DISCUSSION PAPER SERIES

DP14480

INVESTOR-STATE VS. STATE-STATE DISPUTE SETTLEMENT

Henrik Horn and Thomas Tangerås

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

INVESTOR-STATE VS. STATE-STATE DISPUTE SETTLEMENT

Henrik Horn and Thomas Tangerås

INVESTOR-STATE VS. STATE-STATE DISPUTE SETTLEMENT

Abstract

International investment agreements have been intensely criticized, and in particular the ⍰ISDS⍰ mechanisms that enable foreign investors to litigate against host countries. This paper examines the common claim that host countries benefi t from state-state dispute settlement (SSDS), since this yields less litigation. It assumes the standard rationale for ISDS, that SSDS causes political litigation costs. It shows how a host country might indeed bene fit from SSDS, but that there is no presumption that these conditions will prevail. Furthermore, negotiations regarding dispute settlement will plausibly yield ISDS, regardless of the distributional consequences for host countries, since SSDS is Pareto inefficient.

Henrik Horn - henrik.horn@econ-law.se
Research Institute of Industrial Economics (IFN), Bruegel, Brussels and CEPR

Thomas Tangerås - thomas.tangeras@ifn.se
Research Institute of Industrial Economics

INVESTOR-STATE VS. STATE-STATE
DISPUTE SETTLEMENT[1]

Henrik Horn[2]
The Research Institute of Industrial Economics (IFN), Stockholm
Bruegel, Brussels
Centre for Economic Policy Research, London

Thomas Tangerås[3]
Research Institute of Industrial Economics (IFN), Stockholm
Energy Policy Research Group (EPRG), University of Cambridge
Program on Energy and Sustainable Develoment (PESD), Stanford University

March 6, 2020

[2]Address: Research Institute of Industrial Economics (IFN); Box 55665; SE-102 15 Stockholm; Sweden. Email: henrik.horn@ifn.se.
[3]Address: Research Institute of Industrial Economics (IFN); Box 55665; SE-102 15 Stockholm; Sweden. Email: thomas.tangeras@ifn.se.

Abstract

International investment agreements have been intensely criticized, and in particular the "ISDS" mechanisms that enable foreign investors to litigate against host countries. This paper examines the common claim that host countries benefit from state-state dispute settlement (SSDS), since this yields less litigation. It assumes the standard rationale for ISDS, that SSDS causes political litigation costs. It shows how a host country might indeed benefit from SSDS, but that there is no presumption that these conditions will prevail. Furthermore, negotiations regarding dispute settlement will plausibly yield ISDS, regardless of the distributional consequences for host countries, since SSDS is Pareto inefficient.

Contents

1 Introduction

There are currently over 2 600 international treaties in force that seek to encourage foreign direct investment by protecting foreign investors and investment against host country policy interventions. These agreements were traditionally between a developed and a developing country, and addressed only investment. But it has become increasingly common to include the same form of investment protection also in trade agreement, including in agreements between developed countries. Investment agreements have been severely criticized during recent years.[1] For instance, has often been alleged that investment agreements include too onerous obligations to compensate investors in case of regulatory (indirect) expropriation—that is, for measure that have largely the same effect for investors as expropriation, but without formal take-over of assets.[2] Severe criticism has also been directed against the dispute settlement mechanisms in the agreements. Of particular concern has been that the agreements do not only allow contracting states to litigate—*State-State Dispute Settlement* (SSDS)—but typically also allow *foreign investors* to request arbitration—*Investor-State Dispute Settlement* (ISDS).[3,4]

Reflecting the mounting skepticism toward ISDS, some countries have recently sought to change their agreements so as to make it more difficult for investors to litigate outside the regular legal system of the host countries, or in a few cases to exclude entirely the legal standing of investors. For instance, the investment chapter of the US-Australia trade agreement of 2005, which includes investment protection, only allows for SSDS, and this only after investors have exhausted the possibilities to use local legal systems. South Africa is moving in the same direction.[5] Brazil never entered into ISDS agreements, and its new model investment treaty is based on SSDS. In side letters to the Comprehensive and Progressive Trans-Pacific Partnership, New Zealand has excluded ISDS with Australia and Peru, and reduced the scope for ISDS with three other partner countries. The recently revised version of NAFTA has drastically reduced the scope for ISDS; Canada has completely withdrawn from ISDS, and the possibility to use ISDS in investment disputes between Mexico and the US has been substantially reduced.

The purpose of this paper is to contribute to the understanding of the functioning of the ISDS mechanism. A common claim in the policy debate is that *ISDS causes excessive litigation* from a

---

[1] See e.g. Howse (2017) and Stiglitz (2008) for comprehensive overviews and discussions of the criticism against investment agreements.

[2] See Dolzer and Schreuer (2012) for an introduction to International Investment Law, and Bernasconi-Osterwalder (2014) for a discussion of legal aspects of SSDS.

[3] Two remarks on terminology: First, for practical reasons we will use the term "litigation" to denote what more correctly should be denoted "request arbitration" (since litigation is normally used in the context of civil lawsuits). Secondly, the term "ISDS" is often used synonymously with "investment agreements", but we here use it in its literal sense, as referring to a particular type of dispute settlement.

[4] See e.g. Gertz (2017), Johnson et al. (2015), Menon (2018), Roberts (2014), Salacuse (2007), and Trevino (2013) for recent discussions of SSDS, and other non-ISDS mechanisms.

[5] Th requirement that investors must first take their cases to host country courts, can serve as a filter that protects both home and host country governments, since investors might be deterred from bringing disputes if these domestic legal processes are slow, unless having strong cases.

host country point of view, and that host countries therefore would be better off with SSDS only. But other observers, who are also highly critical of certain aspects of investment agreements, argue that legal standing for private investors actually benefits host countries.[6] The purpose of the paper is more specifically to shed some light on this debate, by examining how the interests of the parties investment agreements are affected by the choice of dispute settlement system, and why the vast majority of agreements still allow for both ISDS and SSDS.

International treaties normally allow only contracting states to challenge alleged violations of the agreements. The possibility for private parties to litigate against foreign states under investment agreements is hence a rarity in international law. The formal analysis takes its starting point in a standard explanation in the legal literature for the introduction of ISDS: state-to-state disputes give rise *political or diplomatic tensions* that do not arise, at least not to the same extent, when private parties litigate on commercial grounds.[7] A number of high-profile disputes under investment agreements seems to support this notion. One example is the threat by TransCanada Corporation to litigate against the US regarding USD 15 billion in damages for the Obama administration's decision to disallow the construction of the Keystone XL pipe line; the decision was later overturned by the Trump administration. It is of course just a speculation, but it seems implausible that the Canadian government would have pursued the case had there only been SSDS in NAFTA. Another example is the still ongoing litigation by the energy company Vattenfall against Germany regarding the decision to speed up the phase-out of nuclear energy, made in the wake of the Fukushima accident. Vattenfall is a private limited liability company, but it is fully owned by the Swedish state. It seems unlikely that the Swedish government, which has made repeated commitments in the past to phase out nuclear power in Sweden, would have pursued this dispute. Yet another example is the litigation by Phillip Morris against Australia regarding its tobacco plain packaging legislation. The Obama Administration would hardly have been willing to litigate on behalf of the tobacco company.

The model to be employed builds on the analysis in Horn and Tangerås (2019, "H-T") of the regulatory (or indirect) expropriation rules in investment agreements, but with a simplified form of investment agreement. The economic setting is laid out in Section 2. The interaction takes place in three stages absent an agreement. A representative firm first makes an irreversible investment in a plant. Production in the plant can yield positive externalities for the host country in the form of increased incomes for employees, technology transfers, etc. But production can also have adverse consequences, such as e.g. environmental or health problems. The net effect of the investment on

---

[6]For instance, addressing President Trump concerning the NAFTA renegotiations, Howse (2017) and over two hundred academics argued in favor of legal standing of private investors, but put into question that NAFTA allows investors to by-pass the domestic legal system; see https://www.citizen.org/system/files/case_documents/isds-law-economics-professors-letter-oct-2017_2.pdf.

[7]See e.g. Vandevelde (2005, 174-175). Sykes (2005) points to additional advantages for investor; for instance, an investor might not have enough to offer its government to induce it to litigate on behalf of the investor, and retaliation by the home country need not lead to compensation for investors. Other possible sources of differences are that source country governments might put less value on compensation payments to do their investors, or that ISDS might allow for faster resolution of disputes.

the host country is stochastic, and becomes known after the investment has been sunk. Finally, having observed the shock, the host country decides whether to allow production or regulate, in the latter case effectively shutting down production.

Our framework captures two distortions that are central to problem that indirect expropriation provisions in investment agreements are typically meant to balance. First, foreign investors disregard the positive consequences of their investments for host countries, as well as the regulatory problems that their investment might cause. Second, when making regulatory decisions, host countries disregard investor interests, which creates a tendency toward overregulation. Simple as these distortion are when considered separately, their interaction can be rather complex. But they can lead to situations with underinvestment and overregulation from an ex ante perspective, both from the point of host country welfare, as well as joint welfare.[8] Hence the potential role for an investment agreement. The difficulty for the design of the agreements is how to constrain host country overregulation, while still allowing host countries to regulate when somehow desirable.

Section 3 very briefly describes a few salient features of investment agreements, and describes our formalization of such an agreement as it pertains to regulatory expropriations. We assume that the agreement is negotiated by the source and the host country at the outset of the interaction. It specifies *when* regulation is compensable, *how large* the compensation shall be, and whether investors or the source country government can litigate against the source country regarding refusal to pay compensation.[9] To capture salient features of actual agreements, we assume that the agreement requires the host country to fully compensate investors for foregone operating profits if regulation occurs when the regulatory shock is less severe than an agreed-upon level, denoted the *level of investment protection.* But the agreement allows the host country to regulate without compensation for regulatory shocks that are more severe than this. The level of investment protection is negotiated when the agreement is formed, and the type of dispute settlement system might also be negotiated, depending on the scenario under consideration.

Section 4 characterizes the outcome with an investment agreement. A compensation requirement will induce the host country to abstain from regulating for a range of shocks. But the host country might choose to regulate despite having to pay compensation for a range of more severe shocks. In such instances litigation is required to enact the compensation payments. To capture the above-mentioned standard rationale for ISDS, it is assumed that litigation under SSDS exposes the source country government to political costs that are not borne by private investors when litigating. A shift from ISDS to SSDS will thus impose a form of *enforcement cost* on the source country. This distinction between ISDS and SSDS is formally very simple, but does seem to capture the essence of the standard explanation for the inclusion of ISDS in investment agreements. Also, the implications for the outcome are more complex than might perhaps be thought. Furthermore, some

---

[8]This is not a conventional hold-up problem however, since it might be socially desirable to let the host country regulate in certain situations.

[9]Our companion paper H-T does not allow for different forms of dispute settlement mechanisms.

more elaborate formalizations are likely to yield similar predictions to those derived here, as will be argued below.

To illuminate the role of the form of dispute settlement, we employ two different approaches. In Section 5 we will consider the form of dispute settlement system as exogenously determined. The main purpose of this analysis is to highlight the interests of the respective party with regard to the form of dispute settlement system, interests that should be reflected in negotiations over both the level of investment protection and dispute settlement system. [10] As we will be shown, the consequence of switching from ISDS to SSDS will depend on two factors. The first is the magnitude of the political litigation costs with SSDS. For the agreement to be enforced with SSDS, these costs must be low enough that the source country will find it worthwhile to pursue disputes in order to obtain compensation payments—that is, an *enforcement constraint* need to be fulfilled in order for the agreement to be effective. Otherwise the host country will regulate when this is unilaterally optimal, knowing that there will not be any enforcement. This might appear desirable from a host country perspective. But this fails to recognize that the lack of enforcement will have consequences for firms' willingness to invest, and for the source country's interests in the bargaining over the level of investment protection.

The second factor that affects the outcome of switching to SSDS is whether with ISDS the host country regulates for certain realizations of the regulatory shock despite having to pay compensation. If it does, a switch to SSDS will have the direct consequence that the source country will be exposed to litigation costs. While this is of no concern to the host country as such, the expectation of such costs will affect the source country's benefit from an agreement, and thus the outcome of negotiations over the level of investment protection. To account for the likely full impact of switching to SSDS, we therefore need to trace out the effects for the whole sequence of events: negotiations, investments, regulatory decisions, and litigation decisions. There are here four types of implications.

One possibility is that the outcome is *unaffected* by the switch to SSDS. This will be the case when there are no disputes in the industry in the equilibrium with ISDS, and the enforcement constraint is fulfilled. A second possibility arises when the political litigation costs are low enough that the enforcement constraint is fulfilled with SSDS, but there are equilibrium compensation payments with ISDS. *The negotiated level of investment protection will then change*, but in ambiguous direction. However, *if* the level of protection were to fall somewhat, this would benefit the host country, along the lines suggested in the debate. But there is no particular reason to believe that the level of protection will fall, it might equally well rise.

The other two outcomes arise when the enforcement constraint binds at the level of invest-

---

[10]Investment agreements are sometimes denoted as "*boilerplate*" agreements, due to the remarkable degree of similarity across the thousands of investment agreements in existence; see e.g. Alschner and Skougarevskiy (2016). For instance, the agreements almost always allow for ISDS, and they typically include direct and indirect expropriation clauses. But the agreements tend to differ with regard to their substantive undertakings, and the extent to which include carve-outs from the compensation commitments. Exogenously switching from ISDS to SSDS can hence be seen as a change in the underlying boilerplate.

ment protection that has been negotiated with ISDS, but there are not necessarily any equilibrium compensation payments with ISDS. One possibility is the negotiated level of investment protection *increases sufficiently to provide the host source country with incentive to enforce the agreement.* The switch from ISDS to SSDS will then reduce the expected welfare of the host country. The other possibility is that the enforcement cost is sufficiently high that there is no level of investment protection that is simultaneously acceptable to the host country, and provides the source country with incentive to enforce the agreement. In this case ISDS is the only possible form of dispute settlement, *an agreement with SSDS would effectively unravel.* This unravelling would reduce host country welfare in our main framework, since it is assumed that there is scope for a Pareto improving ISDS agreement. But using an extension of the model we briefly discuss why a host country might at least in theory benefit from the unraveling of an existing agreement in certain industries.

The broader finding from the analysis of an exogenous switch from ISDS to SSDS is that we can indeed identify situations where a host country would benefit from the switch, along the lines suggested by some observers in the policy debate. This requires that the switch reduces the level of investment protection somewhat. But there is no presumption that this will occur. For the source country, a switch to SSDS has a direct negative impact if there are equilibrium compensation payments with ISDS, since it is then exposed to political litigation costs. But the impact on the source country is ambiguous for several reasons. One is that is not clear whether the level of protection will increase or fall. Another reason is that it is not clear whether an increase in the level of protection would benefit source country: it would increase the expected profits of investors. But a higher level of protection might lead to more costly disputes, and this reduces source country welfare. Hence, the effect of switching to SSDS are even more ambiguous for the host country. What is clear however, is that both countries might lose from the switch to SSDS, and that it is not possible for both to benefit.

The purpose of this analysis of an exogenous switch from ISDS and SSDS has been to understand the line-up of interests with regard to dispute settlement, and in particular to examine the notion that host countries would benefit from switching to SSDS. But despite our simple formalization of the difference between ISDS and SSDS, it is not possible to draw any clear conclusions. How should we then understand the fact that almost all actual agreements still allow for ISDS? To shed some light on this we assume in Section 6 that the dispute settlement system is negotiated simultaneously with the level of investment protection. The purpose is then to derive predictions regarding the equilibrium form of dispute settlement. When the two issues are negotiated simultaneously, the door is opened for the parties to trade off their different interests. For a given level of investment protection, the source country prefers ISDS to SSDS to avoid being exposed to political litigation costs, and under ISDS the host country prefers less protection than the source country. It is straightforward to show that when the parties negotiate both the level of protection and the dispute settlement system, they will *settle on ISDS in a broad range of bargaining formats.* This is consistent with practice, in that actual negotiations almost always result in both ISDS and SSDS, although SSDS is hardly

ever used in actual disputes.

It is not very surprising that SSDS is not the equilibrium outcome, in light of the fact that the only difference with ISDS is that SSDS introduces a costly friction: the litigation costs. At the same time, we saw that the host country benefitted from an exogenous switch from ISDS to SSDS, if this reduced the level of investment protection somewhat. So why would not a host country with significant bargaining power ensure that SSDS is the outcome, when both issues are negotiated simultaneously? Even if the host country would be better off with SSDS, since this yields a somewhat lower level of investment protection than with ISDS, this cannot be the equilibrium outcome. The bargaining outcome is assumed to be Pareto efficient, and any outcome with SSDS would be dominated by one with the same level of protection, but with ISDS.

Finally, turning to our contribution to the literature, it can first of all be noted that the theory literature on investment agreements is extremely meager in general, in particular compared to the large literature in trade agreements.[11] The only paper that we are aware of that examines differences between ISDS and SSDS as modes of dispute settlement is by Ossa, Staiger and Sykes (2019), who investigates dispute settlement in trade and investment agreements more broadly, including the choice between ISDS and SSDS. We will return to the relationship between their paper and this paper below.

2 The setting absent an investment agreement

We first describe the model under the assumption of no investment agreement. The setting is a special case of the framework used in H-T. Consider a country that is potential host to foreign direct investment from a source country in a number of industries. The industries might differ with regard to technology, demand, etc. The industries are economically unrelated, to remove complexities that do not seem to be of first-hand importance to the issues at stake here. In each industry investments can be made by a single, foreign firm. We could alternatively have assumed that there is a large number of symmetric investors, without qualitatively affecting the analysis below, as long as these investors are treated identically by the host and the source country. The lack of domestic firms in these industries is potentially less innocuous, since it excludes a role for National Treatment provisions, which typically are included in investment agreements. But it seems reasonable to disregard such complications in a first analysis of dispute settlement. These assumptions also has the advantage that we can omit industry and firm indices.

At the outset, the representative investor in each industry makes an irreversible investment $k \geq 0$ in the host country. The firm's investment cost $R(k) \geq 0$ is a strictly increasing, weakly convex function of the investment k. The investor receives the operating profit $\Pi(k) \geq 0$ if production is

---

[11] Aisbett et al (2010) provide a useful overview over the regulatory takings literature, and an introduction to central legal features of IIAs. Other contributions include Janeba (2019), Kohler and Stähler (2019), Konrad (2017), Schjelderup and Stähler (2019), Stähler (2018), and Horn and Tangerås (2019).

allowed; $\Pi(k)$ is strictly increasing and strictly concave in k, and $\Pi(0) = 0$. For the host country, an investment creates benefits in terms of consumer surplus, employment, technological spill-overs, learning-by-doing in the work-force, and so forth—the exact nature of these benefits is immaterial. After the investments have been undertaken, an industry-specific shock θ is realized that affects the net benefit to the host country of allowing production with investment k. High realizations of θ could represent the arrival of severely adverse information regarding environmental or health consequences of the production process or the goods produced, or other factors affecting the desirability of the investment. The shock is continuously distributed on $[\underline{\theta}, \bar{\theta}]$ with cumulative distribution function $F(\theta)$ and density $f(\theta)$.

Having observed this common-knowledge shock, the host country decides whether to permit or to regulate production. Regulation implies that production is effectively shut down, and thus deprives the firm of its operating profits: $\Pi(k) = 0$. This is the sole consequence of regulation for the source country. The host country welfare is $V(k, \theta)$ in case of production, reflecting the net of the positive and negative effects of the investment. The higher the realization of θ, the less beneficial is the investment for the host country: $V_\theta(k, \theta) < 0$ (subscripts on functional operators denote partial derivatives throughout). $V(k, \theta)$ can be either positive or negative in case of production, and it is zero if there is no production. The marginal net benefit of investment can also be positive or negative, $V_k(k, \theta) \gtreqless 0$.[12] To ensure that there is a role to play for investment and regulation, we assume that for every $k > 0$, the host country prefers production if the shock is sufficiently mild, $V(k, \underline{\theta}) > 0$, and prefers regulating if the shock is sufficiently severe, $V(k, \bar{\theta}) < 0$.

The interaction is solved for backwards in standard fashion. The last stage of the interaction is the decision by the host country whether to allow production or to regulate, given investment k and the realized regulatory shock θ. Since regulation yields zero welfare for the host country, and $V_\theta < 0$ for all $k > 0$ by assumption, it is sequentially rational for the host country to allow production whenever $V(k, \theta) \geq 0$ and to regulate if $V(k, \theta) < 0$. We assume that the host country allows production if indifferent. In that case, the host country regulates if and only if $\theta > \Theta(k)$, where

$$V(k, \Theta) \equiv 0. \tag{1}$$

In each industry the investment is made prior to the realization of the regulatory shock θ, and prior to the regulatory decision. To capture the notion that investors are small relative to their respective markets, we assume that representative investor does not take into consideration how its investment affects the probability of regulation.[13] If the investor expects regulation for $\theta > \theta'$, its

---

[12] Functions $\Pi(k)$, $R(k)$ and $V(k, \theta)$ are assumed to be twice continuously differentiable.

[13] We could e.g. have assumed that the industry is perfectly competitive, in which case it would have been natural to assume that each investor disregards the impact on the probability of regulation. H-T show that the first-best outcome might actually be easier to implement through an investment agreement if firms invest strategically with regard to the host country regulation. But this requires agreements that are contractually more sophisiticated than the type of agreement considered here.

expected profit is $F(\theta')\Pi(k) - R(k)$, and the optimal investment is

$$K(\theta') \equiv \arg\max_{k \geq 0}\{F(\theta')\Pi(k) - R(k)\}$$

with associated first-order condition (FOC)

$$F(\theta')\Pi_k(K(\theta')) = R_k(K(\theta'))$$

It follows from $F(\theta') > 0$, and $R_k > 0$, that $\Pi_k(K(\theta')) > 0$ in the relevant region, and from the second-order condition that $K_\theta(\theta') > 0$. The expected profit is

$$\tilde{\Pi}(\theta') \equiv F(\theta')\Pi(K(\theta')) - R(K(\theta')), \tag{2}$$

which must be non-negative in order for the firm to invest.

The equilibrium absent an investment agreement (k^N, θ^N) will then be given by $k^N = K(\theta^N)$ and $\theta^N = \Theta(k^N)$. Here θ^N is the sequentially rational cut-off value for regulation when investment is k^N, and the investment is k^N when the investor foresees the cut-off level θ^N for regulation.

The equilibrium expected profit is

$$\tilde{\pi}^N \equiv F(\theta^N)\Pi(k^N) - R(k^N)$$

and the equilibrium expected host country welfare is

$$\tilde{v}^N \equiv \int_{\underline{\theta}}^{\theta^N} V(K(\theta^N), \theta)dF(\theta).$$

There are two basic distortions at work. First, the investor disregards both the positive and the negative externalities from the investment that occur in the host country. Second, the host country disregards foreign investor profit in the regulatory decision. These distortions are straightforward, but their interaction is quite complicated. But unilateral investment decisions by foreign investors and unilateral regulatory decisions by host countries typically entail distortions of both investment and regulation.

3 An investment agreement

To remedy the distortions to investment and regulatory decisions, the host country and the home country can enter into an investment agreement at the outset of the interaction, before investments are made. Before formally introducing an agreement, we will say a few words about actual agreements.

As mentioned above, there are considerable similarities across agreements. They typically oblige

host countries to provide foreign investment Fair and Equitable Treatment, and there are non-discrimination rules requesting Most-Favored Nation Treatment and often also National Treatment. Agreements also contain provisions that require compensation in case of expropriation. Importantly for our analysis below, these provisions do not only cover direct expropriation where a host country formally takes over foreign investors' assets. They also cover *regulatory (indirect) expropriation*, which are measures with largely the same effect for investors as direct expropriation, but without formal take-over of assets.

The ambit of regulatory expropriation clauses has been a frequent source of contention in both the case law, and the policy debate. The ambit can be limited in at least two respects. The first lies in the definition of regulatory expropriation. Certain panels have taken the view that all that matters is the effect of government measures for investors, the intent behind the measure is irrelevant. But more recent panels have taken very different views, pointing to the "police powers exemption" in international law that allows states to protect public welfare.[14] The other restriction on the ambit of the rules regarding compensation in case of regulatory expropriation is through *explicit carve-outs* in the agreements, which are becoming increasingly common in agreements. For instance, it might be stated that the agreement shall not preclude the parties from adopting or enforcing any measure that is "necessary to protect human, animal or plant life or health", provided that the measure does not constitute "disguised protection", or "arbitrary or unjustifiable discrimination".

Virtually all investment agreements include rules concerning compulsory dispute settlement, allowing both source country governments and their investors to bring disputes. The disputes can be brought to host country courts, or to international arbitration. If brought to international arbitration, the disputes are normally arbitrated by three-person, ad hoc panels. Agreements often allow the disputes to be confidential at the request of either party. There are very limited possibilities of appealing panel determinations. Most countries are signatories to conventions that require courts in the signatory states to recognize and enforce arbitral awards from any other signatory state, so the enforcement mechanisms are strong compared to those in international agreement concerning e.g. trade or the environment. [15]

To formally capture the above salient features, we consider investment agreements that consists of two components. The first component is a *compensation function T* that stipulates compensation as a function of the underlying variables. With sufficient freedom to design an agreement, the parties should be able to achieve a first-best outcome. A critical issue is therefore the constraints that we impose on the contracting space. Compensation mechanisms in actual investment agreements share several features that are central from a contractual point of view:

---

[14]Often cited examples of the former approach are the panels in *Metalclad v. Mexico,* 1997, and *TECMED v. Mexico,* 2003. An example of the latter approach, which has also been adopted by a number of later panels, is *Methanex v. United States,* 2005.

[15]Dispute settlement mechanisms have also criticized for allowing investors to request arbitration outside host countries' legal systems, for arbitrators' alleged partiality, for the lack of appeal possibilities, for the lack of transparency of the arbitration proceedings and outcomes, and for the incoherence of the case law.

1. Agreements stipulate transfer payments to be made in certain situations if host countries regulate, and only if there is regulation;

2. There are no payments to or from outside parties;

3. There are occasionally carve-outs from compensation requirements for certain types of regulatory measures; and

4. Any compensation equals foregone operating profits.[16]

Investment agreement are hence highly incomplete in several regards. For instance, there is no direct contracting on investment levels or on regulation, investors and the host country retain unilateral discretion over these decisions. Also, there cannot be any direct subsidies, taxes on investors, or punitive damages. To represent the features listed above, for each industry the requested compensation is

$$T(k, \theta, \theta') \equiv \begin{cases} \Pi(k) & \text{if } \theta \leq \theta' \\ 0 & \text{if } \theta > \theta'. \end{cases} \tag{3}$$

That is, compensation is required if the regulatory shock θ is weaker than a threshold value θ', but not for shocks that are more severe than θ'. Furthermore, whenever regulation is compensable, the compensation should equal the foregone operating profits.[17]

The second standard component of the agreement is a specification of *who has legal standing to pursue a dispute*. As mentioned, investment agreements normally allow both investors and states to litigate. But states hardly ever use this option in practice. There are currently over 980 known treaty-based investment disputes. Almost all of the disputes have been pursued by investors rather than governments, despite the possibility for SSDS in these agreements.[18] To avoid introducing strategic interactions between the source country government and its investors regarding who should litigate, it is assumed that the agreement allow for either investor-state or state-state dispute settlement, that is, either ISDS or SSDS. This specification applies to all industries in both countries, as is almost invariably the case in IIAs. But the substantive undertakings—reflected in the level of protection θ' that the agreement specifies—is industry-specific. This is not a self-evident assumption. IIAs have few if any explicit industry specific obligations (the Energy Charter Treaty being a prominent exception). But as argued by H-T, it is conceptually unclear how to compare regulatory treatment across industries: what would it mean to say that an agreement gives the same degree of protection to investment in nuclear power, as to investment in the auto industry? The level of protection that an agreement imposes will thus in practice have to be determined for

---

[16] Some panels have used different criteria, such as the magnitude of the investors' investment.

[17] In H-T we consider a more general form of investment protection scheme that stipulates a level of investment protection that allows the level of investment protection to depend on the magnitude of the investment that has been made in the industry. Is is shown that such a compensation scheme has several desirable efficiency properties.

[18] UNCTAD Investment and from http://investmentpolicyhub.unctad.org/isds.

each industry separately. This is also reflected in treaty texts when stating that their substantive obligations should be interpreted in light of the specific circumstances at hand.

When the host and the source country have entered into an agreement specifying a level of investment protection θ', and a dispute settlement mechanism (ISDS or SSDS), the sequence of events is as follows in each industry:

1. The firm invests;

2. An industry-specific shock θ is realized;

3. The host country decides whether:
   - to allow production;
   - to regulate with compensation; or
   - to regulate without compensation; and

4. The investor or the source country decides whether to litigate.

To capture the central notion that governments face political and/or diplomatic costs when initiating a litigation against foreign states, we assume that SSDS gives rise to a cost $L > 0$ for the source country government that private investors are not exposed to. These political litigation costs could naturally arise in the relationship with the host country. But they can also stem from domestic politics. For instance, a source country government might experience reputational costs if it pursues a dispute on behalf of its tobacco industry. The variable L is not intended to reflect litigation costs in a traditional sense; it is not clear that conventional litigation costs would differ between firms and governments. We therefore set direct litigation costs to zero. An alternative would be to assume that plaintiffs are compensated for litigation costs if the arbitration court rules in their favor. We believe that this simple formalization of the difference between ISDS and SSDS captures the core aspect of the standard explanation for the inclusion of ISDS in investment agreements. We will briefly discuss alternative formulations in Section 7.[19]

When deciding to regulate, the host country also decides whether to pay compensation. There is no direct benefit in the model for the host country to defer paying compensation, face litigation and then be made to pay the compensation with certainty. But for political litigation costs to matter, there must at least potentially be litigation, and this requires that in certain instances the host country regulates without spontaneously paying compensation for compensable regulation. It is not straightforward to explain why rational parties would end up in a dispute in a symmetric

---

[19] Another possible difference between ISDS and SSDS is that with SSDS the government might litigate on behalf of a whole industry, or even investors in general, whereas with ISDS private investors typically litigate regarding their own compensation only. The numerous recent ISDS litigations against Spain regarding renewable energy is an example of the latter.

information-setting as ours (in particular not if disputes are costly).[20] At the same time we observe a large number of investment disputes in practice, despite the litigation costs that are involved. To avoid having to introduce elaborate asymmetric information reasons for the disputes, we assume that in the choice between regulating and paying compensation spontaneously, and regulating without compensation, the host country chooses the latter.[21] Intuitively, this could be explained by political gains for the host country government from being seen to resist challenges by foreign investors, or from discounting.[22]

4 The equilibrium outcome for a given investment agreement

We now derive the equilibrium investment, regulation and litigation for a given investment agreement. This means that we take the level θ' of investment protection in the compensation mechanism $T(k, \theta, \theta')$ and the dispute settlement mechanism, ISDS versus SSDS, as exogenously given. The only difference between the setting with ISDS and with SSDS for given θ', is the magnitude of the political litigation costs. Recall the assumption that $L = 0$ under ISDS and $L > 0$ under SSDS. We can therefore derive the outcome for both dispute settlement mechanisms by solving for the equilibrium for arbitrary $L \geq 0$. We solve the game by backward induction.

4.1 Litigation incentives

Assume that the representative firm has invested $k > 0$, the investment agreement features carve-out compensation $T(k, \theta, \theta')$, and the level of investment protection satisfies $\theta' > \theta^N$. The latter restriction ensures that the agreement offers protection for a range of situations where the host country would regulate if there was no agreement (the agreement is economically meaningless for $\theta' \leq \theta^N$). There will clearly be no litigation if $\theta > \theta'$ since regulation is not compensable in that case. We therefore consider the more interesting case of $\theta \leq \theta'$. The investor always litigates under ISDS when compensable regulation has occurred since the investor then receives compensation $\Pi(k) > 0$ without paying any litigation costs.[23] But since the source country government faces political litigation costs under SSDS, enough compensation must then be at stake relative to the litigation costs for litigation to occur. Specifically, we must have $\Pi(k) \geq L$. Hence, there is a minimal investment level $\bar{K}(L)$ that triggers litigation, where $\bar{K}(0) = 0$ and $\bar{K}(L) > 0$ for $L > 0$.

---

[20] Most formal explanations of disputes assume asymmetrically informed parties, partly stochastic determinations, etc. Aisbett et al (2010) provide an interesting analysis of an investment agreement with less than perfectly informed arbitrators.

[21] For instance, we could instead assume that the host country randomizes between regulating with and without spontaneously paying compensation when it knows that it will ultimately have to pay compensation. As long as there is some strictly positive probability for regulation to occur without compensation, the litigation costs would enter the picture in largely the same way as in what follows.

[22] See e.g. Salacuse (2007, pp.149) for a discussion of the political factors that might motivate governments to take disputes to formal arbitration.

[23] Section 7 discusses the impact of private litigation costs.

The larger is L, the larger the investments have to be to induce the source country government to litigate: $\bar{K}_L(L) > 0$. We will subsequently refer to $k \geq \bar{K}(L)$ as the industry *enforcement constraint*.

Lemma 1 *Under SSDS, there will be litigation if and only if the following three conditions are all met: (i) $k \geq \bar{K}(L)$; (ii) $\theta < \theta'$; and (iii) the host country has regulated without paying compensation.*

4.2 Regulation incentives

Consider next the host country's decision whether to permit production or to regulate under an agreement that stipulates a level of investment protection θ', when the investment is beyond the level that triggers litigation: $k \geq \bar{K}(L)$. The host country never regulates for sufficiently mild shocks $\theta \leq \Theta(k)$ since it then prefers production to regulation regardless of whether there is an agreement or not. It will always regulate if $\Theta(k) \leq \theta' < \theta$, since this requires no compensation according to the agreement, and regulation is desirable from the host country's unilateral perspective. In the intermediate range $\Theta(k) < \theta < \theta'$, regulation requires compensation, and the host country's optimal decision therefore is less clear. Let $\Theta^C(k)$ be the level of the regulatory shock for which the host country is indifferent between allowing production and regulating if it must compensate the industry for its foregone operating profit, i.e.

$$V(k, \Theta^C(k)) \equiv -\Pi(k) < 0,$$

if $V(k, \bar{\theta}) < -\Pi(k)$, and $\Theta^C(k) = \bar{\theta}$ otherwise. The left-hand side of the above expression captures host country welfare when allowing production, and the right-hand side is the host country welfare when production is regulated and the host country pays compensation. Observe that $\Theta^C(k)$ is the jointly optimal threshold for regulation because this is the cut-off level that maximizes the joint ex post welfare of the two countries of allowing production relative to regulating the industry.

For a given level of investment, there will be less regulation under the agreement compared to the situation absent an agreement: $\Theta^C(k) > \Theta(k)$, since $V_\theta < 0$. The host country will hence regulate if $\theta > \max[\Theta(k), \theta']$ since it then unilaterally prefers regulation and there is no compensation requirement, or if $\theta > \Theta^C(k)$ so that the host country prefers regulation *regardless* of whether this requires compensation payments.

Taking into account the subsequent enforcement incentives of the source country or its investors, the host country's incentives with regard to regulation are thus as follows:

Lemma 2 *If the investment agreement stipulates the investment protection level θ', and investment is k, the host country regulates in either of the following situations:*
(i) $k < \bar{K}(L)$ and $\theta > \Theta(k)$;
(ii) $k \geq \bar{K}(L)$ and $\theta' \leq \Theta(k) < \theta$;
(iii) $k \geq \bar{K}(L)$, $\Theta(k) < \theta' < \Theta^C(k)$ and $\theta > \theta'$; and
(iv) $\Theta^C(k) < \theta'$ and $\theta > \Theta^C(k)$.

13

In all four cases of Lemma 2, the regulatory shock is sufficiently severe that the host country would prefer to regulate as long as it does not have to pay compensation. In case (i) where $k < \bar{K}(L)$, there is no enforcement of the agreement. The host country therefore regulates whenever it is unilaterally optimal to do so, which is for $\theta > \Theta(k)$ by definition.

In case (ii) where $k \geq \bar{K}(L)$, there is enforcement for $\theta \leq \theta'$. However, the level $\theta' \leq \Theta(k)$ of investment protection is low enough that it will never constrain the host country. There will thus not be any violation of the agreement for any $\theta < \Theta(k)$. There will be regulation for all $\theta \geq \Theta(k)$, but the agreement does not request any compensation in this case.

In case (iii), investment protection is sufficiently strong, $\theta' > \Theta(k)$, that it in some cases will constrain the host country's behavior. Specifically, the host country will allow production for all $\Theta(k) < \theta \leq \theta'$ because the prospect of having to pay compensation makes it too costly to regulate. However, the host country will regulate for all $\theta > \theta'$ because it can do so without paying any compensation. We illustrate case (iii) in Figure 1. The horizontal axis measures the regulatory shock θ. The two horizontal lines at the bottom show the host country incentive to regulate, as it depends on whether compensation is required or not. The next line depicts the compensation requirement for an agreement with a level of investment protection $\Theta(k) < \theta' < \Theta^C(k)$. The uppermost line shows the resulting optimal behavior. There will be regulation if and only if $\theta > \theta'$ in this case.

(Fig 1: Regulation incentives with $\theta' < \Theta^C(k)$))

Case (iv) is where the host will regulate for $\theta > \Theta^C(k)$ although regulation might require compensation payments to the industry for $\Theta^C(k) < \theta < \theta'$. As further discussed in H-T, this implies that the simple compensation scheme in (3) induces the host country to *fully internalize* the externalities of its regulatory decision except litigation costs. It also means that there will be litigation in order to extract compensation. This scenario is illustrated in Figure 2. It is constructed identically to Figure 1, except that it pertains to the case $\theta' > \Theta^C(k)$. In what follows, we focus most attention to this latter case, since this is where compensable regulation occurs in equilibrium with ISDS, and political litigation costs with SSDS hence matter.

(Fig 2: Regulation incentives with $\theta' > \Theta^C(k)$))

Observe finally that investments have an ambiguous effect on the threshold $\Theta^C(k)$:

$$\Theta_k^C(k) = \frac{V_k(k, \Theta^C(k)) + \Pi_k(k)}{-V_\theta(k, \Theta^C(k))}$$

14

The denominator is positive. Therefore an increase in investment increases the jointly optimal threshold for regulation if and only if the joint marginal value of increased investment is positive, i.e. $V_k(k, \Theta^C(k)) + \Pi_k(k) > 0$. The joint marginal value of investment is generally ambiguous, and therefore the sign of $\Theta_k^C(k)$ is ambiguous.

4.3 Investment incentives

Investors behave non-strategically vis-à-vis the host country's incentive to regulate. As discussed above, if an investor expects industry investment to fall short of the critical level for enforcement $\bar{K}(L)$, the investor expects there to be regulation without compensation for $\theta > \theta^N$.[24] The expected profit is then

$$F(\theta^N)\Pi(k) - R(k)$$

and the optimal investment is $k^N = K(\theta^N)$. For this to be an equilibrium it must hold that $K(\theta^N) < \bar{K}(L)$.

If the expectation is instead that the industry investment will exceed the critical level for enforcement $\bar{K}(L)$, there are two possibilities. If $\theta' \leq \Theta^C(k)$, the investor expects to be allowed to produce for all $\theta \leq \theta'$, and expects regulation without compensation for all $\theta > \theta'$. If instead $\theta' > \Theta^C(k)$, the investor expects that the host country will not intervene for any shock $\theta \leq \Theta^C(k)$, that there will be regulation with full compensation for foregone operating profit for all shocks $\Theta^C(k) < \theta < \theta'$, and there will be regulation without compensation for $\theta > \theta'$. Hence, the cut-off for operating profit will be θ' in both cases, leading to the expected profit

$$F(\theta')\Pi(k) - R(k)$$

regardless of $\Theta^C(k)$. The optimal investment level is therefore $K(\theta') > k^N$. For this to be an equilibrium it is required also that $K(\theta') \geq \bar{K}(L)$.

In the subsequent analysis it will be more convenient to express the enforcement constraint as $\theta' \geq \bar{\Theta}(L)$ rather than $K(\theta') \geq \bar{K}(L)$. We define $\bar{\Theta}(L)$ as the level of investment protection that just suffices to trigger litigation in case of compensable regulation:

$$\Pi(K(\bar{\Theta})) \equiv L.$$

We can thus summarize the equilibrium investment behavior as follows:

Lemma 3 *If the investment agreement stipulates investment protection level θ', the investment will be:*
(i) $K(\theta')$ if $\theta' \geq \bar{\Theta}(L)$; and

---

[24]"Industry" since we assume there to be many firms in the industry, but for expositional reasons formally include a signle investor only.

(ii) k^N *if* $\theta' < \bar{\Theta}(L)$.

4.4 The equilibrium

Having derived the investment, regulation and litigation incentives, we can now characterize the expected welfare of the host country and the source country as functions of the investment protection θ' in the agreement, and the source country litigation costs L. To this end, and for future use, let θ^E be the level of investment protection that triggers an investment level $K(\theta^E)$ such that the host country will allow production for $\theta \leq \theta^E$, and it will regulate for $\theta > \theta^E$ *even if it has to pay full compensation.* θ^E is given by[25]

$$\theta^E \equiv \Theta^C(K(\theta^E))$$

We assume that

$$\theta' - \Theta^C(K(\theta')) > 0 \text{ iff } \theta' > \theta^E. \tag{4}$$

This property is similar to stability conditions used in e.g. oligopoly theory to rule out counter-intuitive comparative statics.[26] It implies that for $\theta' < \theta^E$ there will not be any regulation with compensation payments for any θ, whereas for $\theta' > \theta^E$ there will be a range of θ for which there is regulation with compensation.

Lemmas 1-3 jointly yield the following characterization of the outcome for any given level of investment protection θ':

Proposition 1 *Consider an agreement with the investment protection level $\theta' > \theta^N$ and litigation costs $L \geq 0$.*
(1) If $\theta' \geq \bar{\Theta}(L)$, the agreement yields:
(i)   investment $K(\theta')$;
(ii)   regulation iff $\theta > \min\{\theta'; \theta^E\}$; and
(iii) litigation and compensation payments iff $\theta' > \theta^E$ and $\theta^E \leq \theta \leq \theta'$.
(2) If $\theta' < \bar{\Theta}(L)$, the agreement yields:
(i)   investment k^N;
(ii)   regulation iff $\theta > \theta^N$; and
(iii) no litigation.

Proof: If $\theta' \geq \bar{\Theta}(L)$, the agreement will be enforced if compensable regulation occurs. The investment will hence be $K(\theta')$. If $\theta' \leq \theta^E$, then $\theta' \leq \Theta^C(K(\theta'))$ by (4). There will then be regulation iff $\theta > \theta'$, and since there is no requirement to compensate, there will not be any litigation. If $\theta' > \theta^E$, then $\theta' > \Theta^C(K(\theta'))$. There is regulation if and only if $\theta > \Theta^C(K(\theta'))$,

---

[25] An alternative way of characterizing the threshold θ^E is $V(K(\theta^E), \theta^E) \equiv -\Pi(K(\theta^E))$.
[26] It hence rules out the possibility that regulation occurs only since compensation is not required. Any regulation that the host country undertakes, it would undertake also if it had to pay compensation.

and compensation will be required for $\theta \in (\Theta^C(K(\theta')), \theta')$. Compensation will be enforced through litigation since $\theta' \geq \bar{\Theta}(L)$. The second part of the Proposition follows from Lemma 3.■

4.5 Expected welfare

On the basis of Proposition 1, we can derive the expected welfare of the two countries from an agreement with a given level of investment protection $\theta' > \theta^N$, and for given litigation cost L. Consider first the host country expected welfare. If $\theta' < \bar{\Theta}(L)$, the agreement will not be enforced. Assume therefore that $\theta' \geq \bar{\Theta}(L)$. Then there are two subcases. For $\theta' \leq \theta^E$, the host country will allow production and obtain welfare $V(K(\theta'), \theta)$ for all $\theta \leq \theta'$, and it will regulate without compensation payments for $\theta > \theta'$, in which case host country welfare will be zero. If $\theta' > \theta^E$, the host country will permit production and obtain welfare $V(K(\theta'), \theta)$ for $\theta \leq \Theta^C(K(\theta'))$. It will regulate and pay full compensation for all shocks $\Theta^C(K(\theta')) < \theta \leq \theta'$, in which case its welfare will be $-\Pi(K(\theta'))$. Finally, the host country will regulate without compensation payments for all shocks $\theta > \theta'$, in which case its welfare will be zero. The host country expected welfare $\tilde{V}(\theta')$ can thus be summarized as

$$\tilde{V}(\theta') \equiv \begin{cases} \int_{\underline{\theta}}^{\theta'} V(K(\theta'), \theta) dF(\theta) & \text{if } \theta' \geq \bar{\Theta}(L) \quad \text{and} \quad \theta' \leq \theta^E \\ \int_{\underline{\theta}}^{\bar{\Theta}^C(K(\theta'))} V(K(\theta')), \theta) dF(\theta) - \tilde{P}(\theta')\Pi(K(\theta')) & \text{if } \theta' \geq \bar{\Theta}(L) \quad \text{and} \quad \theta' > \theta^E \\ \tilde{v}^N & \text{if } \theta' < \bar{\Theta}(L). \end{cases} \quad (5)$$

where

$$\tilde{P}(\theta') \equiv F(\theta') - F(\Theta^C(K(\theta'))). \quad (6)$$

is the probability of litigation.

The source country expected welfare equals the expected investor profit minus the expected political litigation costs, if any. The agreement will not be enforced if $\theta' < \bar{\Theta}(L)$. If $\theta' \geq \bar{\Theta}(L)$, investors either are allowed to produce, or are regulated with full compensation for $\theta \leq \theta'$, and regulated without compensation for $\theta > \theta'$. Moreover, the source country will litigate for shocks $\Theta^C(K(\theta')) < \theta \leq \theta'$. The source country expected welfare thus equals

$$\tilde{Y}(\theta', L) \equiv \begin{cases} \tilde{\Pi}(\theta') & \text{if } \theta' \geq \bar{\Theta}(L) \quad \text{and} \quad \theta' \leq \theta^E \\ \tilde{\Pi}(\theta') - \tilde{P}(\theta')L & \text{if } \theta' \geq \bar{\Theta}(L) \quad \text{and} \quad \theta' > \theta^E \\ \tilde{\pi}^N & \text{if } \theta' < \bar{\Theta}(L) \end{cases} \quad (7)$$

In these expressions, the expected profit $\tilde{\Pi}(\theta')$ is given by (2).[27]

In order to ensure that there is scope for an agreement under ISDS, i.e. absent any litigation

---

[27] $\tilde{Y}(\theta', L)$ is continuous in θ', but has a kink at θ^E, since for $\theta' > \theta^E$ the home faces expected enforcement costs.

costs ($L = 0$), we assume that

$$\int_{\underline{\theta}}^{\theta'} V(K(\theta')), \theta) dF(\theta) > \tilde{v}^N \text{ and } \tilde{\Pi}(\theta') > \tilde{\pi}^N \text{ for some } \theta' > \theta^N. \tag{8}$$

5 ISDS vs. SSDS with negotiated investment protection

The previous section derived equilibrium investments, regulation and litigation, with exogenous determined investment protection and dispute settlement. In this section, we endogenize investment protection. The main issue we address is how the form of dispute settlement system affects the expected welfare of the two countries. We will assume that the agreement stipulates ISDS at the outset, and derive implications of switching to SSDS. This could capture a change in the underlying boiler plate for investment agreements, as discussed in the introduction. It will also serve to identify the interests of the parties, interests that will be at play as we consider negotiations over both the level of protection and the form of dispute settlement. We are in particular interested in the notion that ISDS yields too much litigation from the point of view of host countries, and that a switch to SSDS would remedy this problem. Can this notion be supported within our framework?

5.1 Negotiating the level of investment protection

The negotiations occur at the outset of the interaction, before the investment stage. We will not employ any specific bargaining solution, but instead assume that the bargaining maximizes some bargaining function $\hat{B}(\theta', L) = B(\tilde{V}(\theta'), \tilde{Y}(\theta', L)) > 0$ that is strictly increasing in $\tilde{V}(\theta')$ and $\tilde{Y}(\theta', L)$. We also assume that the marginal value of expected national welfare is non-increasing, $B_{VV} \leq 0$ and $B_{YY} \leq 0$, and that welfare levels are weak complements in the bargaining function, $B_{VY} \geq 0$. These properties are compatible, for instance, with the Nash Bargaining solution and with joint welfare maximization. The negotiated outcome is constrained by the two countries' participation constraints

$$\tilde{V}(\theta') \geq \tilde{v}^N \text{ and } \tilde{Y}(\theta', L) \geq \tilde{\pi}^N. \tag{9}$$

Let $\theta_{\max} \leq \bar{\theta}$ be the maximal investment protection acceptable to the host country. If $\theta_{\max} < \bar{\theta}$, then $\theta_{\max}$ is the maximal solution to $\tilde{V}(\theta_{\max}) = \tilde{v}^N$. We assume that the countries will enter into an agreement if and only if both participation constraints are satisfied and at least one country strictly benefits from the agreement. Since $\tilde{V}(\theta') = \tilde{v}^N$ and $\tilde{Y}(\theta', L) = \tilde{\pi}^N$ for $\theta' < \bar{\Theta}(L)$, the negotiated solution must satisfy also the enforcement constraint $\theta' \geq \bar{\Theta}(L)$. It follows by assumption (8) that $\theta_{\max} > \bar{\Theta}(0) = \underline{\theta}$. This condition will also be met under SSDS, for instance if L is sufficiently small. We will therefore proceed under the assumption that $\bar{\Theta}(L) < \theta_{\max}$ under SSDS; we will return to the case where $\theta_{\max} \leq \bar{\Theta}(L)$ at the end of this section.

Let $\Theta^*(L)$ be the level of investment protection that maximizes $\hat{B}(\theta', L)$ over $\theta' \in [\bar{\Theta}(L), \bar{\theta}]$ and subject to (9). If we let $v^*(L) \equiv \tilde{V}(\Theta^*(L))$ be the host country expected welfare and $y^*(L) \equiv$

$\tilde{Y}(\Theta^*(L), L))$ be the source country expected welfare under the negotiated agreement, then $\Theta^*(L)$ satisfies the first-order necessary condition

$$B_V(v^*(L), y^*(L))\tilde{V}_\theta(\Theta^*(L)) + B_Y(v^*(L), y^*(L))\tilde{Y}_\theta(\Theta^*(L), L) \leq 0 \qquad (10)$$

and boundary conditions

$$[\Theta^*(L) - \bar{\Theta}(L)][B_V(v^*(L), y^*(L))\tilde{V}_\theta(\Theta^*(L)) + B_Y(v^*(L), y^*(L))\tilde{Y}_\theta(\Theta^*(L), L)] = 0 \qquad (11)$$

$$\Theta^*(L) \geq \bar{\Theta}(L) \qquad (12)$$

in an equilibrium agreement that features incomplete investment protection $\Theta^*(L) < \bar{\theta}$, and where both countries strictly benefit from the agreement, $v^*(L) > \tilde{v}^N$ and $y^*(L) > \tilde{\pi}^N$. We first establish some fundamental properties of the negotiated agreement under ISDS and then turn to implications of a switch to SSDS.

5.2 Equilibrium investment protection under ISDS

The optimal investment protection $\Theta^*(0)$ under ISDS solves the first-order condition

$$B_{\tilde{V}}(v^*(0), y^*(0))\tilde{V}_\theta(\Theta^*(0)) + B_{\tilde{Y}}(v^*(0), y^*(0))\tilde{Y}_\theta(\Theta^*(0), 0) = 0 \qquad (13)$$

in an interior optimum $\Theta^*(0) < \bar{\theta}$. The marginal effect of investment protection on the source country under ISDS,

$$\tilde{Y}_\theta(\theta', 0) = f(\theta')\Pi(K(\theta')) > 0$$

with $\theta' = \Theta^*(0)$, is positive under ISDS because it increases expected profits without causing any litigation costs. This is implies by (13) that $\tilde{V}_\theta(\Theta^*(0)) < 0$. This implies directly that at the negotiated outcome $\Theta^*(0)$ with ISDS, the host country would benefit from a marginal *reduction,* and the source country from a marginal *increase,* in investment protection. In fact *any* increase $\theta' > \Theta^*(0)$ in investment protection would strictly benefit the source country and strictly hurt the host country. The first result is obvious from $\tilde{Y}_\theta(\theta', 0) > 0$. To see the second result, suppose $\tilde{V}(\theta') \geq \tilde{V}(\Theta^*(0))$ for some $\theta' > \Theta^*(0)$. We would then have $B(\tilde{V}(\theta'), \tilde{Y}(\theta', 0)) > B(v_0^*, y_0^*)$ by the assumption that the bargaining function is strictly increasing in both arguments. This would violate the presumed optimality of $\Theta^*(0)$. Note also that the source country would be lose from any reduction in the investment protection below $\Theta^*(0)$. But the host country would also lose if the decrease is large enough, by the assumption that the host country benefits from a higher level of protection than θ^N.

Hence, under ISDS the negotiated level of protection reflects a trade-off between the marginal benefit of increased investment protection from the viewpoint of the source country, and the marginal cost thereof from the viewpoint of the host country. Any increase in investment protection above

19

$\Theta^*(0)$ would strictly benefit the source country and strictly hurt the host country under ISDS: $\tilde{Y}(\theta', 0) > \tilde{Y}(\Theta^*(0), 0)$ and $\tilde{V}(\theta') < \tilde{V}(\Theta^*(0))$ for all $\theta' > \Theta^*(0)$. The host country hence prefers a lower level of protection than $\Theta^*(0)$, although not too much lower—it is still prefers a strictly higher level than θ^N.

Proposition 2 *Under ISDS, the source country would prefer a higher level of investment protection than the negotiated level $\Theta^*(0)$, $\tilde{Y}_\theta(\Theta^*(0), 0) > 0$, and the host country would prefer at least a marginally lower level of protection, $\tilde{V}_\theta(\Theta^*(0)) < 0$.*

Figure 3 illustrates these results. In both graphs, the horizontal axis shows the level of investment protection θ'. The upper graph shows the expected welfare of the source country, which is increasing in the level of investment protection from the level of θ^N. The lower graph gives the expected welfare of the host country under the simplifying assumption that this welfare is concave. The expected host country welfare reaches its global maximum at $\hat{\theta}$, after which it declines until hitting the reservation level for the host country at $\theta_{\max}$. Hence, there will be an interior solution $\Theta^*(0) \in (\hat{\theta}, \theta_{\max})$ that reflects the relative bargaining strength of the two parties. Since the graph assumes that $\theta^E < \hat{\theta} < \theta'$, it depicts a situation with equilibrium compensation payments: $\Theta^*(0) > \theta^E$.

(Fig. 3: The outcome of negotiations with ISDS)

5.3 Implications of switching from ISDS to SSDS

We will now compare the outcome with ISDS to that with SSDS. Our model allows for an analytically convenient method of analyzing the difference between the two dispute settlement systems, since the difference ultimately stems from difference in the magnitude L of the political litigation costs. We can therefore interpret an increase in L as a step from ISDS toward SSDS.

As can be seen from the expressions from expected welfare (5) and (7), the introduction of political litigation costs can affect the outcome through two distinct mechanisms, both of which will influence the negotiated level of investment protection, and the expected welfare of the parties. One is that the litigation costs can change the source country's benefit from the agreement, and thereby *affect the source country's bargaining position.* The other mechanism is that the litigation costs can *reduce the source country's incentive to enforce the agreement.* Intuitively, the first effect can arise already for small L, whereas the latter is more plausible when L is large. We consider each in turn, but we first establish the situation in which neither of these consequences of SSDS is present.

5.3.1 The outcome is the same

A first possibility is that ISDS and SSDS yield the same allocations. This will be the case if the negotiated level of protection with ISDS is low enough that there are no compensation payments

in equilibrium, specifically $\Theta^*(0) \leq \theta^E$, and the equilibrium level of investment protection under ISDS can be implemented also under SSDS, i.e. $\bar{\Theta}(L) < \Theta^*(0)$. Then, the magnitude of L has no direct effect on allocations, and so the negotiated outcome will be identical under the two dispute settlement systems: $\Theta^*(L) = \Theta^*(0)$.

Proposition 3 *Assume that $\Theta^*(0) \leq \theta^E$, so that the negotiated investment protection under ISDS does not yield any litigation, and that the same level of investment protection can be enforced also under SSDS, $\bar{\Theta}(L) < \Theta^*(0)$. The form of dispute settlement system then has no impact on equilibrium investment protection, or on expected welfare, in any of the two countries.*

Proof: Assume that $\Theta^*(L) = \Theta^*(0)$. $\Theta^*(L)$ can then be implemented under SSDS by the assumption that $\bar{\Theta}(L) < \Theta^*(0)$. Since there is no litigation in equilibrium under either dispute settlement system if $\Theta^*(L) = \Theta^*(0) \leq \theta^E$, it follows that $\hat{B}(\Theta^*(L), L) = \hat{B}(\Theta^*(0), 0)$. It can then not be beneficial to deviate from $\Theta^*(L) = \Theta^*(0)$ under SSDS:

(i) It cannot be beneficial to deviate to $\theta' < \bar{\Theta}(L)$ because then $\hat{B}(\Theta^*(L), L) = \hat{B}(\Theta^*(0), 0) > \hat{B}(v^N, \pi^N, 0) = \hat{B}(\theta', L)$, where the strict inequality follows from assumption (8).

(ii) It cannot be beneficial to deviate to $\theta' \in [\bar{\Theta}(L), \theta^E]$, $\theta' \neq \Theta^*(L)$, because then $\hat{B}(\Theta^*(L), L) = \hat{B}(\Theta^*(0), 0) \geq \hat{B}(\theta', 0) = \hat{B}(\theta', L)$, where the inequality follows from the assumed optimality of $\Theta^*(0)$ under ISDS, and $\hat{B}(\theta', 0) = \hat{B}(\theta', L)$ because there is no litigation for $\theta' \leq \theta^E$.

(iii) It cannot be beneficial to deviate to $\theta' > \theta^E$ because $\hat{B}(\Theta^*(L), L) = \hat{B}(\Theta^*(0), 0) \geq \hat{B}(\theta', 0) > \hat{B}(\theta', L)$, where the strict inequality follows from the presence of litigation costs L under SSDS for $\theta' > \theta^E$.

The equality of the levels of investment protection in the two dispute settlement systems implies that the expected host country welfare levels are the same, $\tilde{V}(\Theta^*(L)) = \tilde{V}(\Theta^*(0))$. And it implies together with the absence of litigation in equilibrium that the source country expected welfare is the same, $\tilde{Y}(\Theta^*(L), L) = \tilde{Y}(\Theta^*(L), 0) = \tilde{Y}(\Theta^*(0), 0)$.∎

Proposition 3 identifies the situation where there should be no controversy over the dispute settlement system: there are no equilibrium compensation payments under ISDS, and switching to SSDS would not cause any enforcement problems given investment protection $\Theta^*(0)$. But when either of these conditions is violated, matters are different. We begin by considering the case where enforcement is not a problem, but the negotiated agreement yields expected compensation payments under ISDS.

5.3.2 The source country's bargaining position is affected

Assume that the negotiated level of investment protection is sufficiently high under ISDS, for instance because of source country bargaining power, that it yields compensation payments and therefore litigation in equilibrium, $\Theta^*(0) > \theta^E$. Assume also that litigation is relatively inexpensive

under SSDS, in the sense that it would be worthwhile for the source country to litigate if the host country regulates for $\theta < \Theta^*(L)$, that is, that $\bar{\Theta}(L) < \Theta^*(L)$.

How the negotiated level of protection depends on L

The impact of the litigation cost L on the level of negotiated investment protection will be central for what follows. In the present case, it is given by the FOC (10) when fulfilled with equality. Total differentiation for an arbitrary L yields

$$\Theta_L^*(L) = \frac{1}{\hat{B}_{\theta\theta}}[(B_{YV}\tilde{V}_\theta + B_{YY}\tilde{Y}_\theta)\tilde{P} + B_Y\tilde{P}_\theta] \tag{14}$$

Since $\hat{B}_{\theta\theta}$ is negative by the assumed SOC for the bargaining problem, the sign of the right-hand side is the opposite of the sign of the bracketed term. By assumption, $B_{VY} \geq 0$, $B_{YY} \leq 0$, $B_Y > 0$, and $\tilde{P} > 0$. The sign of $\Theta_L^*(L)$ thus in general partly depends on the sign of

$$\tilde{Y}_\theta(\Theta^*(L), L) = \tilde{\Pi}_\theta - L\tilde{P}_\theta \gtrless 0 \tag{15}$$

Hence, with SSDS, the level of protection affects source country expected welfare partly by affecting the expected investor profits, and partly by affecting the expected enforcement costs. The latter effect is of inherently ambiguous direction, regardless of $L \geq 0$:

$$\tilde{P}_\theta(\theta') = f(\theta') - f(\Theta^C(K(\theta')))\Theta_k^C(K(\theta'))K_\theta \gtrless 0. \tag{16}$$

A higher level of investment protection will affect both the upper and the lower bound of the range of θ for which there is litigation. It will increase the upper limit, but it will have unclear effects for the lower bound $\Theta^C(K(\theta'))$, due to the ambiguous sign of Θ_k^C.[28]

We can thus distinguish between two cases with regard to how the level of protection is affected by the magnitude of L. First, assume that $\tilde{P}_\theta(\Theta^*(L)) < 0$, as could be the case if $\Theta_k^C > 0$. This ensures that $\tilde{Y}_\theta(\Theta^*(L), L) > 0$, which in turn implies by the FOC for the bargaining problem that $\tilde{V}_\theta(\Theta^*(L)) < 0$. The negotiated level of investment protection then increases in L: $\Theta_L^*(L) > 0$. Second, if $\tilde{P}_\theta(\Theta^*(L)) > 0$, it is ambiguous whether $\Theta_L^* \gtrless 0$.

Implications for how a switch to SSDS affects expected welfare levels

To capture the implications of a shift from ISDS to SSDS, let us consider first the case where the litigation cost L with SSDS is marginally larger than 0. We can then analyze the switch from ISDS to SSDS as a marginal increase in L at $L = 0$. We have by Proposition 2 that $\tilde{V}_\theta(\Theta^*(0)) < 0$, so the host country will benefit from switch to SSDS if and only if the level of investment protection falls, that is, if and only if $\Theta_L^*(0) < 0$. The implications for the source country are unambiguously

---

[28]The stability condition (4) ensures that θ' grows faster than $\Theta^C(K(\theta'))$ as we increase θ', but we here also have to take account of the marginal frequencies.

negative:

$$\frac{d\tilde{Y}(\Theta^*(0),0)}{dL} = -\tilde{P}(\Theta^*(0)) + \tilde{Y}_\theta(\Theta^*(0))\Theta^*_L(0) \tag{17}$$

An increase in L will have a direct adverse effect by making the enforcement more costly, as captured by the first term. And the resulting fall in the level of protection will also affect the source country adversely, since $\tilde{Y}_\theta(\Theta^*(0),0) > 0$ by Proposition 2. Hence:

Proposition 4 *Assume that the negotiated investment protection under ISDS yields compensation payments with positive probability, $\Theta^*(0) > \theta^E$, and that $\Theta^*(0) > \bar{\Theta}(0)$. Assume that the litigation cost L with SSDS is marginally larger than 0. A switch from ISDS to SSDS will then:*
(i) Increase the negotiated level of protection if this reduces the probability of litigation, that is, if $\tilde{P}_\theta < 0$. The switch will reduce host country expected welfare, and will ambiguous impact on source country expected welfare.
(ii) Have ambiguous impact on the negotiated level of investment protection if $\tilde{P}_\theta > 0$. If it increases the level of protection, the host country loses, and the impact on the source country is ambiguous. If it reduces the level of protection, the host country benefits, and the source country loses.

We have thus far considered a switching to a SSDS system with a very low litigation cost L. What if instead increase $L > 0$ marginally? The effects would qualitative largely be the same, as long as is the enforcement constraint is respected, and there is a strictly positive probability for litigation at the outset ($\tilde{P}(\Theta^*(L)) > 0$). Formally, the only difference is that for L sufficiently large, the source country will lose from an increase in the level of investment protection $\tilde{Y}_\theta < 0$ if $\tilde{P}_\theta > 0$, as can be seen from (15). The reason would then be that an increase in the level of protection would increase litigation costs more than it would increase investors' expected profits. However, such a setting would require $\tilde{V}_\theta > 0$.by the FOC for the bargaining problem. So the source country would then prefer less investment protection, and the host country more protection, than what is negotiated. Why formally feasible, we find this setting less plausible empirically, and therefore disregard it. Hence, the findings in Proposition 4 can be generalized to switches to SSDS for a range of $L > 0$.

Figure 4 illustrates a case where the host country would indeed benefit from to SSDS relative to ISDS, due to a lower level of protection. The dashed curves reproduce the ISDS setting from Figure 3. In the lower graph, the solid curve shows the expected welfare of the host country when the enforcement constraint is low enough, $\Theta^*(L) > \bar{\Theta}(L)$ not to constrain the bargaining problem. The upper graph illustrates how the expected source country welfare is lower with SSDS than with ISDS due to the litigation costs; this is illustrated by the thinner curve, which has a kink at θ^E. Formally, this comes from the fact that $\tilde{Y}_L < 0$ in this case. But since the enforcement constraint implies that there will be no enforcement unless $\theta' \geq \bar{\Theta}(L)$, the relevant part with SSDS is the thicker drawn curve.

(Fig 4: Possible outcomes of negotiations with SSDS).

5.3.3 The enforcement constraint binds

We next consider the case where the enforcement constraint is binding under SSDS, in the sense that the level of protection with ISDS $\Theta^*(0)$ will not be enforced by the host country: $\bar{\Theta}(L) \in [\Theta^*(0), \theta_{\max})$. This is equivalent to the political costs of litigation being sufficiently high under SSDS since

$$\bar{\Theta}_L(L) = \frac{1}{\Pi_k(\bar{K}(L))K_\theta(\bar{\Theta}(L))} > 0.$$

To remove the considerations examined in the previous subsection, assume also that there are no compensation payments in equilibrium with ISDS, $\Theta^*(0) \leq \theta^E$.

Since investment protection under SSDS must not fall below the threshold $\bar{\Theta}(L)$, $\Theta^*(0) \leq \bar{\Theta}(L) \leq \Theta^*(L)$. If either of these inequalities is strict, the host country will strictly prefer ISDS to SSDS: $\tilde{V}(\Theta^*(L)) < \tilde{V}(\Theta^*(0))$ by Proposition 2. If they are both equalities, the host country is indifferent between ISDS and SSDS. Hence, the host country at least weakly prefers ISDS to SSDS because of the (weakly) lower level of investment protection under ISDS when litigation under SSDS is costly.

For the source country, a switch to SSDS would have ambiguous consequences:

$$\tilde{Y}(\Theta^*(L), L) - \tilde{Y}(\Theta^*(0), 0) = \tilde{\Pi}(\Theta^*(L)) - \tilde{\Pi}(\Theta^*(0)) - \tilde{P}(\Theta^*(L))L \gtrless 0$$

On the one hand, SSDS yields (weakly) better investment protection, which increases the expected industry profit; this is captured by the first two terms above. On the other hand, there are expected litigation costs under SSDS that may or may not dominate the benefit of improved investment protection. We summarize these results as:

Proposition 5 *Assume that political litigation costs with SSDS are sufficiently high that the outcome with ISDS would not be enforced by the source country: $\bar{\Theta}(L) \in [\Theta^*(0), \theta_{\max})$. Negotiated investment protection then is weakly higher under SSDS than ISDS: $\Theta^*(L) \geq \Theta^*(0)$ with strict inequality if $\bar{\Theta}(L) > \Theta^*(0)$. The host country expected welfare is weakly lower under SSDS than ISDS, whereas the effect of the dispute settlement system on the source country is ambiguous.*

Note that in this case the impact of the switch to SSDS does not come from making it costly to enforce equilibrium compensation payments; we have assumed away such payments. Instead, the switch to SSDS will make the source country unwilling to enforce compensation payments should the host country regulate for $\theta < \Theta^*(0) \leq \theta^E$. This would be opportunistically exploited by the host country, which would then regulate for all $\theta > \Theta(K(\Theta^*(0)))$. The level of investment protection will thus have to be higher with SSDS in order to give the source country sufficient incentive to enforce the agreement.

24

5.3.4 Investment protection is no longer enforced

The final possibility is where litigation costs are so large that $\bar{\Theta}(L) \geq \theta_{\max}$. Then there is no level of investment protection high enough to induce the source country to enforce litigation under SSDS and low enough to ensure that the host country benefits from an agreement featuring SSDS. If this is the case, a switch to SSDS would effectively make the agreement unravel. It is then only possible to implement an ISDS agreement:

Proposition 6 *If litigation costs under SSDS are sufficiently high ($\bar{\Theta}(L) \geq \theta_{\max}$), then the only feasible agreement is one that features ISDS.*

A switch to SSDS would in the above case obviously harm both countries; this follows trivially from the assumption (8) that there is scope for a negotiated agreement with ISDS. However, there are intuitively many plausible reasons why a host country in practice might benefit from the unravelling of protection commitments in certain sectors. Let us therefore briefly take a step outside the above framework to capture such a notion.

We assumed implicitly above that the agreement was designed to address distortions in the sector it applied to. But actual agreements normally cover whole economies, so for our formalization to depict such agreements, we need to assume that the agreement specifies industry-specific commitments. However, even though actual agreements have certain flexibilities that makes it possible to interpret their provisions in light of the specific circumstances at hand in each industry, they do not appear to be fully flexible in this regard. They appear to be at least partly designed with "typical" industry features in mind.[29]

To illustrate in a very simple fashion how this might change the implications of switching from ISDS to SSDS, assume that our source country can invest in two economically unrelated industries A and B in the host country, two representing a large number. But the investment agreement specifies a single level of protection θ' that applies to both industries. Also for the sake of simplicity, assume that the source country is the dominant party in the negotiations, and can thus set θ' unilaterally; the same issue can arise with bargaining. Consider the case with ISDS. The source country will then propose an agreement θ^M that gives the host country just as high expected welfare as the host country would get in the no-agreement situation:

$$\tilde{V}^A(\theta^M) + \tilde{V}^B(\theta^M) \equiv \tilde{v}^{AN} + \tilde{v}^{BN}$$

If the two industries differ in the welfare they generate for the host country for any common level investment protection, the host country must lose from the inclusion in the agreement of one

---

[29]There are many reasons why actual agreements might impose obligations that the host country would prefer to escape. For instance, host countries can disagree with how arbitration panels interpret their vaguely drafted agreements; host countries might have had erroneous expectations regarding the consequences of their agreements for investment, or the benefits from the investments; or exogenous changes might have occurred that the agreements, such as increasing levels of development.

of the sectors. For instance, suppose that the host country would always regulate industry B absent an agreement due to adverse externalities. There would then not be any investment absent an agreement, and the host country expected welfare would be unaffected by this industry. But with the agreement the expected welfare is negative since there will be instances where the host country will have to compensate the investor for regulating an investment it did want to have in the first place, or abstain from regulating it: $\tilde{V}^B(\theta^M) < \tilde{v}^{BN}$.

Suppose that there is now a switch to SSDS. If the political litigation costs are sufficiently high in industry B that the source country will not enforce the commitments in this sector, the agreement would effectively unravel there. This would by itself *benefit* the host country, while it would harm the source country. But while we have here identified a mechanism through which a host country might benefit from an unraveling of an agreement in specific instances, we have not established a plausible argument for why a host country should advocate a shift to SSDS for an agreement as a whole.

First, the switch to SSDS would presumably induce a renegotiating of the level of investment protection. In the example here, it would induce the source country to change the level of investment protection to be designed specifically for industry A. It will now only have to ensure that the host country expected welfare does not fall short of $v = \tilde{v}^{AN}$. Whether the end effect would be to make the host country better off will depend on whether the source country will find it optimal to push host country welfare to this level, in light of the litigation costs in this industry, and it will depend on whether $\tilde{v}^{AN} \gtrless \tilde{v}^{BN}$.

Second, there is nothing to suggest that the unravelling of the agreement will occur only in industries where the host country would prefer no agreement. It might also occur in industries where the host country would be better off with the existing agreement than with the no-agreement outcome. Hence, as long as the same dispute settlement system is to be used across all industries, and the levels of investment protection are unchanged, for the switch to SSDS to be beneficial to the host country, the effective unravelling of protection commitments must be desirable "on average", across industries.

5.3.5 Conclusions regarding the incentives to shift to SSDS

Our main concern here has been to examine to notion that host countries would benefit from removing the possibility for foreign investors to litigate. The distributional effects of switching from ISDS to SSDS are ambiguous for both countries despite our simple formalization of political litigation costs, including the assumption that they only accrue to the source country. But our framework does identify certain circumstances under which a switch from ISDS to SSDS would benefit the host country for a range of L by reducing the negotiated level of protection. But there is no presumption that the level of investment protection will fall. The analysis hence gives at best weak support for the notion that excess litigation under ISDS could be a rationale for host country

discontent with ISDS, and that host countries should therefore seek to remove ISDS from their agreements.

6   Negotiating dispute settlement and investment protection

In the previous section we treated the choice of dispute settlement system as exogenous, in order to see how ISDS and SSDS might have different distributional consequences on the host and source country. We showed how these implications depend on the negotiated level of investment protection under the two different regimes, and on the magnitude of political litigation costs under SSDS. We now consider simultaneous negotiation regarding investment protection and dispute settlement. As pointed out above, most actual agreements allow for both ISDS and SSDS, and in practice only ISDS is used. Can this be understood within our framework? In particular, if we here assume a setting in which the host country prefers SSDS as per the previous section, will not the outcome of negotiations over both the level of investment protection and the form of dispute settlement lead to SSDS, if the host country has sufficient bargaining power?

The bargaining outcome is now given by the solution to

$$\max_{(\theta',L')\in[\underline{\theta},\bar{\theta}]\times\{0,L\}} B(\tilde{V}(\theta'),\tilde{Y}(\theta',L'))$$

subject to the two countries' individual participation constraints, $\tilde{V}(\theta') \geq \tilde{v}^N$ and $\tilde{Y}(\theta',L') \geq \tilde{\pi}^N$, and the enforcement constraint $\theta' \geq \bar{\Theta}(L)$. We also assume that $\bar{\Theta}(L) < \theta_{\max}$. Otherwise an agreement featuring SSDS would not be feasible.

Consider any proposed agreement that features SSDS and the negotiated level of investment protection $\Theta^*(L)$. If the parties instead switched to ISDS, and maintained the same level of protection, equilibrium investment and regulation would remain unchanged. ISDS must therefore achieve as high a level of B as SSDS. However, if there is litigation under SSDS, the source country will face litigation costs $\tilde{P}(\Theta^*(L))L$, which will result in a lower expected welfare level, and thus a strictly lower value of B, for any $\Theta^*(L)$. Hence, ISDS weakly maximizes B. It thus follows directly from the assumed Pareto efficiency of the bargaining solution that the negotiated dispute settlement system will be ISDS. This will have distributional consequences, since with ISDS the negotiated level of protection will be $\Theta^*(0)$, this being what maximizes B for ISDS.

Hence:

Proposition 7 *The agreement will feature ISDS and the equilibrium investment protection be equal to $\Theta^*(0)$ when the host and the source country negotiate both the form of dispute settlement and the level of investment protection. If $\Theta_L^*(0) < 0$ there exists a range of $L > 0$ for which the host country would have been better off with SSDS exogenously imposed.*

Intuitively, the fact that the negotiation is conducted over two instruments—the level of invest-

ment protection θ' and the dispute settlement mechanism—opens up a "trading opportunity" for the two countries. For a given level of investment protection, the source country prefers ISDS to SSDS to avoid being exposed to political litigation costs, and for a given dispute settlement mechanism, the host country prefers less protection than the source country under ISDS. Consequently, both parties prefer an agreement with ISDS rather than SSDS and negotiate the investment protection to split the gains from the agreement, or alternatively they are indifferent. This is consistent with the fact that almost all investment agreements include ISDS.

The host country would sometimes prefer to negotiate the level of investment protection with SSDS rather than ISDS, but is unable to enforce this according to Proposition 7. Why could the host country not do this if it had enough bargaining power? To illustrate the role of host country bargaining strength, assume that $B(\tilde{V}, \tilde{Y})$ is the maximand to a Nash Bargaining problem with the status quo utilities $\tilde{v}^N$ and $\tilde{\pi}^N$ normalized to zero:

$$B(\tilde{V}(\theta'), \tilde{Y}(\theta', L)) \equiv \tilde{V}(\theta')^\alpha \tilde{Y}(\theta', L)^{1-\alpha}, \tag{18}$$

where the parameter α can be interpreted to reflect the bargaining power of the parties. The FOC with regard to the level of investment protection can be written as

$$\alpha \frac{\tilde{V}_\theta}{\tilde{V}} + (1 - \alpha) \frac{\tilde{Y}_\theta}{\tilde{Y}} = 0$$

Hence, as all bargaining power is shifted toward the host country—as α goes to unity—the FOC for the negotiated level of investment protection converges to $\tilde{V}_\theta = 0$, that is, the negotiated outcome converges to the host country's most preferred level of protection $\hat{\theta}$. But along this path, the maximand B is still maximized for $L = 0$ as long as any weight is attached to source country welfare. Hence, the relative bargaining power of the host country is reflected in the magnitude of $\Theta^*(0)$, but not in the choice of dispute settlement system.

Finally, Ossa, Staiger and Sykes (2019) investigate the choice of dispute settlement system in trade and investment agreements, under the assumption of incomplete enforcement. In contrast to the present setting, their investment provision applies to direct expropriation, and the agreement is designed to maximize the expected joint surplus of the host and source country. Ossa et al. (2019) focus on the role of imperfect information in arbitration, assuming that an arbitration court receives a noisy signal regarding whether a host country intervention was efficient. The host country's decision is assumed to be overturned if regulation is found to be jointly inefficient, but sustained otherwise; this would in our setting correspond to the special case where the agreement stipulates compensation if and only if $\theta \leq \Theta^C(k)$. In their model a foreign firm has a stronger incentive to request arbitration than the foreign government by an assumption that the perceived benefit of winning is larger for the firm than the government.[30] ISDS has two counteracting consequences for

---

[30] An assumption that the host country government has an incremental fixed arbitration cost under SSDS is formally

efficiency relative to SSDS: the firm's stronger incentive to request arbitration will dampen the host country's desire to expropriate the firm. But there is also an increased likelihood of non-intervention under ISDS because of the probability of an incorrect court ruling in the complainant's favor. The net effect on efficiency is ambiguous in general, but the authors identify a set of conditions under which the first positive effect of ISDS dominates the second negative effect. Ossa et al. (2019) thus for the most part address other issues than those studied here. But their findings complements those made here, for instance by showing that the efficiency of ISDS does not critically hinge on complete enforcement.

7   Alternative formulations of the ISDS/SSDS distinction

We have formalized the standard explanation for the difference between SSDS and ISDS—political and/or diplomatic enforcement costs with SSDS—in an analytically very simple fashion. But simple as it is, it does appear to capture essential aspects of the issue. But there are of course alternative ways of understanding and modeling the difference between the two forms of dispute settlement systems.

First, a common claim in favor of ISDS is that it provides for speedier resolution of disputes, and thereby saves costs for investors. Another explanation for why SSDS differs from ISDS might be that source country governments put less weight on the profits of their investors, compared to the investors themselves. To illustrate, assume that both investors and the source country government face a standard (non-political) litigation costs G for legal counseling. Assuming that no mistakes are made in arbitration, the investor will litigate in case of compensable regulation if $\Pi(k) \geq G$. With SSDS, the source country government faces the same process costs, but also the political costs L. The source country government potentially places a smaller weight $\gamma \leq 1$ on the profits of the investor relative to the administrative costs, than do investors. Hence, the government will litigate if and only if $\gamma\Pi(k) \geq G + L$. An alternative interpretation is that γ reflects discounting of future compensation payments, if the arbitration process is slower under SSDS. Since $G < (G + L)/\gamma$, these alternative explanations for the difference between ISDS and SSDS share the basic feature with the framework employed in the analysis above, in that the there will be a range of disputes for which the source country will refrain from litigating, while private investors will litigate.[31]

The outcomes under these different approaches might be qualitatively different, however. The political enforcement costs L will only affect the outcome in situations where there is litigation, and the same is true for the discounting of the profit. But if the source country government places less weight on the profit of an investor when deciding whether to litigate on behalf of the investor,

---

equivalent to an assumption that investors and the government face the same arbitration cost, but the government places a lower weight on the foregone operating profits than do investors. Some alternative formulations are briefly discussed in the ensuing section.

[31] Another potential difference between the dispute settlement systems is that state-state disputes might go to bodies such as the International Court of Justice, while investor-state disputes are arbitrated by ad hoc panels.

the government will presumably also place less weight on the investor's profit when negotiating the agreement in the first place. This implies, for instance, that industries of less value to the source country government will have lower negotiated investment protection.[32]

Second, we assumed above that the political/diplomatic enforcement is unrelated to the regulatory problem facing the host country, that is, L and θ are uncorrelated. By implication, the decision to enforce the agreement is unrelated to the severity of regulatory problem. Consequently, SSDS might cause the agreement to collapse in industries where production would be desirable to the host country (and thus be optimal from a joint point of view). But it seems plausible that the political enforcement costs and the severity of the regulatory problem can be positively correlated: it should be particularly costly politically for source countries to pursue cases in instances where the host country regulations are addressing severe regulatory problem. For instance, it would likely cause a government a substantial loss of political goodwill if it were to litigate on behalf of a tobacco producer regarding non-discriminatory host country health measures against smoking.

Finally, we assumed that the source country carries enforcement costs only when litigating in order to obtain compensation. An alternative would be to assume that enforcement efforts are required more generally to induce the host country to abide by the agreement whenever it would unilaterally prefer to regulate, that is, for all $\theta > \Theta(k)$. The source country expected welfare would then be

$$\tilde{Y}(\theta', L) \equiv \tilde{\Pi}(\theta') - [F(\theta') - F(\Theta(K(\theta')))]L$$

whenever the enforcement constraint is fulfilled. However, the qualitative properties of the model would be much the same as above, except the large magnitude of enforcement costs would probably reduce investment protection $\Theta^*(L)$ under ISDS.

8   Concluding remarks

There have been calls in the policy debate for changes to the dispute settlement system in investment agreements. A common suggestion has been to restrict the possibilities for private investors to litigate against host countries. Some countries have also moved in this direction with regard to their agreements. The vast majority of investment agreements still allow for ISDS, however. This paper is to the best of our knowledge the first economic analysis of distributional implications of ISDS versus SSDS in bilateral investment treaties.

The paper focuses on the standard explanation for why SSDS yields less litigation than ISDS— the political/diplomatic costs that source country governments face when enforcing agreements on behalf of their investors. We have used this framework to identify circumstances under which a host country would benefit from an exogenous switch from ISDS to SSDS, by reducing the level of

---

[32] There is thus an distinction between the case where a home government does not put as much weight on investor profits or compensation payments, as captured by $\gamma < 1$, and where it does not want to hold the axe when enforcing the payments, or at least not be seen to hold the axe, as captured by the cost L.

investment protection somewhat. A reduction in the level of protection will allow the host country to regulate without paying compensation for a larger set of regulatory shocks, and it might reduced expected compensation payments, along the lines proposed in the debate. But our framework offers no compelling reason to believe that a switch from ISDS to SSDS will indeed reduce the level of investment protection. More generally, an exogenous switch to SSDS might benefit one of the parties, or it might reduce expected welfare of both parties. But it cannot benefit both parties.

The inefficiency of SSDS is even more clear in the setting where the parties negotiate both the level of protection, and the form of dispute settlement system. The Pareto efficient outcome is ISDS, so it will be the outcome for a broad range of bargaining formats. But we have also seen that this feature can have strong distributional implications, since it is not necessarily to the benefit of both parties.

Our analysis has been based on a very simple formalization of the difference between ISDS and SSDS. But it does seem to fairly well capture the standard explanation for the introduction of ISDS in investment agreements. Also, several other empirically plausible formalizations seem to have similar properties to the one employed here. To make sense of the critique against ISDS we thus have to bring in additional elements into the analysis, the fact that investors are more prone to litigate than their governments does not necessarily make SSDS desirable even for host countries. For instance, criticism has been directed in the debate in the debate against the possibility for foreign investors to request arbitration outside host countries' legal systems.

References

Aisbett, Emma, Larry Karp and Carol McAusland (2010). "Police Powers, Regulatory Takings and the Efficient Compensation of Domestic and Foreign Investors." *The Economic Record* 86(274), September, 367–383.

Alschner, Wolfgang and Dimitriy Skougarevskiy (2016). "Convergence and Divergence in the Investment Treaty Universe – Scoping the Potential for Multilateral consolidation." *Trade, Law and Development* 8(2), 331-362.

Bernasconi-Osterwalder, Nathalie (2014). "State-State Dispute Settlement in Investment Treaties". International Institute for Sustainable Development Best Practice Series. October.

Dolzer, Rudolf and Christoph Schreuer (2012). *Principles of International Investment Law.* Oxford, UK: Oxford University Press.

Gertz, Geoffrey (2017). "Renegotiating NAFTA: Options for Investment Protection". *Globalviews* 7, March. Brookings Institution.

Howse, Robert L. (2017). "International Investment Law and Arbitration: A Conceptual Framework." In Helene Ruiz-Fabri (ed.), *International Law and Litigation*, New York: Nomos Press.

Horn, Henrik and Thomas Tangerås (2019). "Economics and Politics of International Investment Agreements". Working Paper No. 1140, Research Institute for Industrial Economics (IFN), Stockholm.

Janeba, Eckhard (2019). "Regulatory Chill and the Effects of Investor State Dispute Settlement." *Review of International Economics* 27(4), 1172-1198.

Johnson, Lise, Lisa Sachs and Jeffrey Sachs (2015). "Investor-State Dispute settlement, Public Interest and US Domestic Law". Columbia Center for Sustainable Development Policy Paper, May.

Kohler, Wilhelm and Frank Stähler (2019). "The Economics of Investor Protection: ISDS versus National Treatment." Mimeo, April 1.

Konrad, Kai A. (2017). "Large Investors and Permissive Regulation: Why Environmentalists May Dislike Investor-State Dispute Settlement." *European Economic Review 98,* 341-353.

Menon, Trishna (2018). "Developing Country Opposition to an Investment Court: Could State-State Dispute Settlement be an Alternative?". Kluwer Arbitration Blog, February 17.

Ossa, Ralph, Robert W. Staiger and Alan O. Sykes (2019). "Disputes in International Investment and Trade." Preliminary draft, March.

Roberts, Anthea (2014). "State-to-State Investment Treaty Arbitration: A Hybrid Theory of Interdependent Rights and Shared Interpretative Authority". *Harvard International Law Journal 55(1),* 1-70.

Schjelderup, Guttorm and Frank Stähler (2017). "Investor State Dispute Settlement and Multinational Firm Behavior." Mimeo, July 10.

Stiglitz, Joseph E. (2008). "Regulating Multinational Corporations: Towards Principles of Cross-Border Legal Frameworks in a Globalized World Balancing Rights with Responsibilities." 2007 Grotius Lecture. *American University International Law Review* 23, 451-558.

Stähler, Frank (2018). "An Optimal Investor State Dispute Settlement Mechanism." Mimeo, March 1.

Salacuse, Jeswald W. (2007). "Is There a Better Way? Alternative Methods of Treaty-Based Investor-State Dispute Resolution". *Fordham International Law Journal 31(1)*, 138-185.

Sykes, Alan O. (2005). "Public vs. Private Enforcement of International Law: Of Standing and Remedy". *Journal of Legal Studies 34,* 631.

Trevino, Clovis J. (2013). "State-to-State Investment Treaty Arbitration and the Interplay with Investor-State Arbitration Under the Same Treaty". *Journal of International Dispute Settlement 5,* 199-233.

Vandevelde, J. Kenneth (2005). "A Brief History of International Investment Agreements". *U.C. Davies Journal of International Law & Policy 12(1),* 157-194.
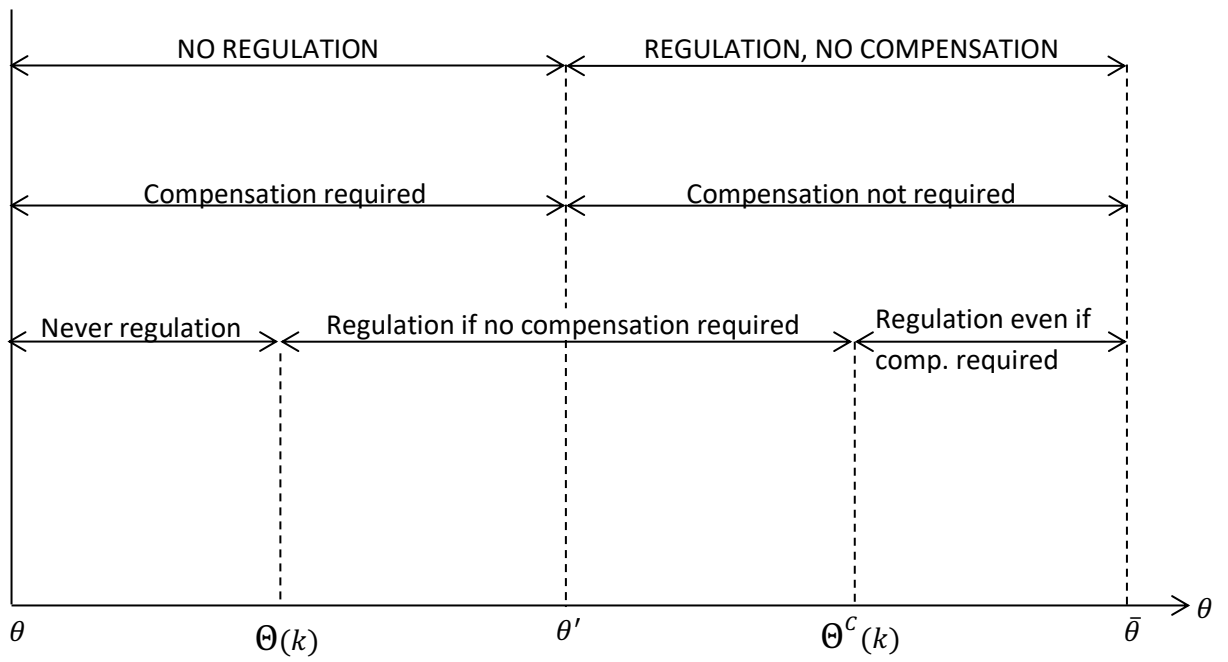
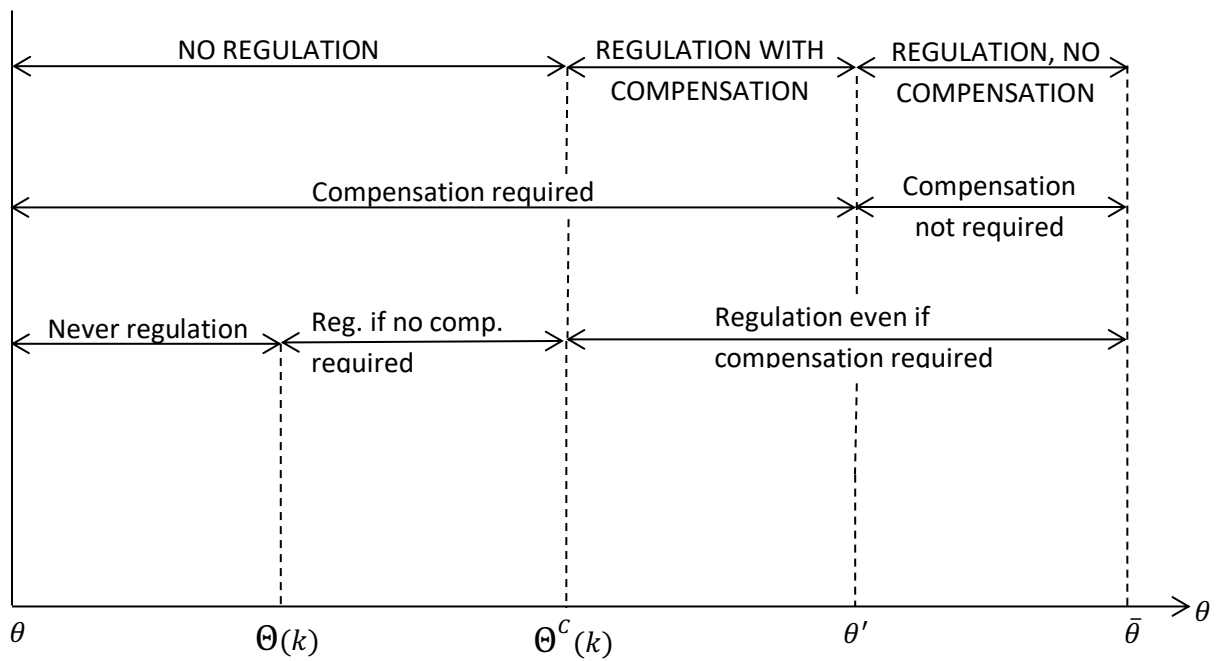**Figure 1: Regulation incentives with $\theta' < \Theta^C(k)$**

Figure 2: Regulation incentives with $\theta' > \Theta^C(k)$
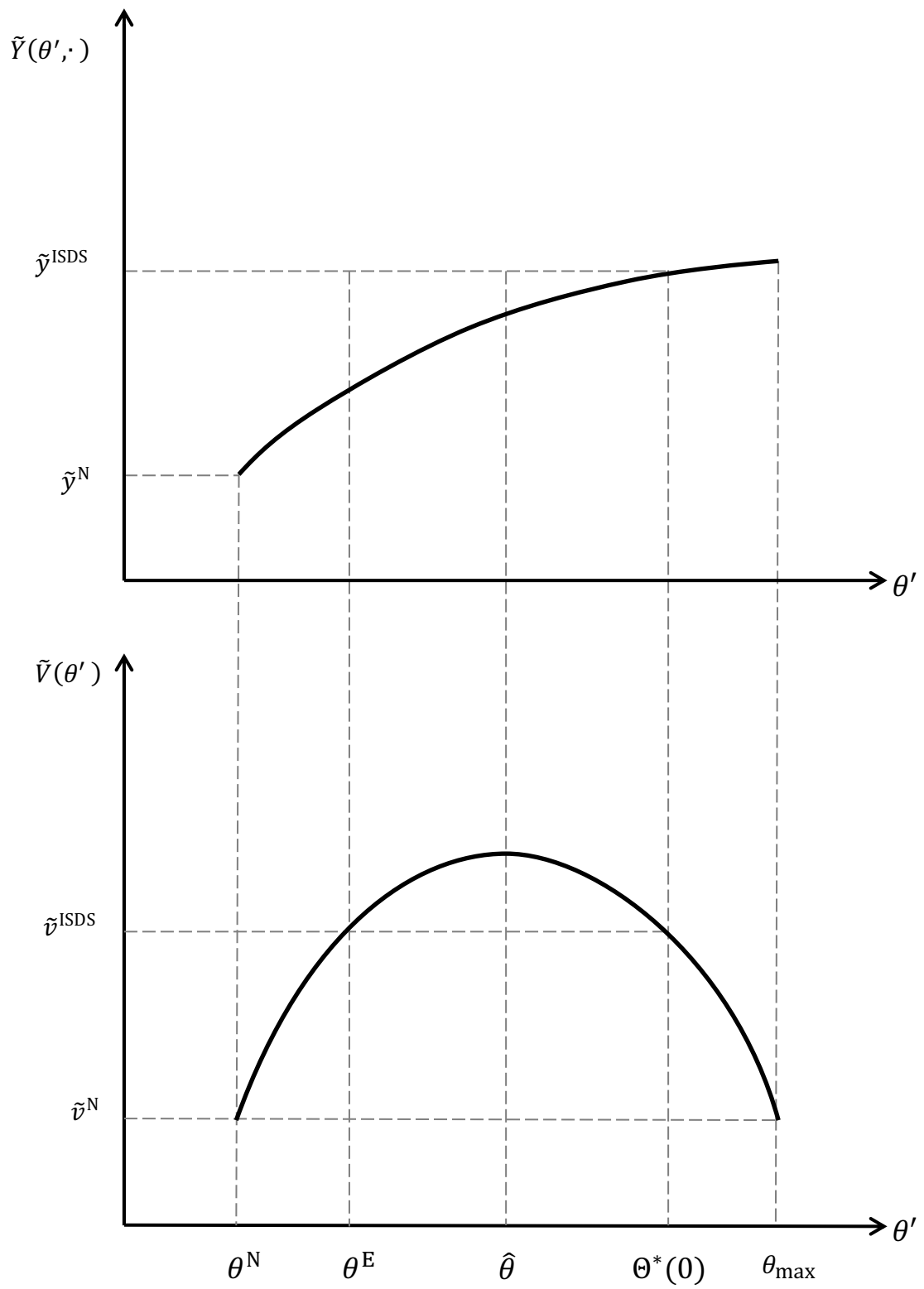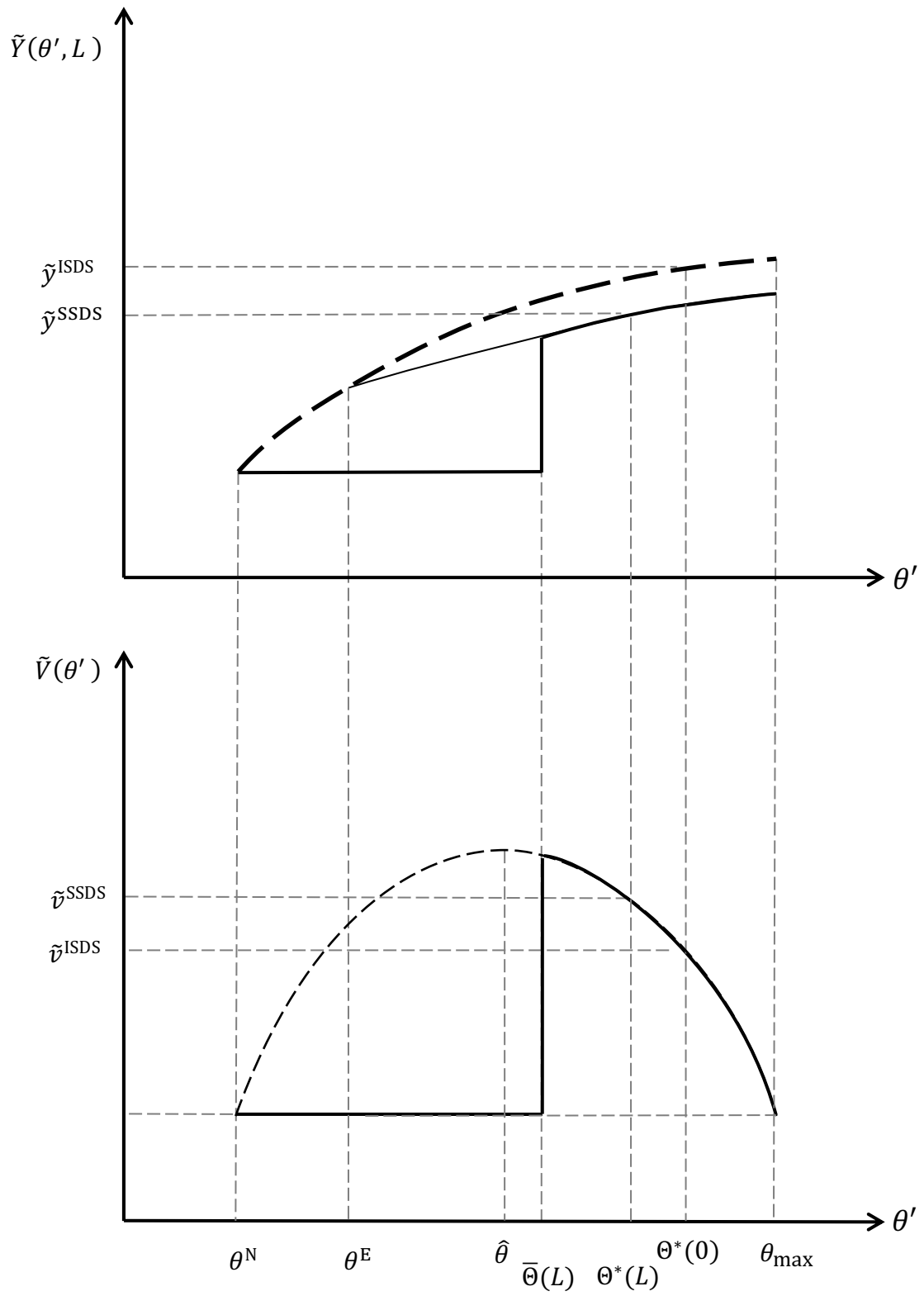
Figure 3: The outcome of negotiations with ISDS

Figure 4: Possible outcomes of negotiations with SSDS