{\mu}_f^*(s)) \), we have \( \iota = T \) and a dispute; if DSB quality
      is low in the sense that \( q_k(s) \geq 1 - \bar{\mu}_f^*(s) \), 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 \( q_k(s) \leq \bar{\mu}_f^*(s) \),
      and a dispute if \( q_k(s) > \bar{\mu}_f^*(s) \).
   
   2. If the Host government is unconstrained: If DSB quality is high in the sense that
      \( q_k(s) \leq \bar{\mu}_f^*(s) \), we have \( \iota = T \) and no dispute; if DSB quality is intermediate in the
      sense that \( q_k(s) \in (\bar{\mu}_f^*(s), 1 - \bar{\mu}(s)) \), we have \( \iota = T \) and a dispute; if DSB quality
      is low in the sense that \( q_k(s) \geq 1 - \bar{\mu}(s) \), we have \( \iota = FT \) and no dispute.

Note that Lemma 3 has a completely analogous structure to Lemma 1 if the Host government
is unconstrained by domestic institutions, in the sense that the equilibrium actions described by
Lemma 3 then mirror those of Lemma 1 and follow an intuitive sorting along the dimension of
DSB quality: if the DSB quality is high, the Host government makes the efficient policy choice
and there is no dispute; if the DSB quality is intermediate, the Host government chooses \( \iota = T \)
and there is a dispute; and if the DSB quality is low, the Host government chooses the inefficient
policy and there is no dispute. The only substantive difference between Lemma 3 and Lemma
is that the Host government is now constrained by domestic institutions with probability $\tilde{p}$ which forces it to implement the efficient policy no matter what. For future reference, we define

the sets $\tilde{\sigma}_1^{FT} \equiv \{s \in \sigma^{FT} \mid qk(s) \leq \tilde{\mu}(s)\}$, $\tilde{\sigma}_2^{FT} \equiv \{s \in \sigma^{FT} \mid qk(s) \in (\tilde{\mu}(s), 1 - \tilde{\mu}^*_T(s))\}$, and $\tilde{\sigma}_3^{FT} \equiv \{s \in \sigma^{FT} \mid qk(s) \geq 1 - \tilde{\mu}^*_T(s)\}$, as well as $\tilde{\sigma}_1^{T} \equiv \{s \in \sigma^{T} \mid qk(s) \leq \tilde{\mu}^*_T(s)\}, \tilde{\sigma}_2^{T} \equiv \{s \in \sigma^{T} \mid qk(s) \in (\tilde{\mu}^*_T(s), 1 - \tilde{\mu}(s))\}$, and, $\tilde{\sigma}_3^{T} \equiv \{s \in \sigma^{T} \mid qk(s) \geq 1 - \tilde{\mu}(s)\}$.

With Lemma 3 in hand, we can now roll back to the first stage and solve for the optimal investment incentives offered by the Host government. The Host government can induce any level of investment $I^*$ by offering up-front investment incentives of $\{r^* - E_s [\rho_f^* (I^*, s)]\}$ $I^*$, where $E_s [\rho_f^* (I^*, s)]$ is the expected return to the Foreign investor on an investment of $I^*$ under an investment treaty with standing choice $f \in \{G^*, G^* & I^*\}$. In the Appendix, we derive the expression for $E_s [\rho_f^* (I^*, s)]$ and use it and (4.3) to show that the expected Host government surplus associated with Foreign investment level $I^*$ is given by:

$$E_s [\tilde{\omega}_f (I^*, s)] = (1 - \tilde{p}) \sum_{s \in \tilde{\sigma}_1^{FT}} p(s) [CS (I^*, s) + PS (I^*, s)]$$

$$+ (1 - \tilde{p}) \sum_{s \in \tilde{\sigma}_2^{FT}} p(s) \left\{CS (I^*, s) + PS (I^*, s) + qk (s) \tilde{\Gamma} (I^*, s) - \tilde{c} (I^*, s) - \tilde{c}_f^* (I^*, s)\right\}$$

$$+ (1 - \tilde{p}) \sum_{s \in \tilde{\sigma}_3^{FT}} p(s) \left\{CS (I^*, s) + PS (I^*, s) + \tilde{\Gamma} (I^*, s)\right\}$$

$$+ (1 - \tilde{p}) \sum_{s \in \tilde{\sigma}_3^{T}} p(s) \left\{qk (s) [-\tilde{\Gamma} (I^*, s)] - \tilde{c} (I^*, s) - \tilde{c}_f^* (I^*, s)\right\}$$

$$+ (1 - \tilde{p}) \sum_{s \in \tilde{\sigma}_3^{T}} p(s) [-\tilde{\Gamma} (I^*, s)]$$

$$+ \tilde{p} \sum_{s \in \sigma^{FT}} p(s) [CS (I^*, s) + PS (I^*, s)]$$

$$+ \tilde{p} \sum_{s \in \sigma^{T}} p(s) \left\{qk (s) [-\tilde{\Gamma} (I^*, s)] - \tilde{c} (I^*, s) - \tilde{c}_f^* (I^*, s)\right\}$$

$$- r^* I^*.$$
lines of (4.11) record this ex-post expected joint surplus in the five sets of states where it is non-zero for the case that the Host government is unconstrained in its ex-post policy choice (which happens with probability $1 - \bar{p}$). And the next two lines of (4.11) add the ex-post expected joint surplus that is generated if the Host government is constrained to implement the first-best ex-post policies (which happens with probability $\bar{p}$). This interpretation can be confirmed by noting that the efficient policy yields a joint surplus of $CS(I^*, s) + PS(I^*, s)$ in states $s \in \sigma^{FT}$ and 0 in states $s \in \sigma^T$, the inefficient policy yields a joint surplus of $CS(I^*, s) + \kappa PS(I^*) = [CS(I^*) + PS(I^*) + \Gamma_s(I^*, s)]$ in states $s \in \sigma^{FT}$ and $-\Gamma_s(I^*, s) < 0$ in states $s \in \sigma^T$, the DSB makes a mistake with probability $qk(s)$, and invoking the DSB costs $c(I^*, s) + c_f(I^*, s)$.

The interpretation of $E_s[\tilde{\omega}_f(I^*, s)]$ carries with it an important implication: for any investment level $I^*$ that the Host government wishes to induce with its program of up-front investment incentives, the expected Host government surplus will be maximized by the choice of standing in an investment agreement that maximizes the ex-post expected joint surplus for the Host government and the Foreign investor (i.e., all but the last line of (4.11)). Put differently, while the optimal choice of standing in our analysis of trade agreements maximized the expected joint surplus of the Home and Foreign governments because those agreements addressed an inefficiency created by a government-to-government international policy externality, (4.11) reveals that the optimal choice of standing in an investment agreement maximizes the ex-post expected joint surplus of the Home government and the Foreign investor, because the investment agreement addresses an inefficiency created by a government-to-investor policy commitment problem.

We record this observation in:

**Lemma 4.** For an investment agreement that addresses an inefficiency created by a government-to-investor policy commitment problem, the optimal choice of standing maximizes the ex-post expected joint surplus of the Host government and the Foreign investor.

Lemma 4 suggests the possibility that the optimal choice of standing could be different across trade and investment agreements, despite the similar structures of Lemmas 1 and 3, due to the distinct problems that trade and investment agreements are designed to solve. We will return to this observation below when we evaluate the case for including ISDS in an investment agreement.
Finally, using the expression for $E_s[\bar{\omega}_f(I^*, s)]$ in (4.11), we can solve for the optimal level of FDI in the presence of an investment treaty with standing choice $f \in \{G^*, G'^* \cup I^*\}$, which we denote by $\bar{I}_f^*$, defined implicitly by $\frac{\partial E_s[\bar{\omega}_f(\bar{I}_f^*, s)]}{\partial I} = 0$. The expected Host government surplus under such an investment treaty, $E_s[\bar{\omega}_f(\bar{I}_f^*, s)]$, is then given by (4.11) evaluated with $I^* = \bar{I}_f^*$.

It is illuminating to compare the expression for $E_s[\bar{\omega}_f(\bar{I}_f^*, s)]$ in (4.11) with the expression for the expected Host government surplus under a program of up-front investment incentives but no investment treaty given in (4.6), also evaluated at $\bar{I}_f^*$, $E_s[\bar{\omega}(\bar{I}_f^*, \psi_C, s)]$. With these expressions it is straightforward to confirm that, as long as DSB quality satisfies $q_k(s) > \bar{\mu}_f^*(s)$ so that the union of the sets $\sigma_{2, f}^T$ and $\sigma_{3}^T$ is non-empty, there exists a critical level of $\bar{p} \in (0, 1)$ above which $E_s[\bar{\omega}(\bar{I}_f^*, \psi_C, s)] > E_s[\bar{\omega}_f(\bar{I}_f^*, s)]$, which with $E_s[\bar{\omega}(\bar{I}_f^*, \psi_C, s)] \geq E_s[\bar{\omega}(\bar{I}_f^*, \psi_C, s)]$ then implies $E_s[\bar{\omega}(\bar{I}_f^*, \psi_C, s)] > E_s[\bar{\omega}_f(\bar{I}_f^*, s)]$, implying in turn that the Host government cannot improve upon a program of up-front investment incentives by adding an investment agreement. Intuitively, this simply reflects the fact that with $\bar{p}$ above this critical level and hence with the Host government’s commitment problem largely (but still not completely) solved by its own domestic institutions, the introduction of an investment agreement with a DSB that will sometimes be utilized in equilibrium to challenge efficient takings (for $s \in \bar{\sigma}_{2, f}^T \cup \bar{\sigma}_{3}^T$) and that will sometimes make mistakes is not worth the added commitment potential that it provides.

We summarize with:

**Proposition 2.** Unless the quality of the DSB is above a threshold level, the introduction of an investment treaty can lead to efficiency gains and benefit the Host government relative to a stand-alone program of offering up-front investment incentives to Foreign investors if and only if the quality of domestic institutions in the Host country is sufficiently weak.

Notice that the possibility described by Proposition 2, that a purely unilateral intervention could dominate an international agreement for addressing inefficiencies in the context of investment policies, does not arise for the case of trade agreements. This difference reflects the different nature of the inefficiencies addressed across the two kinds of agreements, a domestic inefficiency in the case of investment treaties (stemming from a lack of commitment power on the part of the Host government) and an international inefficiency in the case of trade agreements (stemming from the terms-of-trade externalities associated with unilateral trade policy choices). And this difference has an important implication: with a domestic program of up-front incentives for Foreign investors in place, there is no guarantee that the perceived loss of Host-country
sovereignty that comes with the introduction of an (optimally designed) investment agreement and its associated dispute settlement procedure will be worth it in terms of the enhanced efficiency of investments that the agreement delivers. In particular, Proposition 2 suggests that countries with strong domestic institutions, such as might be expected of industrialized countries, may have legitimate concerns over the loss of sovereignty associated with the dispute settlement procedures of international investment agreements.

Henceforth we will assume that the Host-country commitment problem is sufficiently severe to allow the introduction of an investment treaty to improve upon up-front investment incentives by themselves, at least when standing is optimally allocated in the investment treaty. We next consider the optimal allocation of standing.

4.4. The case for including ISDS in an investment agreement

We now evaluate the case for including ISDS in an investment agreement. As with our analysis of standing in trade agreements in section 3, this requires that we adopt a stance on the Foreign complainant’s cost of filing and payoff from winning in court under each choice of standing, and in this regard we impose analogous assumptions to our earlier Assumption 2 and Assumption 3. In particular, we assume that the cost of filing for the Foreign government is the same as the cost of filing for the Foreign investor, namely

$$c^*_G(I^*, s) = c^*_I(I^*, s) \equiv c^*(I^*, s) \quad \text{for all } s \text{ and all } I^* \quad \text{(Assumption 2')}$$

which amounts to the parameter restriction $c^*_G(s) = c^*_I(s) \equiv c^*(s)$ for all $s$ since then $c^*_G(I^*, s) = c^*(s) \text{PS}(I^*) = c^*_I(I^*, s)$ for all $s$ and all $I^*$. And we assume that there is at least one marginal filing state in which the payoff from winning in court is greater for the Foreign investor than it is for the Foreign government:

$$|\gamma^*_I(I^*, s)| > |\gamma^*_G(I^*, s)| \quad \text{for a least one marginal filing state } s \text{ and all } I^* \quad \text{(Assumption 3')}$$

which amounts to the parameter restriction $1 - \gamma^*_I(s) > \gamma^*_G(s)$ for a least one marginal filing state $s$ since then $|\gamma^*_I(I^*, s)| = \text{PS}(I^*, s) > \gamma^*_G(s)\text{PS}(I^*, s) = |\gamma^*_G(I^*, s)|$ for this state $s$ and all $I^*$.

Similar to Assumption 2 and Assumption 3, Assumption 2’ and Assumption 3’ ensure that there is at least one state where a filing would not occur under SSDS alone but would occur if ISDS were also included, and hence ensures that when ISDS is included there is at least...
one state where filing reflects Foreign investor over public interests. As discussed in section 2, here we have in mind that the Foreign government applies a “political filter,” which takes into account the broader political, diplomatic and public relations repercussions of winning in court against Home. For example, because of the perceived implications for US jobs, the United States Trade Representative Robert Lighthizer recently expressed a dim view of American companies litigating expropriations through US agreements that contained ISDS provisions, stating:

“...The agreements allowed companies to litigate disputes with foreign governments over expropriations and other issues not through local courts, but through so-called investor-state dispute-settlement provisions. In doing so, the federal government effectively purchased political risk insurance for any American company that wanted to send jobs abroad.” (New York Times Op Ed, May 11 2020).

It is clear from Lighthizer’s statement that, if he had been in charge of the decision to invoke the ISDS provisions of these agreements, there would have been less litigation: this is the essence of Assumption 3′.34

Given Assumption 2′ and Assumption 3′, are there conditions (model parameter ranges) under which ISDS could be part of an optimally designed investment treaty? To answer this question, we do not derive expressions for the expected efficiency loss under each choice of standing relative to the first-best outcome as we did for the analysis of trade agreements in the previous subsection, because the difference in optimal investment levels across each choice of standing complicates such comparisons in the context of investment treaties. Instead we hold investment at the optimal level under an investment treaty with SSDS (I∗G), and we use (4.11) to calculate ΔG∗&I∗,G∗ ≡ Es[ωG∗&I∗(I∗G, s)] − Es[ωG∗(I∗G, s)], seeking conditions under which ΔG∗&I∗,G∗ is positive, which then provide sufficient conditions for Es[ωG∗&I∗(I∗G, s)] > Es[ωG∗(I∗G, s)] and hence for ISDS to be part of an optimal investment treaty, given that Es[ωG∗&I∗(I∗G, s)] ≥ Es[ωG∗&I∗(I∗G, s)].

34 As another example that motivates our Assumption 3′, if the series of tobacco plain-packaging disputes brought by Philip Morris under the ISDS provisions of various US BITs had instead been brought by the United States government, it seems plausible in light of the public controversy surrounding these disputes that the United States government would have enjoyed diminished gains from a win at court relative to the gains that would be enjoyed by Philip Morris (reflecting, for example, the political costs of having the name of the United States government associated with litigation aimed at weakening the health regulations of other countries and dealing with various constituencies in the United States on these issues).
To write down an expression for $\Delta_{G^{*} \& I^{*}, G^{*}}$, we first define the sets

$$
\tilde{\Omega}^T_S \equiv \{ s \in \sigma^T | qk (s) \in (\bar{\mu}_J^{*}, (s), \bar{\mu}_G^{*}, (s)) \} \quad \text{and} \\
\tilde{\Omega}^{FT}_S \equiv \{ s \in \sigma^{FT} | qk (s) \in [1 - \bar{\mu}_G^{*}, (s), 1 - \bar{\mu}_J^{*}, (s)) \}.
$$

As defined, the set $\tilde{\Omega}^T_S$ describes states in $\sigma^T$ where the Host government would implement a taking without court challenge under SSDS but would face litigation under ISDS, while the set $\tilde{\Omega}^{FT}_S$ describes states in $\sigma^{FT}$ where the Host government would implement a taking with impunity under SSDS but would face litigation under ISDS. With these new sets defined and using (4.11), we can now write

$$
\Delta_{G^{*} \& I^{*}, G^{*}} = - \sum_{s \in \tilde{\Omega}^T_S} p (s) \left[ qk (s) \tilde{\Gamma}(\bar{I}_G^{*}, s) + \bar{c} (\bar{I}_G^{*}, s) + \bar{c}^*(\bar{I}_G^{*}, s) \right] \\
- (1 - \bar{\rho}) \sum_{s \in \tilde{\Omega}^{FT}_S} p (s) \left[ qk (s) |\tilde{\Gamma}(\bar{I}_G^{*}, s)| + \bar{c} (\bar{I}_G^{*}, s) + \bar{c}^*(\bar{I}_G^{*}, s) - |\tilde{\Gamma}(\bar{I}_G^{*}, s)| \right].
$$

Expression (4.12) summarizes the costs and benefits associated with the additional litigation arising when ISDS is added to SSDS in the investment agreement. On its surface, it is completely analogous to our earlier expression (3.4) pertaining to trade agreements. As with (3.4), the first line of (4.12) describes the impacts of the extra litigation that challenges an efficient policy and hence introduces nothing but court error and litigation costs, and this line is clearly negative. And as with (3.4), the second line describes the impacts of the extra litigation that challenges an inefficient policy. However, unlike with (3.4) where the sign of the second line is also negative and ensures that ESDS in a trade agreement is never desirable, the sign of the second line of (4.12) is ambiguous, raising the possibility that in some circumstances the inclusion of ISDS in an investment agreement could be warranted.

This can be confirmed by using (4.3) to express the second line of (4.12) in equivalent form

$$(1 - \bar{\rho}) \sum_{s \in \tilde{\Omega}^{FT}_S} p (s) \left\{ [ (1 - qk (s)) PS (\bar{I}_G^{*}, s) - \bar{c}^* (\bar{I}_G^{*}, s) ] - [ (1 - qk (s)) \kappa PS (\bar{I}_G^{*}, s) + \bar{c} (\bar{I}_G^{*}, s) ] \right\},
$$

and noting that the first term in square brackets must be positive by the filing condition for Foreign investors in the set of states $s \in \tilde{\Omega}^{FT}_S$, that is, for states in $\sigma^{FT}$ where the Host government would implement a taking with impunity under SSDS but would face litigation under ISDS. Evidently, the entire expression will then be positive if the second term in square brackets is sufficiently small, which is guaranteed if $\kappa$ and $\bar{c}$ are sufficiently small, that is, if
the Host government is highly inefficient in orchestrating takings when expropriation is not socially efficient and bears little cost of defending itself in court. Moreover, the set of states over which this expression is summed, $\Omega^{FT}$, will be larger the smaller is $\gamma_{G^*}(s)$, that is, the greater the divergence is between the Foreign government and Foreign investors in the payoff to filing according to Assumption 3'; and more weight will be given to the expression the lower is $\bar{p}$, and hence the lower the quality of domestic institutions in the Host country. Finally, note that the first line of (4.12) will approach zero as the probability of states $s \in \sigma^T$ approaches zero. Hence, with the second line of (4.12) guaranteed to be positive if $\kappa$ and $\bar{c}$ are sufficiently small and large if $\gamma_{G^*}(s)$ and $\bar{p}$ are also small, adding ISDS to SSDS in an investment agreement will be desirable if it is sufficiently rare for a taking to be socially efficient.

Summarizing, we may now state:

**Proposition 3.** Whether Foreign investors, in addition to their governments, should have standing to bring disputes in an optimally designed investment treaty depends on a number of tradeoffs. But if expropriation is socially efficient only in unusual circumstances, if the quality of domestic institutions in the Host country is sufficiently bad, if the Host government is highly inefficient in orchestrating takings when expropriation is not socially efficient and bears little cost of defending itself in court, and if the Foreign government faces high political costs of initiating disputes, then it is optimal to give Foreign investors standing to bring disputes in an investment treaty. That is, under these conditions an optimally designed investment treaty should include both SSDS and ISDS.

The key to understanding Proposition 3 in relation to Proposition 1 is to recall that investment agreements are fundamentally government-to-investor agreements, as Lemma 4 indicates; and this means that Foreign investors will be overly litigious at the margin of litigation in the context of investment agreements, just as the Foreign government is overly litigious at the margin of litigation in the context of a government-to-government trade agreement. But in the context of investment agreements, the Foreign government, who by Assumption 3' is less litigious than Foreign investors, may then exhibit a degree of filing that is either closer to or farther from the socially efficient level relative to the filing behavior exhibited by the Foreign investor, raising an ambiguous possibility for the desirability of adding ISDS to SSDS that cannot arise under the analogous Assumption 3 in the context of trade agreements where the Foreign government is overly litigious and adding ESDS to SSDS can only make things worse.
More broadly, Propositions 1 and 3 imply that, in contrast to trade agreements, for investment agreements there is a case to be made for going beyond SSDS and including standing for private agents in the form of ISDS, though even here the case for ISDS is far from absolute. The intuition for these conclusions can be understood in two steps. A first step follows from our earlier observation that a trade agreement is a contract between governments to address a government-to-government international policy externality, while an investment agreement is a contract between the Host government and Foreign investors to address a government-to-investor policy commitment problem. So from the point of view of the filing decision the Foreign government is the principal in a trade agreement while the Foreign investor is the principal in the investment agreement. From this perspective, if the Foreign export industry is an imperfect agent of the Foreign government (which we impose in Assumption 3), it does not make sense to give Foreign exporters standing in trade disputes; and if the Foreign government is an imperfect agent of the Foreign investor (which we impose in Assumption 3'), it does not make sense to give the Foreign government standing in investment disputes. According to this first step, then, trade agreements should adopt SSDS while investment agreements should adopt ISDS, so that the filing authority always remains with the principal. The second step is then to observe that Foreign complainants never internalize the costs that filing a dispute imposes on the Home/Host government, and this negative externality leads to a general tendency to overfile relative to efficient litigation levels. Hence, beginning from the position that trade agreements should adopt SSDS, it would never make sense to introduce the possibility of an even more aggressive filer in the form of ESDS (as is implied by Assumption 3); whereas beginning from the position that investment agreements should adopt ISDS, it might make sense to restrict standing to a less aggressive filer in the form of SSDS (as implied by Assumption 3').

4.5. Standing for disputes over commitments to investors more generally

We have analyzed commitment issues with respect to foreign investors within the context of investment treaties. Arguably, similar issues may arise in the context of trade agreements: indeed, Yarbrough and Yarbrough (1992) argue that a central role for trade agreements is to allow importer governments to make policy commitments to foreign exporters who must make sunk investments in order to export to their markets.35 In principle, our analysis above could

---

35See also McLaren (1997) whose analysis of a trade agreement between a large and a small country turns this argument for trade agreements on its head.
be applied more or less directly to trade agreements wherever these agreements are designed to address such commitment issues, with an ESDS mechanism playing the role in trade agreements that is played by ISDS mechanisms in investment treaties.

However, we view such commitment issues as less important in the context of trade agreements than they are in the context of investment treaties, because the issue of sunk investments is likely to be more important in the context of FDI than in the context of exporting. There are several reasons to think that this distinction is important. First, and most obviously, there is a lack of any outright expropriation threat to the investments of exporters, in contrast to the case for FDI. But beyond this, in a multi-country world the investments made by exporters will commonly have alternative uses to produce exports for other markets — and to this extent therefore will not be sunk — whereas FDI would continue to be largely sunk and therefore highly susceptible to hold-up by the Host country in a multi-country world.

If one accepts this distinction, then it follows that the ex ante investment problem is more important in the context of investment treaties than it is in the context of trade agreements. And if setting up ISDS involves some fixed-cost component so that it is not worth doing below some minimal level of hold-up threat, then this distinction could account for the inclusion of ISDS provisions in investment treaties when the conditions of Proposition 3 are satisfied but no analogous inclusion of ESDS provisions in trade agreements to handle commitment issues there. Likewise, if it is difficult to create an ESDS mechanism that limits private standing to cases where serious commitment problems arise, but denies it for market access disputes, this would provide a further rationale for the exclusion of ESDS from trade agreements even when those agreements are designed to address important commitment issues.

5. Nature of the Remedy and Remedial Period

We have assumed thus far that when a case is filed and the court sides with the complainant, the defendant has no choice but to “cease and desist” the policy that has been found to be illegal. In reality, however, convicted violators of trade and investment agreements have the alternative option to continue their violation and compensate the injured party through some form of damage payment. There is a fundamental difference in the nature of damage payments across trade and investment agreements which is the source of much public controversy: while trade agreements allow the injured party to engage in reciprocal retaliation, investment treaties provide explicitly for cash payments.
In our Online Appendix, we extend our baseline models of sections 3 and 4 to explore this difference in the nature of the remedy. Rather than simply assuming that the Home or Host government switches to $\tau = FT$ or $\iota = FT$ when convicted by the DSB, 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, when convicted by the DSB, the Home or Host government switches to $\tau = FT$ or $\iota = FT$ in order 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 (with the trade agreement featuring SSDS and the investment treaty also including ISDS). Second, when cash payments are the remedy, we assume that these cash payments amount to costless transfers so that surplus can be efficiently transferred internationally (although incorrectly computed cash damages can of course still distort the behavior of the home/host government). 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.

Formally, in the Online Appendix we prove:

**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. (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.

In effect, Proposition 4 implies that it makes sense that investment treaties provide for cash
payments while trade agreements do not, provided that the cash value of the harm suffered by a foreign investor who is subject to a taking is sufficiently easy to quantify relative to the cash value of the harm suffered by a foreign government whose exporters face trade protection. In practice, and as we have noted in section 2, this condition seems plausible at least at a broad level, especially in light of the difficulty of placing a monetary value on the kinds of issues, such as lost jobs and distributional considerations, that typically dominate the trade concerns of real-world governments, and the relative ease by comparison of assessing lost profits from a (regulatory or outright) taking.

We have also thus far assumed that litigation is effectively instantaneous. When a case is filed, it is adjudicated immediately and the parties comply with the ruling immediately; hence, there is no possibility of any “pre-compliance harm” to exporters or investors. In our Online Appendix, we augment our baseline models from sections 3 and 4 to allow for the possibility of pre-compliance harm, and consider in the context of both trade agreements and investment agreements the following question: Can our models make sense of why trade agreements adopt prospective remedies while investment treaties adopt retrospective remedies? That is, can our models help us understand the conditions under which it might be optimal for investment treaties to include retrospective damages in the remedy (damages for harm suffered before the case is finally adjudicated) and for trade agreements to limit the remedy to only prospective damages (damages that would arise after adjudication if the ruling is not obeyed)?

To capture these alternative remedial periods, 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. We refer to $\delta$ as the degree of “litigation delay,” but $\delta$ should be interpreted more broadly: we have in mind all the costs of delay in adjudication including lost profits on export sales, losses experienced due to impairment of sunk investments, and so on. 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.

As we discussed in section 2, we think of lower values of $\delta$ as reasonable for the case of trade disputes, where the pre-ruling harm is attributable mostly to delay in securing market access, thus resulting in some lost profits if exports must be diverted to other markets temporarily, and so in our analysis of trade agreements we highlight model results with that parameter range in mind. By contrast, we think of higher values of $\delta$ as reasonable for the case of investment disputes, where the pre-ruling harm is usually attributable to policy actions that diminish
or destroy the returns to a sunk investment with a finite lifespan, and so in our analysis of investment treaties we highlight model results with that parameter range in mind.

To make our points on the remedial period, 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 in the Online Appendix collapse to the original models of sections 3 and 4, a feature that makes our comparisons easier but is not necessary for our results.

In the Online Appendix we prove:

**Proposition 5.** Provided that the quality of the DSB is above a threshold level: (i) 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; (ii) 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.

In effect, Proposition 5 can be understood as follows. If court quality is sufficiently high, it is optimal to adopt prospective remedies in trade agreements such as the WTO while it is optimal to adopt retrospective remedies in investment treaties because (i) the degree of pre-ruling harm is typically more severe in the context of investment disputes as compared to trade disputes so that prospective remedies become a poor option for effective investment agreements, and (ii) the available means of making international transfers are much less efficient in the context of trade disputes as compared to investment disputes, making retrospective remedies prohibitively expensive in the context of trade agreements and therefore unattractive in that context.

---

36 As 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).
International investment agreements employ dispute settlement procedures that differ markedly from their counterparts in trade agreements along three key dimensions: standing, the nature of the remedy, and the remedial period. In this paper we have developed parallel models of trade and investment agreements and have employed them to study these differences.

Our main finding is that many of the observed differences can be understood as deriving from the fundamentally different problems that trade and investment agreements are designed to solve. We have shown that the typical dispute settlement process in trade agreements - one that emphasizes state-to-state dispute settlement, tariff retaliation, and prospective damages - can be justified based on the government-to-government nature of trade agreements in combination with key features of the economic environment in which trade agreements operate. And we have shown that the typical dispute settlement process in investment agreements - one that emphasizes investor-state-dispute settlement, cash damages, and retrospective damages - can be justified based on the government-to-investor nature of investment agreements, in combination with key features of the economic environment in which investment agreements operate.

Our analysis also identifies some important qualifications to the wisdom of these design features, particularly with respect to the optimal design of investment agreements, thereby offering a potential explanation for the strong political controversy that investment agreements elicit. First, we have shown that relative to a stand-alone program of up-front investment incentives, investment agreements with an imperfect court can lead to efficiency gains only if the problem governments have in making commitments to foreign investors absent such an agreement is sufficiently severe. For liberal democracies with strong domestic institutions, this may help explain why some observers view such agreements as not worth the perceived sacrifice in sovereignty. Second, we have shown that including ISDS in addition to SSDS in an investment agreement is only optimal if governments are sufficiently poor agents for their investors: where government and investor incentives are only moderately misaligned, excluding ISDS from investment agreements may be the better option. And finally, to the extent that the purpose of an investment agreement is primarily one of securing market access rather than helping governments make commitments to foreign investors, our findings indicate that the design features of its dispute settlement procedures should not stray far from those of the typical trade agreement.
References


7. Appendix

7.1. Derivation of equation (4.11)

Given the equilibrium actions described by Lemma 3, the expected return on a level of FDI \( I^* \) from expected ex-post operating profits is given by:

\[
E_s [\rho_f^* (I^*, s)] = (1 - \bar{p}) \sum_{s \in \sigma^T \cup \sigma^T_u} p(s) \tilde{P}(I^*, s) \\
+ (1 - \bar{p}) \sum_{s \in \sigma^T_{2,f}} p(s) \left\{ \left[ 1 - qk(s) \right] \tilde{P}(I^*, s) - \frac{\bar{c}_f^*(I^*, s)}{I^*} \right\} \\
+ (1 - \bar{p}) \sum_{s \in \sigma^T_{1,f}} p(s) \left[ qk(s) \tilde{P}(I^*) - \frac{\bar{c}_f^*(I^*, s)}{I^*} \right] \\
+ \bar{p} \sum_{s \in \sigma^T} p(s) \tilde{P}(I^*, s) \\
+ \bar{p} \sum_{s \in \sigma^T_{2,f} \cup \sigma^T_{3,f}} p(s) \left[ qk(s) \tilde{P}(I^*, s) - \frac{\bar{c}_f^*(I^*, s)}{I^*} \right].
\]

Keeping in mind that the foreign investor must be offered an up-front investment incentive \( \{ r^* - E_s [\rho_f^* (I^*, s)] \} \) \( I^* \) in order to be willing to invest \( I^* \) given its outside option, the Host government’s expected payoff can be written as:

\[
E_s [\omega_f(I^*, s)] = (1 - \bar{p}) \sum_{s \in \sigma^T} p(s) CS(I^*, s) \\
+ (1 - \bar{p}) \sum_{s \in \sigma^T_{2,f}} p(s) \left\{ CS(I^*, s) + qk(s) PS(I^*, s) + qk(s) \bar{\Gamma}(I^*, s) - \bar{c}(I^*, s) \right\} \\
+ (1 - \bar{p}) \sum_{s \in \sigma^T_{3,f}} p(s) \left[ CS(I^*, s) + PS(I^*, s) + \bar{\Gamma}(I^*, s) \right] \\
+ (1 - \bar{p}) \sum_{s \in \sigma^T_{1,f}} p(s) \left\{ qk(s) \left[ -PS(I^*, s) - \bar{\Gamma}(I^*, s) \right] - \bar{c}(I^*, s) \right\} \\
+ (1 - \bar{p}) \sum_{s \in \sigma^T_{3,f}} p(s) \left\{ -PS(I^*, s) - \bar{\Gamma}(I^*, s) \right\} \tilde{P}(I^*, s) + \bar{p} \sum_{s \in \sigma^T} p(s) CS(I^*, s) \\
+ \bar{p} \sum_{s \in \sigma^T_{2,f} \cup \sigma^T_{3,f}} p(s) \left\{ qk(s) \left[ -PS(I^*, s) - \bar{\Gamma}(I^*, s) \right] - \bar{c}(I^*, s) \right\} \\
- \{ r^* - E_s [\rho_f^* (I^*, s)] \} I^*.
\]

Plugging the expression for \( E_s [\rho_f^* (I^*, s)] \) into the above expression for \( E_s [\omega_f(I^*, s)] \) yields equation (4.11) in the main text.
Figure 1: Timing of events (trade agreement)
Figure 2a: Illustration of Lemma 1

Figure 2b: From SSDS to ESDS
Figure 3: Timing of events (investment agreement)