

THE ECONOMIC STRUCTURE OF INTERNATIONAL INVESTMENT AGREEMENTS WITH IMPLICATIONS FOR TREATY INTERPRETATION AND DESIGN

By Alan O. Sykes*

ABSTRACT

This Article argues that international investment agreements (IIAs) serve a dual economic function—to discipline host country policies that impose international externalities on foreign investors, and to curtail inefficient risks associated with agency costs, risk aversion, asymmetric information, and time inconsistency problems that uneconomically increase the cost of imported capital in host countries. It draws on the economic analysis to explain central features of IIAs and their evolution over time, and to address various controversial issues in international investment litigation.

The economic analysis of international investment agreements (IIAs) is in its early stages, despite the fact that IIAs in force now number in the thousands.¹ Most of the research to date focuses on the possible role of IIAs in policing expropriation and “regulatory takings.”² A prime motivation for this focus is the civil society debate over a possible tension between investor-state dispute settlement (ISDS) under IIAs and national regulatory sovereignty,³

* Professor of Law and Warren Christopher Professor in the Practice of International Law and Diplomacy, Stanford Law School. I have benefited immensely from related collaborative work with Ralph Ossa and Bob Staiger, from the comments of Joe Bankman and Louis Kaplow, from participants in workshops at Harvard, Northwestern, and Stanford Law Schools, from comments at the 2019 annual meeting of the American Law & Economics Association, and from numerous suggestions by *AJIL* reviewers.

¹ The majority of international investment agreements are stand-alone bilateral investment treaties (BITs), of which there are now over 2,300 in force. Other international investment agreements in force, numbering over three hundred, are incorporated within broader economic treaties such as the North American Free Trade Agreement (NAFTA) and its potential successor, the United States-Mexico-Canada Agreement (USMCA). See UNCTAD, International Investment Agreements Navigator, available at <http://investmentpolicyhub.unctad.org/IIA>.

² See, e.g., Henrik Horn & Thomas Tangerås, *Economics and Politics of International Investment Agreements* (IFN Working Paper No. 1140, 2017); JONATHAN BONNITCHA, LUAGE POULSEN & MICHAEL WAIBEL, *THE POLITICAL ECONOMY OF THE INVESTMENT TREATY REGIME*, ch. 5 (2017), and sources cited therein; JONATHAN BONNITCHA, *SUBSTANTIVE PROTECTION UNDER INVESTMENT TREATIES: A LEGAL AND ECONOMIC ANALYSIS* (2014); Emma Aisbett & Jonathan Bonnitcha, *Compensation Under Investment Treaties – As If Host Interests Mattered* (UNSW Law Research Paper No. 18-80, 2018); Emma Aisbett, Larry Karp & Karol McAusland, *Police Powers, Regulatory Takings and the Efficient Compensation of Domestic and Foreign Investors*, 86 *ECON. RECORD* 367 (2010); Emma Aisbett, Larry Karp & Karol McAusland, *Compensation for Indirect Expropriation in International Investment Agreements: Implications of National Treatment and Rights to Invest*, 1 *J. GLOBALIZATION & DEV.* Article 6 (2010); Eckhard Janeba, *Regulatory Chill and the Effects of Investor-State Dispute Settlements* (mimeo 2016).

³ See, e.g., Elizabeth Warren, *The Trans-Pacific Partnership Clause Everyone Should Oppose*, *WASH. POST* (Feb. 25, 2015), at https://www.washingtonpost.com/opinions/kill-the-dispute-settlement-language-in-the-trans-pacific-partnership/2015/02/25/ec7705a2-bd1e-11e4-b274-e5209a3bc9a9_story.html?utm_term=.c96b0fdd01ac; Public Citizen, *Investor-State Dispute Settlement: Extraordinary Corporate Power in “Trade” Deals*, at <https://www.citizen.org/our-work/globalization-and-trade/investor-state-system>.

and the associated fear that ISDS litigation may produce undesirable “regulatory chill.”⁴ This analysis has close ties to the broader literature on takings and regulatory transitions.⁵ A few other matters have received attention, such as the question whether developing countries benefit from entering IIAs,⁶ the rationale for private rights of action in IIAs (in contrast to their absence in other areas of international law, such as trade treaties),⁷ and the question whether host countries can escape their treaty obligations due to “necessity” or other unanticipated contingencies.⁸ But many central features and legal issues associated with IIAs have received little consideration from an economic perspective. This Article seeks to fill a number of the lacuna, as well as to deepen the economic understanding of issues that have received prior attention.

The economic explanation for IIAs begins with the premise that counterparties must expect to be better off with IIAs than without them or they would not agree to them. An economic theory of IIAs thus necessarily entails a search for sources of mutual gain—the “efficiencies” from concluding them. A focus on “efficiencies” in no way denies, however, that the attainment of joint gains through IIAs may be accompanied by a greater sense of fairness, that the orchestration of joint gains through treaty-making may eliminate international conflict and promote peaceful international relations, that IIAs may have some favorable impact on domestic rule of law institutions, or that IIAs may aid in promoting economic development.⁹ Indeed, any such effects can themselves be viewed as a source of “efficiency,” at least in a political sense, although they are not the emphasis here. Accordingly, the analysis in this Article should be seen as complementary to these other accounts of IIAs.

Likewise, a focus on efficiencies does not rule out the possibility that parties may have expectations about the benefits of a treaty that prove incorrect after the fact. We return to this issue when discussing some of the literature on whether developing countries benefit from IIAs, whether ISDS is in the best interest of treaty signatories, and indirectly in discussing how treaty provisions have evolved in response to some controversies in arbitration. The claim here is *not* that IIAs flawlessly promote efficiency in every respect, that there is no room for improvement in any of their provisions, or that parties will never renegotiate or withdraw from IIAs. The claim is more modest—that properly crafted and interpreted treaty provisions

⁴ See, e.g., Vicki Been & Joel C. Beauvais, *The Global Fifth Amendment? NAFTA's Investment Protections and the Misguided Quest for an International "Regulatory Takings" Doctrine*, 78 N.Y.U. L. REV. 30 (2003).

⁵ See, e.g., Lawrence Blume & Daniel L. Rubinfeld, *Compensation for Takings: An Economic Analysis*, 72 CAL. L. REV. 569 (1984); Lawrence Blume, Daniel L. Rubinfeld & Perry Shapiro, *The Taking of Land: When Should Compensation Be Paid?*, 99 Q. J. ECON. 71 (1984); Louis Kaplow, *An Economic Analysis of Legal Transitions*, 99 HARV. L. REV. 509 (1986).

⁶ See Andrew Guzman, *Why LDCs Sign Treaties that Hurt Them: Explaining the Popularity of Bilateral Investment Treaties*, 38 VA. J. INT'L L. 639 (1998); Ryan J. Bubbs & Susan Rose-Ackerman, *BITs and Bargains: Strategic Aspects of Bilateral and Multilateral Regulation of Foreign Investment*, 27 INT'L REV. L. & ECON. 291 (2007); LAUGE POULSEN, *BOUNDED RATIONALITY AND ECONOMIC DIPLOMACY: THE POLITICS OF INVESTMENT TREATIES IN DEVELOPING COUNTRIES* (2015). There is also a substantial empirical literature that seeks to identify the effect of IIAs on direct foreign investment flows, which is discussed briefly in Section 1.E.

⁷ Alan O. Sykes, *Public Versus Private Enforcement of International Economic Law: Standing and Remedy*, 34 J. LEG. STUD. 631 (2005) (hereinafter *Public Versus Private Enforcement*).

⁸ Anne van Aaken, *International Investment Law Between Commitment and Flexibility: A Contract Theory Analysis*, 12 J. INT'L ECON. L. 507 (2009); Alan O. Sykes, *Economic "Necessity" in International Law*, 109 AJIL 296 (2015).

⁹ See Sergio Puig & Gregory Shaffer, *Imperfect Alternatives: Institutional Choice and the Reform of Investment Law*, 112 AJIL 361 (2018).

can ameliorate a range of inefficiencies that would arise in their absence, and that central features of existing IIAs have their genesis in this economic logic.

Section I develops the central theoretical points regarding the possible efficiency gains from investment treaties. Section I.A concerns the “international externality problem” associated with what may be viewed as “monopsony power” in the host country, which arises when host countries offer investment opportunities to investors that are more profitable than opportunities available elsewhere. In such cases, host countries have the leverage to extract some of the returns on foreign investments (which can also be termed “surplus” or “profits”) through a range of different policies, including but not limited to constraints on market access for foreign investors. This international externality can lead to inefficiently low levels of inbound investment, and to distortions in other policies affecting established investors whenever host countries reap the benefits of the policy but foreign investors bear some of the costs. These inefficiencies are analogous to those associated with “terms of trade” externalities in international trade, which afford the predominant theoretical account of the economic rationale for trade agreements. As in the trade area, properly crafted IIAs can curtail these externalities and enhance global efficiency.

Section I.B discusses the “risk reduction” function of IIAs. This section assumes an absence of international externalities for the purpose of isolating other possible inefficiencies that do not involve them. Accordingly, the analysis assumes that host countries bear the full costs of their policies toward foreign investors through adjustments in the cost of imported capital. Under these circumstances, inefficiencies may still arise from certain risks to investors that host countries can benefit from eliminating. These inefficient risks are associated with agency costs, investor risk aversion, asymmetric information issues, and “time inconsistency” problems (which concern the exploitation of investors after they incur irreversible capital costs in the host country). IIAs can enhance the ability of host countries to eliminate these inefficient risks, which can reduce the cost of imported capital by an amount that exceeds the costs of the associated constraints on host country action.

Sections I.A and I.B also consider the challenges of enforcing IIAs. The analysis suggests that certain commitments, such as market access commitments, may require state-to-state dispute settlement and cannot be obtained through an IIA without sufficient reciprocity between the counterparties. Other commitments are in the unilateral interest of host countries and may be found in IIAs that do not offer significant reciprocity. These include commitments to curtail inefficient risks that harm the host country and that represent inferior mechanisms for the extraction of surplus from foreign investors. ISDS may play a vital role in enforcing such commitments, and Section I.B includes a discussion of the possible benefits and costs of the ISDS mechanism.

Section I.C combines the international externality and risk reduction accounts and considers how the analysis is modified when an IIA seeks to achieve both kinds of objectives. Section I.D addresses the relative importance of the two types of functions in practice, and the reasons why only some IIAs contain market access commitments. Section I.E concludes with a discussion of some alternative accounts of IIAs in the literature.

Section II draws on the economic analysis to develop the economic logic behind the central substantive obligations common to most IIAs. These include the national treatment and most-favored-nation obligations, the minimum standard/fair and equitable treatment principle, rules concerning “expropriation,” and rules regarding capital transfers. Among other

things, it emphasizes how these obligations can play a constructive role in relation to both the international externality and risk reduction rationales for IIAs.

Section III develops the implications of the economic perspective for a range of prominent controversies that have arisen in ISDS litigation, and for the subsequent evolution of IIA treaty provisions on controversial issues. Topics include disputes over the definition of covered “investors” and “investments,” the question whether arbitral jurisdiction should turn on the investment’s contribution to “development,” the function and interpretation of the most-favored-nation and national treatment commitments, the function of the minimum standard/fair and equitable treatment principle, the definition and consequences of “expropriation,” and the question whether IIAs undesirably constrain regulatory sovereignty.

I. THE ECONOMIC RATIONALE FOR IIAs

It is now a commonplace in economic treatments of international law to observe that treaties are contracts, albeit among states rather than private actors, and that parties to contracts do not enter them (absent coercion) unless they yield expected benefits to all parties.¹⁰ What are the expected benefits from IIAs? This section suggests that there are two sources of benefits: measures to reduce international externalities associated with the exercise of host country monopsony power, and curtailment of investment risks that uneconomically increase the cost of imported capital.

A. Host Country Monopsony and the International Externality Problem

Many areas of international law can be understood as addressing cross border externalities. The canonical framework for economic analysis in this context posits that national policy decisions often affect the welfare of foreign nationals, and that national policymakers frequently ignore or undervalue the welfare of foreigners. Consequently, when nations make policy choices unilaterally that have effects abroad, negative international externalities are inefficiently large and positive international externalities are inefficiently small. The role of international law is to escape these inefficient choices and move nations toward behavior that yields benefits for all concerned.¹¹ Modern theorizing about the function of international trade agreements fits nicely into this framework, and similar insights suggest an important role for IIAs as well.

1. The Modern Theory of Trade Agreements

Policy choices by trading nations affect the returns to economic activity by foreign firms, and lead to inefficiencies when nations act unilaterally that trade agreements can address. This phenomenon has now been studied in a variety of settings including perfect competition, imperfect competition, and supply chains that embody relationship-specific investments.¹² For present purposes it is enough to illustrate with the competitive case.

¹⁰ See, e.g., ERIC A. POSNER & ALAN O. SYKES, *ECONOMIC FOUNDATIONS OF INTERNATIONAL LAW*, ch. 6 (2013).

¹¹ See, e.g., Alan O. Sykes, *When Is International Law Useful?*, 45 N.Y.U. J. INT’L L. & POL. 787 (2013).

¹² An excellent up-to-date survey is Gene M. Grossman, *The Purpose of Trade Agreements*, in *HANDBOOK OF COMMERCIAL POLICY*, VOL. 1A (Kyle Bagwell & Robert W. Staiger eds., 2016).

Under conditions of competition, a tariff or other trade impediment will cause inefficiency in the importing country as more costly domestic production substitutes for less costly imports, and as some consumers are priced out of the market by higher prices, resulting in a loss of consumer surplus. These inefficiencies lie at the heart of the traditional economic case for “free trade.”¹³ Harry Johnson observed decades ago,¹⁴ however, that if the market of an importing nation is large enough to “matter” to foreign exporters—that is, if access to the market affects the prices that foreign exporters receive for their production—then a tariff of appropriate magnitude (the “optimum tariff”) can nevertheless enhance the economic welfare (national income) of the importing nation (subject to an important caveat about possible retaliation). This possibility arises whenever the supply curve of imports slopes upward. Then, a tariff induces foreign suppliers to move down their supply curve and to the left, lowering their price. Part of the tariff revenue thus represents an extraction of foreign producer surplus or profit, the magnitude of which can exceed the other domestic costs of a tariff. This reduction of foreign producer surplus constitutes the international externality associated with import tariffs imposed by a country large enough to “matter” to its foreign suppliers.

Another way to understand the optimum tariff concept is to observe that when access to an importing country’s market “matters” to foreign exporters, the importing country’s consumers *collectively* have a degree of monopsony power in that market. A “monopsonist” in economics possesses “buyer power”—an ability to affect the price of what it purchases. The classic textbook monopsonist exploits its buyer power by curtailing purchases to reduce the price charged by its suppliers. The price savings on the units that it purchases can more than compensate the monopsonist for the lost benefits of foregone purchases. An optimum tariff on imports in effect exploits the collective monopsony power of consumers in the importing nation¹⁵—it leads then to reduce their purchases, in turn reducing the price that the importing nation pays to foreign suppliers.¹⁶ By contrast, if the importing country lacks monopsony power, foreign suppliers will not reduce their prices and any tariff will be passed through in full to domestic consumers. The “optimum tariff” under those conditions is zero for a country that wishes to maximize its national income.

Modern theorizing has generalized this insight to settings in which governments use trade policy to achieve distributional as well as efficiency objectives and use policy instruments other than tariffs. The general lesson is that a variety of commercial policy choices, regulatory

¹³ See, e.g., PAUL R. KRUGMAN, MAURICE OBSTFELD & MARC J. MELITZ, *INTERNATIONAL ECONOMICS: THEORY AND POLICY* 198–202 (11th ed. 2018).

¹⁴ Harry G. Johnson, *Optimum Tariffs and Retaliation*, 21 *REV. ECON. STUDS.* 142 (1953–54).

¹⁵ For clarity, however, this theory of trade agreements does not require that importing nations *deliberately* exploit national monopsony power, or even that their trade policy officials understand the concept. Johnson’s original model of the optimum tariff is written from the perspective of an optimizing government pursuing its national welfare, and the modern economic literature often refers to importing governments “manipulating” their terms of trade (the price of imports in relation to exports). But it is enough that when governments make decisions without regard to their effects on the welfare of foreigners and a harmful externality arises for foreign firms, there will be a tendency for importing governments to engage excessively in whatever behavior produces the externality. Put differently, when an important interest group is missing from a political process, the resulting political equilibrium will tend to burden that group for the benefit of others to an excessive degree.

¹⁶ The monopsony analogy also explains why the “pecuniary” externality associated with tariffs yields global inefficiency. Pecuniary externalities do not produce inefficiency if all actors are price takers, but the importing country in the optimum tariff setting has the power to affect price.

policy choices,¹⁷ tax policies, and so forth can reduce demand for imports and cause their prices to fall, effectively extracting profits from foreign suppliers. Because of this negative international externality, importing governments will make policy decisions that are globally inefficient when they underweight the well-being of foreigners in their decision making, and international cooperation through trade agreements can ameliorate the problem.

2. *International Externalities and IIAs*

Host country policies that impose costs on foreign investors can extract profits from those investors when the host country has monopsony power in relation to foreign investors, much as a tariff may extract profits from foreign exporters who reduce their prices in response to the tariff. As an example, the host country might impose a tax on inbound foreign investment, which investors will absorb in whole or in part in preference to foregoing investment opportunities. As another example, a host country might impose costly regulatory burdens on foreign-owned investors operating within its territory when investors are unable to pass along the cost to their customers and instead absorb some or all of it.

Such costs may be imposed up front before the foreign investment is undertaken, as in the first example, or after the fact, as in the second. What matters is not the timing of the costs imposed on investors, but the fact that investors *absorb* some or all of the costs rather than passing them along to the host country through some price mechanism (such as the price of goods that the investors produce and sell to the host country, or the interest rate charged to the host country for imported capital). This observation differentiates the international externality issue—the subject of this section—from the issues in Section B to follow, where the costs imposed on foreign investors are assumed to be passed on to the host country through the cost of imported capital.¹⁸

To complete the analogy to the theory of international trade agreements, however, it is necessary to argue that host countries can indeed possess monopsony power in their markets for imported capital. At first blush, this proposition might seem dubious. Global capital markets are enormous,¹⁹ and it seems implausible that an individual capital-importing country would have the ability to affect global returns to capital, let alone the developing countries that sign many IIAs.

The error in this thinking lies in conceptualizing “capital” as homogenous. Capital in the auto industry, for example, is not fungible with capital in the airplane or sugar industry, but instead includes intellectual, managerial, and human capital that is specialized to auto production and cannot earn comparable returns in other industries. Put differently, capital becomes specialized once certain “sunk costs” have been incurred that relate to particular

¹⁷ For further discussion in the regulatory context, see Robert W. Staiger & Alan O. Sykes, *International Trade, National Treatment and Domestic Regulation*, 40 J. LEG. STUD. 149 (2011).

¹⁸ The reader may note that even if the host country lacks monopsony power, an investor that has made irreversible capital investments in the host country may have to absorb costs imposed by the host country if they come as a “surprise” after the fact (and hence are not priced into the required rate of return on capital up front). Host countries will likely have difficulty engaging in such behavior systematically, however, as rational investors will adapt their expectations to take account of it.

¹⁹ One recent estimate suggests that the combined global bond and stock markets have a valuation of roughly U.S.\$160 trillion. See *5 Bond Market Facts You Need to Know*, MOTLEY FOOL, at <https://www.fool.com/knowledge-center/5-bond-market-facts-you-need-to-know.aspx>.

uses of capital. A patent on automobile technology, for example, resulting from sunk investments in R&D, is likely useless for producing other products. Accordingly, imagine an American auto manufacturer that wishes to sell automobiles into an important market like China, but faces heavy tariffs and sizeable transportation costs and thus has trouble exporting successfully. The automaker will value the opportunity to set up a production facility in China to overcome these costs and reach Chinese consumers more effectively, a fact that gives China leverage to extract surplus from the automaker through various types of investment restrictions.

Likewise, specialized capital will value new investment opportunities in places where complementary factors of production are relatively inexpensive. Consider a company that specializes in footwear production and imagine that a large labor force of footwear manufacturing workers is available at fairly low wages in Vietnam. The footwear manufacturer will value the opportunity to open a production facility in Vietnam not because it expects to sell a lot to Vietnamese consumers, but because it can hire a labor force at favorable wages, reduce its costs, and sell a greater volume of footwear at a higher profit margin. The added profit available from investment in Vietnam again gives the government leverage to extract surplus from the investor in exchange for the opportunity to invest.

The first scenario suggests that host country monopsony will be important whenever the host country market is large enough to consume a considerable portion of the output from an investor's local production facility, and trade costs (transportation costs and trade barriers) are significant enough to advantage localized production. The second scenario suggests that host country monopsony will be important when the host country has abundant inputs into the production process that are relatively inexpensive compared to other sources of supply, and that are not cheaply tradeable (such as labor services or other inputs that are expensive to transport or face important trade barriers).

Specialized capital investment thus explains why access to foreign markets often "matters" to foreign investors. In such cases, costly restrictions on foreign investment create a negative externality for investors just as tariffs create a negative externality for exporters. The externality takes the form of a reduction in the return on investment to foreign investors with specialized capital. From the perspective of the capital-importing country, this amounts to a reduction in the cost of capital, analogous to the reduction in the price of imported goods due to a tariff. One can imagine host countries deliberately exploiting their monopsony power over specialized capital imports to extract surplus from foreign investors. One can also imagine policies that simply have the *effect* of imposing a negative externality on foreign capital whatever their underlying motivation, arising from a political equilibrium in which the welfare of foreigners is undervalued.

As in the trade area, the political impetus for policies that harm foreign investors may often be old-fashioned protectionism. Inbound foreign investment can be a substitute for inbound trade in the face of trade barriers,²⁰ and the owners of domestic capital may see their returns diminished by competition from inbound foreign capital. Well-organized domestic capitalists may then prevail on their governments to afford protection from that competition. The political dynamics of this "investment protectionism" are somewhat different from the dynamics

²⁰ See, e.g., Emily J. Blanchard, *Reevaluating the Role of Trade Agreements: Does Investment Globalization Make the WTO Obsolete?*, 82 J. INT'L ECON. 63 (2010).

of trade protection, to be sure, as barriers to inbound investment will protect domestic capital but may create no benefit or even cause harm to domestic labor because foreign investors often employ significant amounts of domestic labor in their operations. The political support for investment protectionism may thus be weaker, on average, than the demand for trade protection. Nevertheless, the political equilibrium may favor investment protectionism at least some of the time.

Likewise, even when investment protectionism is not the *motive* for policies that burden foreign investors, policies that extract surplus from them can create the same inefficiency by reducing the amount of inbound investment below its globally efficient level. If the host country imposes discriminatory taxes on foreign investors when the investors cannot pass them along, for example, their response may be to curtail the quantity of investment uneconomically.

It bears emphasis, however, that distortions in the amount of inbound investment are by no means the only inefficiency associated with international externalities. Whenever the benefits of a policy inure to the host country but some of the costs are borne by others whose well-being is undervalued by policymakers, the host country may choose to pursue the policy even if the aggregate costs exceed the benefits.

Accordingly, it is instructive to enumerate some of the policies affecting foreign investors that may be (globally) inefficient due to international externalities. An obvious possibility is a tax on inbound capital, akin to the optimum tariff, which directly extracts surplus from foreign investors by reducing the volume of investment and thereby lowering the rate of return that investors demand. Such taxes could be administered in a number of ways. A per unit output tax on the production of foreign-owned firms is a possibility. Alternatively, a fee might be required for an “investment license.” Such a tax might have an advantage over an output tax that it does not affect the marginal costs of foreign-owned firms and avoids the resulting distortion of consumer prices. Discriminatory corporate income taxation offers another possibility.

Stepping away from purely competitive market structures, a host country might employ mandatory technology transfer policies that force foreign investors to share rents from their intellectual property with the host country (a common complaint about China in recent years²¹). “Local content requirements” are another possibility. If a foreign-owned firm in the host country is required to purchase local input products instead of less expensive foreign inputs, some transfer of surplus can occur from the foreign-owned firm to local input suppliers. Likewise, if a foreign-owned firm faces an upward sloping supply of foreign inputs, local content requirements can also divert surplus to local suppliers from foreign input suppliers, who lower their prices when the foreign-owned firm reduces its purchases of foreign inputs.²²

Regulatory requirements are another possible source of surplus extraction.²³ If foreign investors absorb regulatory compliance costs by accepting a lower rate of return on investment in response to them, the host country transfers surplus to itself in the form of regulatory

²¹ See China – Certain Measures on the Transfer of Technology, Request for Consultations by the European Union, WT/DS549/1/Rev. 1 (Jan. 8, 2019).

²² An analysis of these two cases involving local content requirements may be found in Kyle Bagwell & Alan O. Sykes, *India – Measures Affecting the Automotive Sector*, in *THE WTO CASE LAW OF 2002* (Henrik Horn & Petros C. Mavroidis eds., 2005).

²³ See, e.g., Aisbett, Karp & McAusland, *supra* note 2.

benefits for which its citizens need not pay in full. Inefficiently high levels of regulation may result.²⁴

“Export performance requirements” are a further option. Investors may be required to export some proportion of their output, for example, or to export as much by value as they import. When foreign-owned firms are subject to such requirements that force them to deviate from what they would do if allowed to maximize profit without constraint, the effect is comparable to a tax on their investment that can transfer surplus to the host country.²⁵

Restrictions on the transfer of funds out of the host country may also have an effect equivalent to a tax. The host country might require, for example, that the earnings of foreign investors be converted into foreign exchange at a below market rate set by the host government before they can be transferred out of the country.

As Section II will argue, the substantive obligations common to most IIAs address many of these kinds of policy distortions. They involve a variety of disciplines on fiscal measures, regulatory measures, and other policies that can extract profits from foreign suppliers of specialized capital. Some IIAs—especially those embedded in modern preferential trade agreements—also include market access commitments governing the establishment of new investments, which restrict the up-front exercise of monopsony power over inbound investment flows. More will be said about these matters below.

3. *Reciprocity and Enforcement*

It is not in the interest of nations acting unilaterally to eliminate their ability to exploit their monopsony position to extract profits from foreign investors (although they may nevertheless commit not to extract profits through inferior mechanisms that unnecessarily destroy investor surplus, a matter discussed further in Section I.C). Instead, international cooperation is likely to be necessary to orchestrate complete forbearance from the exploitation of monopsony, and treaties are a natural vehicle for this purpose. As in the trade area, governments will negotiate commitments on behalf of their overseas investors and must offer something in return to induce counterparties to make those commitments.

Such commitments are most easily reached in settings where they yield reciprocal benefits to investors—investors in Country A benefit from commitments by Country B, and vice versa.²⁶ In any bilateral relationship where investment flows are significant in both directions, such reciprocal benefits can arise.

Many IIAs arise in settings where investment flows are highly asymmetric, however, such as in many bilateral investment treaties (BITs) between developed and developing countries. If investment flows are mainly in one direction, it is less clear how commitments to eschew the exploitation of monopsony yield mutual benefits. To a degree, this problem may be overcome by issue linkage. Opportunities for issue linkage are most obviously in play when commitments for investment are embedded in larger economic arrangements, such as free trade

²⁴ The same phenomenon arises in the international trade context when a “large” importing country imposes regulatory compliance costs that exporters partially absorb by lowering their prices. See Staiger & Sykes, *supra* note 17.

²⁵ See Dani Rodrik, *The Economics of Export Performance Requirements*, 102 Q. J. ECON. 633 (1987).

²⁶ For a formal model contrasting unilateral and bilateral policies, see Janeba, *supra* note 2.

agreements. A commitment by a host state not to exploit its monopsony power over inbound investment when it exports little capital itself to the counterparty, for example, can be obtained in exchange for tariff concessions or other trade-related measures. Informal issue linkage is also possible even in agreements limited to investment matters, with policies such as foreign aid, military security, and trade preferences for developing countries in play in the background.

A further question arises as to how these commitments will be enforced. Governments will likely play a key role via state-to-state dispute settlement, especially with respect to market access commitments that benefit future investors. Future investors may represent a diffuse group of actors facing a collective action problem in relation to private enforcement actions and would also have great difficulty proving damages from an as yet non-existent investment. For commitments that are in the unilateral interest of host countries, however, ISDS is a stronger candidate for enforcement. Sections I.B and I.C will have more to say about this distinction.

B. The Risk Reduction Function of IIAs

The rationale for IIAs advanced in the last section rests on the existence of international externalities affecting the rate of return to foreign suppliers of specialized capital. In this section, we will ignore such externalities, and consider a rationale for IIAs that would arise even if host country policies do not create any international externality. Thus, for purposes of the analysis below, we may assume that the host country is a “price taker” in the international capital market—investors have a fixed, required *risk-adjusted* rate of return. The cost of imported capital to a host country is then simply a function of the perceived riskiness of investment, and any increase in perceived risk will pass through to the host country in the form of a higher risk premium on imported capital. Under these conditions, no international externality to investors results from host country policies that create risk. Instead, host countries internalize the costs associated with risk, and it is in their *unilateral* interest to eliminate any risks that increase the cost of capital by an amount greater than the benefits to the host country of the actions or policies that create the risk. Risks satisfying this criterion may be termed “inefficient risks,” and this section considers several circumstances under which inefficient risks may arise.²⁷ Along the way, one must also inquire as to the best mechanisms for eliminating inefficient risks and, in particular, the possible role of IIAs. For example, do IIAs dominate unilateral commitments by host countries, embodied in domestic law or in contracts, and if so why? To what extent can insurance markets eliminate inefficient risks?

1. Potential Sources of Inefficient Risk

We begin with a benchmark example in which investment risk is not a source of inefficiency. We then relax the assumptions used to construct the benchmark case to illuminate ways that inefficient risks may arise, and to lay the groundwork for showing how legal instruments can eliminate them.

²⁷ Just as capital becomes specialized and vulnerable to the exercise of monopsony because of sunk investments in intellectual property, human capital, and the like, other investment risks arise because of sunk investments. If investments were fully reversible and investors could always leave with their original capital intact, they would face no risk and have no reason to demand any more than the required riskless rate of return.

a. Benchmark Case(s)

For the benchmark case, assume the following:

- (1) The host country is a unitary actor that maximizes national economic welfare (national income);
- (2) The supply of foreign investment capital is perfectly elastic (unlimited at a fixed rate of return). The riskless rate of return per unit of invested capital required by foreign investors is r , and investors are risk neutral—they are thus indifferent between a certain return of r and an expected return of r . The host country is also risk neutral; and
- (3) When the returns to investment are uncertain, investors have accurate probability estimates *ex ante* regarding all contingencies that affect their returns.

With these assumptions in mind, suppose that investors will contribute an amount of capital K to some investment project, so that the risk-free return they would require is $(1 + r)K = R$. We may presume that the project yields some surplus above R that is enjoyed by the host country, but the exact amount does not matter. Suppose further, however, that the host country may take an action with some probability that will transfer a fraction of the investors' gross returns to the host government. An expropriation of investment returns is one example. A default on government bond payments owed to foreigners is another.

Following assumptions (2) and (3), investors know this information *ex ante*, and they will not invest unless their *expected* return equals R . Accordingly, to induce the investors to participate in the now risky investment, the investors' gross returns before the possible transfer to the host government must exceed the required risk-free return of R by an amount exactly equal to the expected value of the transfer. From the host government's perspective, therefore, the transfer is *at best* a wash—any increase in national income from the expected value of the transfer will be fully offset by a loss of national income due to the higher share of investment returns that investors demand as compensation for bearing the risk of the transfer.²⁸ This zero-sum outcome is in fact the best case scenario from the host government's perspective, because the increase in gross returns required by investors to compensate them for bearing risk may not be possible and the investment may not be made at all—the host government then loses the share of the surplus from the investment that it could earn by protecting investors from the risk of transfer.²⁹

²⁸ Formally, let the gross return to investors (in the state of the world where the host country does *not* transfer a portion to itself) be denoted Y . The host country with probability p will transfer a fraction α of the investors' gross returns to itself. The value of Y will represent the solution to the following expression:

$$R = (1 + r)K = (1 - p)Y + p(1 - \alpha)Y = (1 - \alpha p)Y$$

Hence

$$Y - R = \alpha p Y$$

The expected value of the transfer to the host country is $\alpha p Y$. But the investors require an increase in their gross return Y to offset that amount exactly so that the expected return remains equal to R .

²⁹ To illustrate, suppose that $\alpha p = 1/2$. Then, the gross return to investment Y in the state where the transfer does not occur must equal $2R$ if the investment is to be made at all. Returns at that level may be impossible to achieve.

In summary, under the assumptions of the benchmark case, the host country *internalizes* the costs of any policy that transfers returns from investors to itself. It cannot profit from doing so, therefore, and can only hurt itself if the effect of such policies is to scare socially valuable investment away. A national income-maximizing government would have a unilateral incentive to steer clear of policies that cause such losses, and there is no problem that an investment agreement is needed to solve.

The illustration above considers investment risk associated with a policy that transfers a portion of the returns to investors directly to the host country. That assumption describes certain scenarios such as public expropriation and default on public debt, but other policies do not involve direct transfers to the host government. Consider, for example, a regulatory policy that yields some social benefit from regulatory compliance. The cost of compliance to investors burdened with a regulatory requirement bears no necessary relationship to the social benefit from such a requirement—the benefit may be greater or smaller than the compliance cost.

Hence, one might modify the previous illustration to assume that the host country will discover new information suggesting benefits to the host country from a regulatory policy that will impose a compliance cost on investors. But under assumption (3), investors know exactly the probability of this new information arising and can forecast perfectly how the government will react to it. In this full information setting, the national-welfare maximizing host country would never undertake a regulatory measure unless the social gain from regulation exceeded the compliance cost borne by investors—if and only if the regulation is socially efficient. If the host country behaved otherwise, the increase in the cost of imported capital³⁰ would exceed the social benefits of regulation and national income would fall.

In sum, by maintaining the strong assumptions of the benchmark case, we can readily extend it to host country policies that do not amount to simple transfers from the investor to the host country. This extension does not create inefficient investment risk, or any other problem that requires a solution. The host country fully internalizes the costs imposed on investors and will engage in policies that burden investors only if the social benefits exceed the costs.

The reader is no doubt thinking at this point, however, that the assumptions required for these results are unrealistic much of the time. We now consider the implications of relaxing them in various ways.

b. Host Country Objectives and Agency Costs

The benchmark case above implicitly assumes that the host country is a *unitary actor* that *maximizes national economic welfare*. This assumption—two assumptions actually—may often be unrealistic. Relaxing them introduces the possibility that investment risk is socially inefficient and *may* suggest a role for legal instruments to reduce it.

And in a more general framework, one could imagine a partial loss of surplus to the host government if the investment still occurs but on a smaller scale.

³⁰ More precisely, the cost of the regulatory policy will be passed onto the host state in full, possibly through the price of imported capital, or possibly through the prices charged by the enterprise that the investor owns. The exact mechanism does not matter to the argument.

Consider first the national welfare maximization assumption. It is entirely possible that host country decision makers give more weight to some domestic constituencies than others because some constituencies are poorly organized. This situation can lead to a number of possible inefficiencies (in relation to the national income benchmark). Perhaps regulators undervalue the benefits of regulation to diffuse and unorganized groups of individual citizens, for example, leading to a tendency toward under-regulation. Or perhaps regulators overvalue the benefits of regulation because the beneficiaries are well organized and the domestic constituencies bearing much of the costs are not.

Such inefficiencies in the behavior of the host country government, however, do not necessarily create a problem that the host country would wish to solve. Under the remaining assumptions of the benchmark case, foreign investors are indifferent because they are fully compensated *ex ante*. Domestic officials in the host country may then be quite content to behave inefficiently in relation to their domestic citizenry, responding to political support from well-organized interests and ignoring those that are poorly organized. Accordingly, host countries may at times find it desirable to create or tolerate investment risk that is “inefficient” measured against a national income maximization standard, but we have no reason to expect international agreements to solve this fundamentally domestic “problem.”³¹

The unitary actor assumption, however, is a potentially different matter. Governments are not unitary actors in general but are comprised of individual agents that pursue a variety of interests in their official capacities, interests that can be at odds with the overall social interest of the host country. As an easy example, consider the case of a corrupt public official with the power to impose substantial costs on foreign investors for the official’s personal benefit, such as by requiring costly favors or bribes. Investors may price such costs into the required rate of return *ex ante* and, as a result, the host country may internalize the cost of corruption and the attendant inefficiencies (for example, the award of government contracts to inferior suppliers). But the host country can enhance its national income by eliminating corruption.

As another illustration, imagine that foreign investors at times become embroiled in litigation. And suppose that domestic courts exhibit bias against them, perhaps animated by the fact that the investors are foreign. As a result of this judicial bias, decisions are rendered that produce inefficiency. Perhaps damages are awarded against foreign investors that have done nothing wrong, for example, leading to various forms of inefficient behavior to try to avert future damages awards. Here again, even if investors have priced these risks *ex ante*, the host country may prefer to eliminate the problem because the reduction in the cost of imported capital can exceed any benefits to the host country from judicial bias.

These examples illustrate the general problem of agency costs in the host country that result in actions that fail to promote the national interest. The question then becomes one of how best to eliminate them. Domestic legal instruments will be the preferred option in some cases—perhaps the problem of corrupt officials is best handled through tough criminal penalties for corruption under domestic law. But what if the enforcers of such laws are themselves corrupt?

Similarly, a host country might undertake to solve the problem of bias against foreign litigants by creating rights of appeal domestically. But what if the appellate courts are also biased? And what if appellate litigation is extremely slow and costly?

³¹ Indeed, national income maximization is a controversial normative benchmark, and deviations from that maximand, presumably for distributive purposes, need not be seen as a “problem” at all.

A question then arises as to the possible role of IIAs, and particularly ISDS, in addressing agency costs that create inefficiencies in relation to foreign investors. Plainly, nothing will be accomplished under our benchmark assumptions if the only effect of IIAs is to *compensate* foreign investors for manifestations of agency costs. The direct cost of compensation will simply offset reductions in the cost of imported capital dollar for dollar. But IIAs can be useful if they lead to diminished agency costs. They may generate publicity about misconduct by agents that reduces their propensity to engage in misconduct, for example. Later sections will say more about the role of IIAs as a possible solution to these types of problems.

c. Risk-Bearing Inefficiencies

Another assumption in the benchmark case above is that all actors are risk neutral. “Risk” per se is then of no concern to investors, who care only about their expected returns.

This assumption may be realistic much of the time. Foreign investors will often be large companies owned by diversified shareholders. Foreign portfolio investors will also typically be diversified. Country-specific investment risks to foreign investors will tend to be uncorrelated and thus diversifiable. The non-diversifiable risks that remain will relate primarily to global economic conditions that are not connected to any commitments by host countries under IIAs.

Nevertheless, some foreign investors may be exposed to risk that is not readily diversifiable. Closely held entities, partnerships, and individual proprietorships may involve individuals who hold a high percentage of their wealth in a particular investment. Similarly, the managers of foreign-owned operations may be required to make substantial relationship-specific investments that are worthless if they lose their jobs, and their compensation demands may then increase as they are exposed to greater risk.

In any setting where investors (or their employees) exhibit significant risk aversion, a possible efficiency gain arises if that risk can be transferred to a less risk-averse entity, subject to a standard moral hazard caveat if investors can affect the likelihood or magnitude of the risk. If neither the investor nor the host country has much control over the risk, an insurance entity is a logical choice for laying off the risk. If actions by the host country create or affect the risk, however, the host country may be the superior risk bearer. Placing the risk on the host country incentivizes the host country to minimize it.

Accordingly, IIAs have clear potential to improve risk allocation in situations involving risk-averse investors. Damages awards to investors in ISDS compensate them for adverse events and eliminate the attendant risk and the associated risk premium. In contrast to the benchmark case above, the risk-adjusted required rate of return demanded by investors will decline by more than the value of the expected value of payments from the host government to the investor, yielding a net (expected) gain to the host country. Further gains may arise if placing the risk on the host country induces cost-effective measures to reduce the risk.

d. Imperfect Foresight, Asymmetric Information, and Deception

The third assumption of the benchmark case—that investors accurately assess the probability of all contingencies that may affect their returns—is surely an unrealistic notion in many settings. It can be softened to a degree, however, without affecting the basic conclusions. The analysis changes importantly, however, in the presence of information asymmetries between the investors and the host country.

Consider first the possibility that investors make errors in their assessments of potential future contingencies. They may overestimate or underestimate their probability, or overestimate or underestimate the magnitude of their impact on investment returns. Assume further, however, that these errors are unbiased, and that the host country has no information that would help improve the accuracy of investor estimates.

Under these circumstances, investors are correct on average about the effects of host country policies on their investment returns. The risk-adjusted rate of return that they demand will allow them to earn their required rate of return on average, and the cost of imported capital will be the same for host countries on average if investors are risk neutral. The assumption that host countries have no more information than investors rules out the possibility that host countries can identify cases where investors are misestimating risks and disclose such information selectively.

The incentives for host countries would then be roughly the same as in the benchmark case. They will expect to pay the full costs that their policies impose on investors because those expected costs will be reflected in the cost of imported capital. Any policy drift toward inefficient policies that impose greater costs on investors than the benefits they create for the host country would tend to backfire as capital costs rise by more than the gains from such policies.

The situation is different, however, if for some reason investors' expectations are systematically mistaken. But why would they be? An important possibility relates to the problem of "asymmetric information," a familiar problem that has been studied in a wide range of contexts in economics. In one type of canonical model, host countries might be of two "types"—reliable investment environments and unreliable investment environments. But each country's type may be private information unknown to investors, leading them to treat each host country as if it were a "weighted average" of the two types, resulting in an inefficiently low cost of capital for unreliable types and an inefficiently high cost of capital for reliable types.

A variety of possible solutions may exist. Disclosure by the government of relevant information may help, but such disclosure may not be credible. Ratings agencies such as Moody's may try to address the problem, and insurers might offer some form of default insurance, but they too may suffer from the same challenges of obtaining accurate information.

Now consider IIAs that require compensation to investors who suffer impaired returns under specified conditions. Such agreements will be more costly to unreliable types, and only reliable types may be willing to sign them. If so, they serve as useful "signals" and have the potential to facilitate a "separating equilibrium" in which reliable types sign IIAs and identify themselves,³² allowing investors to draw correct inferences about investment risk in each potential host country.³³

³² See ERIC RASMUSEN, *GAMES AND INFORMATION*, ch. 11 (4th ed. 2007) on the concept of signaling and separating equilibria.

³³ None of this is to deny that situations may arise in which investors have better information than host governments about some risk or hazard associated with an investment and may fail to disclose it in the hope that they can escape regulation and earn greater profit. See Puig & Shaffer, *supra* note 9. This possibility does not alter the analysis in the text, however, as such investors neither require treaty protection nor benefit from protection under sensible IIAs that allow host countries the freedom to engage in sensible regulation. See the discussion of regulatory issues in Section III *infra* (sections on fair and equitable treatment and expropriation). One could imagine hypothetical IIAs that allow host countries to sue investors for taking advantage of their superior information in a

An analytically similar problem arises when host countries (or their agents) deliberately undertake to deceive investors about the risks that they face, or carelessly mislead them. Officials seeking to attract foreign investors may offer false assurances, paint rosy pictures about the likelihood of regulatory approvals, and the like. Investors will not be ignorant of this possibility and will include the risk of deception or careless misrepresentation *ex ante* in their required rate of return unless the host country promises compensation for any harm that is proven to result from such conduct. The costs of a promise to compensate will be lower for host countries that eschew deceptive practices, and it may afford an effective way for them to distinguish themselves. A prospect of monetary liability may also impose a significant political cost on officials responsible for misleading or deceiving investors, and thereby discourage such behavior.

e. Sunk Costs, Time Inconsistency, and the Commitment Problem

Sunk costs in investments are ubiquitous as has been noted. Capital becomes specialized and vulnerable to international externalities because of certain sunk costs, and investment risk in general exists because investments are not fully reversible. Here, however, the concern relates to sunk costs incurred by a foreign investor for the purpose of doing business in a particular host country, costs that will often entail large in-country sunk investments in physical and local human capital. A waterworks cannot be picked up and moved to another city. An automobile factory will have some liquidation value in its real estate and equipment, but likely much less than the cost to build it. Investments in educating and training a local workforce for specialized employment usually cannot be recouped if the venture fails. And so on.

It is widely understood in economics that once such sunk costs have been incurred, the stream of returns necessary to recoup them becomes vulnerable to various actions that may appropriate all or part of the returns, often termed a “holdup problem.”³⁴ Once an investor has incurred the sunk costs to develop an oilfield, for example, the government may expropriate the oilfield leaving the investor with nothing. Alternatively, the government might impose a stiff tax that renders the investment unprofitable after the fact. Or the government may enact a new regulatory policy with compliance costs that leave the investor with negative returns.³⁵

This class of problems is distinguishable from those already considered—one need not invoke “agency costs” to explain them, investors need not be risk averse, and host countries need not have any superior information. Rather, the problem lies with the “time inconsistency” of host country interests.³⁶

To elaborate, let the term *ex ante* refer to the time before sunk costs are incurred, and *ex post* to refer to the time after sunk costs are incurred. Assume that the host government cannot deceive investors *ex ante* about the government’s *ex post* behavior and that investors accurately anticipate it.

manner that imposes unreasonable costs on the host state, but there is no apparent reason why host countries cannot address such matters in their domestic legal systems.

³⁴ Key insights were developed with reference to the challenges of writing complete contracts between private actors. See Benjamin Klein, Robert G. Crawford & Armen Alchian, *Vertical Integration, Appropriable Rents and the Competitive Contracting Process*, 21 J. L. & ECON. 297 (1978). Here, as discussed further below, the problem can be viewed as a contracting problem between the investor and the host state.

³⁵ Time inconsistency issues are a primary focus of the limited existing literature on the economics of IIAs. See sources cited in note 2 *supra*.

³⁶ See James R. Markusen, *Commitment to Rules on Investment: The Developing Countries “Stake,”* 9 REV. INT’L ECON. 287 (2001).

Host governments will then benefit *ex ante* if they refrain from taking actions *ex post* that inefficiently impair investment returns. Investors know, however, that once sunk costs have been incurred, the host government can benefit *ex post* by seizing the returns to sunk investments through expropriation, taxation, regulation, or other policies that benefit the host country, even if such actions are inefficient because the costs to investors exceed the benefits to the host country.³⁷ Concomitantly, even if the host government promises *ex ante* not to behave in this fashion, the government has an incentive to renege on that promise *ex post* and its promise is not credible. The cost of capital will then increase inefficiently, quite possibly to the point that the investment appears hopelessly unprofitable and will not be made at all.

The fundamental problem here is thus a “commitment problem.” Inefficient risk arises because host governments are unable to make credible commitments *ex ante* to eschew inefficient behavior *ex post*.³⁸ When the host country is a price taker in international capital markets, however, it internalizes all the costs of this commitment problem and has a powerful incentive to solve it.

In that regard, market forces may sometimes restrain governments from *ex post* opportunism and help them to ameliorate the commitment problem. In particular, governments that refrain from opportunism may develop reputations for doing so over time and enjoy a reduced cost of capital as a result. More generally, governments are often repeat players in their interactions with certain foreign investors, which enhances the opportunities for strategic interaction to sustain efficient outcomes.³⁹

Nevertheless, these market mechanisms will often be imperfect, and host governments can then benefit from legal measures that enhance the credibility of their commitment to refrain from inefficient *ex post* behavior. A potential mechanism in that regard is a credible promise to compensate investors for any returns lost due to inefficient actions. The next section will reflect further on commitment mechanisms and the role of IIAs and ISDS.

2. Reducing Inefficient Risk

The last section identifies possible sources of inefficient risk to investors, and hints at some options for reducing it. This section elaborates those options.

³⁷ This class of problems bears considerable resemblance to aspects of the international externality problem discussed in Section I.A, in that the same kinds of policies may be involved and the underlying issue arises because investors have incurred certain kinds of sunk costs. The analytical difference in this section is that the costs of host country policies are borne in full by the host country via attendant adjustments in the cost of capital, and hence no international externality arises.

³⁸ Interestingly, a parallel arises once again to the literature on international trade agreements. Even in the absence of international externalities, trade agreements may help governments to “tie their hands” in the face of time inconsistency problems. Standard models involve a government that wishes to commit *ex ante* to efficient trade policy that eschews protectionism and discourages wasteful domestic investment in industries that are not competitive. Once investment occurs, however, the government knows that it will face overwhelming political pressure to afford protection in adverse scenarios. Trade agreements may enable the government to resist those pressures *ex post* and thus make the desired credible commitment *ex ante*. See, e.g., Grossman, *supra* note 12, at 415–21. The difference is that the trade literature concerns the problem of making credible commitments to domestic firms, while the problem for IIAs to solve here involves the need for credible commitments to foreign investors.

³⁹ This proposition follows from the well-known “folk theorem.” See ERIC RASMUSEN, GAMES AND INFORMATION 131 (4th ed. 2007).

As a preliminary, it bears emphasis that *ex post* compensation to investors for the materialization of inefficient risk is simply a transfer from one party to another. Indeed, to the extent that it requires a costly legal process, compensation mechanisms are costly transfers, and some associated efficiency must exist to justify these costs. Otherwise, it is more efficient to let investors protect themselves by demanding higher returns *ex ante*.

In addition, although the focus of this Article is on IIAs, the sources of inefficient risk identified above are not intrinsically “international.” Agency costs arise in closed economies, as does risk aversion, asymmetric information between the government and the private sector, and the possibility of inefficient government action due to time inconsistency problems. To justify solutions that apply to foreign investors in particular, and that are embodied in international rather than domestic law, one must embrace some account of why domestic solutions of general applicability are inadequate. We begin with that issue.

a. The Role of IIAs

In this section we have focused on inefficient risks that arise when host countries are price takers in international capital markets, thus ruling out by assumption international externalities. The risk-adjusted rate of return to investors compensates them for whatever risks they face. What role then is left for international agreements, or for any policies that apply specifically to foreign investors? The answer focuses on the transaction costs of reducing inefficient risk using the domestic legal systems of host countries.

Government agents, like host country governments as a whole, may give little weight to the welfare of foreigners. They will often tend to be more attentive and responsive to the interests of domestic actors who vote, lobby, and otherwise possess political efficacy, and will exercise their discretion to favor domestic constituencies. This behavior plainly undermines the efficacy of domestic legal obligations that may extend to foreigners.

As an illustration, assume that a host country has a general legal principle in its domestic legal system to protect investors against expropriation without adequate compensation. The principle is enforced in domestic courts. But the domestic courts are biased against foreigners, and rule against them on liability and damages issues when the foreign investors should win on the merits. The result is a problem that afflicts foreign investors in particular, and that calls for a solution tailored to them. Such problems may be especially acute in developing countries where courts may be incompetent or corrupt.⁴⁰

An obvious response to this type of problem is to shift adjudicative authority to a competent neutral forum. A panel of international arbitrators with no connection to either party is a possibility, as is a conventional arbitral system where each party picks an arbitrator and the two party-selected arbitrators agree on a third neutral.⁴¹

⁴⁰ To be sure, large multinational companies may at times exercise undue influence over host governments through corruption or the exploitation of superior information and are not always going to be at a disadvantage in the domestic legal system. See Puig & Shaffer, *supra* note 9. This fact does not affect the utility of IIAs where foreign investors are vulnerable, it simply means that foreign investors are not always vulnerable. And where foreign investors have availed themselves of opportunities to benefit themselves through corruption or deception, various doctrines have evolved to deny the investors the benefits of IIA protection. See, e.g., ALOYSIUS P. LLAMZON, CORRUPTION IN INTERNATIONAL INVESTMENT ARBITRATION (2014).

⁴¹ Once again, however, one must acknowledge the possibility that arbitral tribunals may also be biased, perhaps in favor of foreign investors. See Puig & Shaffer, *supra* note 9.

Neutral arbitration is only a partial solution, however, if domestic legal rules are substantively deficient. This too is likely to be a more serious problem in developing countries with weak rule of law traditions and self-serving political leaders.

Contractual solutions for the substantive and procedural inadequacies in domestic law are possible. For investments that are large enough to justify the costs, the investor and the host state might negotiate a contract providing the investor with important substantive protections that domestic law may omit as well as a contractual commitment to arbitrate in a neutral forum. But contracts are costly and may not be economical at all in relation to smaller investments. Off the rack rules that work well most of the time for foreign investors can economize significantly on the transaction costs of contracts. Furthermore, contracts can always be breached or abrogated. What is to prevent a host country from refusing to honor its commitment to neutral international arbitration in the event of a dispute, or refusing to pay a contractual judgment against it?

These observations all point to a role for IIAs. They present an array of options for off the rack substantive rules. Although they will never be comprehensive enough to eliminate the need for customized investment contracts in all cases,⁴² they provide substantial transaction cost savings much of the time. Parties to IIAs typically agree at the time of signing to submit disputes to neutral international adjudication, and adjudicators will accept jurisdiction on that basis. Withdrawal from an IIA is likely to be considerably more costly for a host country than a simple breach of an individual investment contract because of the broader reputational considerations. Parties to IIAs are also typically parties to the New York Convention regarding the enforcement of arbitral awards, which commits them to honor awards when presented for enforcement and enhances the chances that a successful claimant can collect a judgment.⁴³

In sum, IIAs can overcome or at least ameliorate many of the problems associated with efforts by host countries to use domestic law to protect foreign investors from inefficient risks. They are by no means perfect, but they have the potential to: (1) enhance the credibility of the commitment to submit to neutral adjudication; (2) enhance the quality of substantive investor protection rules; (3) reduce the transaction costs of putting efficient substantive and procedural protections in place for foreign investors; and (4) enhance the credibility of a commitment by the host country to pay judgments obtained by foreign investors.

b. Enforcement and the Role of ISDS

It remains to consider in greater detail how IIAs concerning inefficient risk are enforced. Section I.A argued that state-to-state dispute settlement affords the natural enforcement mechanism for IIAs embodying market access commitments, as only government counterparties are likely to overcome the collective action problems faced by future investors, as well as the difficulties of proving damages to investments not yet undertaken. ISDS is only a good candidate for the enforcement of commitments to established investors. But if those

⁴² A further implication of this point is that the substantive commitments of IIAs are best treated as default rules, and that sophisticated investors and host countries should be able to contract around these defaults if they are unsuitable to a particular situation. Arbitral tribunals should not treat treaty provisions as non-waivable by sophisticated investors. For extensive discussion of this issue see Julian Arato, *The Private Law Critique of International Investment Law*, 113 AJIL 1 (2019).

⁴³ The text of the Convention and its role in international arbitration cases is surveyed extensively at <http://www.newyorkconvention.org>.

commitments relate to international externalities, host countries will likely be reluctant to agree to them without reciprocal commitments of value, although they may agree to eschew inferior mechanisms for extracting surplus from investors even without reciprocity (more about that in Section I.C). How does the analysis change when host countries internalize the costs of inefficient risks through the cost of imported capital?⁴⁴

It is in the host country's interest to eliminate inefficient risk when the host country internalizes its cost. Private rights of action are especially attractive in this respect for two reasons. First, investors harmed by inefficient host country conduct cannot always count on their home governments to bring actions on their behalf. Resource constraints, diplomatic and security concerns, and domestic politics may lead home governments to prefer not to enforce the rules in any given case. Second, even if a home government would bring a complaint in the context of a state-to-state dispute settlement procedure, the aggrieved investor may gain little if the result is nothing more than a promise by the host government not to misbehave again in the future (the essential remedy, for example, at the end of state-to-state dispute settlement in the World Trade Organization (WTO)). Money damages restore the investor to the position it would have enjoyed absent the inefficient behavior. Absent compensation, the risk-adjusted required rate of return *ex ante* would increase to cover the expected cost of uncompensated inefficient conduct.⁴⁵

But what ensures that the host country will abide by the ISDS procedure and outcome? With reference to time inconsistency issues noted above, for example, why does the host country not simply ignore the results of ISDS *ex post*? Is a credible state-to-state dispute settlement prospect necessary to ensure that ISDS outcomes will be respected?

State-to-state dispute settlement may sit in the background for this purpose, but it will not be necessary if the host country's reputation in the capital markets offers an adequate substitute. A host country that refuses to abide by ISDS rulings destroys its credibility with future investors and undermines its own rationale for entering IIA commitments to reduce inefficient risk in the first place. This observation affords a further advantage to IIAs not yet emphasized—ISDS litigation may be much more visible and well publicized than litigation in the domestic courts of a host country. The existence of investment disputes is usually public and can be tracked on various official and unofficial websites that monitor investor-state litigation.⁴⁶ Decisions by arbitral tribunals are increasingly detailed and publicly available. In the cases where losing governments refuse to pay judgments, and some such cases do exist, their recalcitrance generally becomes public as well.

Thus, even if individual investors have limited leverage against host countries, the ultimate enforcement of commitments to eliminate inefficient risks may lie with the capital markets. Host country reputations will be based on reasonably good information, and host countries will then tend to internalize the costs of opportunism.

If ISDS has the virtues noted above, however, why has ISDS become so controversial in recent years? Various politicians and non-governmental organizations are critical of ISDS as noted in the introduction. Some prominent economists such as Dani Rodrik, Jeffrey Sachs,

⁴⁴ For a formal model addressing these issues, see Ralph Ossa, Robert W. Staiger & Alan O. Sykes, *Dispute Settlement in International Trade and Investment Agreements* (mimeo 2019).

⁴⁵ See Sykes, *Public Versus Private Enforcement*, *supra* note 7.

⁴⁶ Published arbitral decisions may be accessed promptly, for example, at www.italaw.com.

and Joseph Stiglitz have also questioned the wisdom of IIAs on the grounds that ISDS creates a cause of action against socially desirable policy changes, chills beneficial regulation in a number of areas, and lacks adequate procedures.⁴⁷

The analysis in Section III will take issue with several specific claims in this regard, arguing that the purported dangers are exaggerated, often rest on misunderstandings of prior arbitral decisions, and have been addressed by changes in the language and interpretation of more recent IIAs. But the critique of ISDS nevertheless raises an important theoretical point. If the treaty obligations subject to ISDS are poorly designed and provide investors with a legal basis to challenge host country policies that are in fact socially efficient, ISDS becomes not an instrument for reducing inefficient risks but a source inefficient rent-seeking by investors. The same problem arises if arbitral tribunals make errors in identifying manifestations of inefficient risk due to the vagueness of treaty obligations, arbitrator bias, or the tribunals' inability to secure the information necessary to make a correct decision.⁴⁸

In short, ISDS has the potential to serve as a valuable mechanism for controlling inefficient risks, and its utility in that regard offers a powerful explanation for its widespread prevalence in modern IIAs. But errors and inaccuracy in arbitral decisions, as well as poorly drafted substantive obligations, can undermine the value of ISDS and lead IIA signatories to reject it. Some of the recent developments in treaty practice may reflect this concern. Brazil's relatively new model for Cooperation and Facilitation Investment Agreements (CFIA), and the handful of IIAs resulting from it thus far, relies entirely on state-to-state dispute settlement not only for market access issues but for protecting established investors from policies such as unjustified expropriation.⁴⁹ India has terminated a number of its BITs in recent years, and seeks to replace them based on a new Indian model BIT that limits ISDS. Most importantly, the new Indian model requires the investor to pursue domestic remedies for a period of five years before resorting to ISDS.⁵⁰ Ecuador, Indonesia, and Australia have also expressed skepticism toward the utility of ISDS in recent years, largely motivated by high-profile investment litigation, although only Ecuador has moved to eliminate its involvement with ISDS.⁵¹

⁴⁷ See Dani Rodrik, *What Do Trade Agreements Really Do*, 32 J. ECON. PERSP. 73 (2018); Lise Johnson, Lisa Sachs & Jeffrey Sachs, *Investor-State Dispute Settlement, Public Interest and U.S. Domestic Law* (CCSI Policy Paper, May 2015), available at <http://ccsi.columbia.edu/files/2015/05/Investor-State-Dispute-Settlement-Public-Interest-and-U.S.-Domestic-Law-FINAL-May-19-8.pdf>; Joseph E. Stiglitz, *Towards a Twenty-First Century Investment Agreement*, in YEARBOOK ON INTERNATIONAL INVESTMENT LAW AND POLICY 2015–16 (Lise Johnson & Lisa Sachs eds., 2018).

⁴⁸ Ossa, Staiger & Sykes, *supra* note 44, develop a formal model in which the desirability of ISDS turns importantly, inter alia, on the accuracy of tribunals. If tribunals are perfectly accurate in identifying inefficient decisions by host governments, disputes do not arise because neither party can benefit from opportunism and it does not occur in equilibrium. If tribunals are horribly inaccurate, then the parties cannot benefit from ISDS. The case of intermediate accuracy creates the possibility of disputes arising in equilibrium, and ISDS will prove more valuable the greater the accuracy of the tribunal, the greater the cost of host country opportunism, and the greater its probability.

⁴⁹ See notes 60–61 *infra* and accompanying text.

⁵⁰ See Model Text for the Indian Bilateral Investment Treaty, Art. 15, available at <https://investmentpolicyhub.unctad.org/Download/TreatyFile/3560>.

⁵¹ Ecuador terminated its BITs after a multi-billion-dollar judgment against it on behalf of Occidental Petroleum. See Cecilia Olivet, *Why Did Ecuador Terminate All Its Bilateral Investment Treaties?*, TNI (May 25, 2017), at <https://www.tni.org/en/article/why-did-ecuador-terminate-all-its-bilateral-investment-treaties>. Australia threatened to withdraw from its BITs and the ISDS system in response to the plain-packaging tobacco litigation filed by Phillip-Morris. Australia changed its view after a change in government, however, and ultimately prevailed in the plain-packaging case—the current position of Australia remains favorable to ISDS. See Australian

Accordingly, ISDS is not a panacea for the problem of inefficient risks to established investors. But if the substantive obligations enforceable through ISDS are well-designed and tribunals are able to administer them with sufficient accuracy, ISDS offers important advantages over state-to-state dispute settlement and local remedies under domestic law.

c. Why Not Rely on Political Risk Insurance Instead?

Where the objective is to eliminate inefficient risk for investors, a potential alternative to IIAs is the insurance market. Indeed, markets exist for the sale of “political risk” insurance, which covers losses to investors in the event of certain covered risks associated with investments in developing countries. A prime example of such insurance is that offered by the Multilateral Investment Guarantee Agency (MIGA), an arm of the World Bank. MIGA insurance is available to cover losses from currency conversion or transfer restrictions, expropriation, breach of contract, war, terrorism and civil disturbance, and payment defaults by governments and state-owned enterprises.⁵²

Several reasons exist, however, as to why these insurance products are not an adequate substitute for IIAs. First, unless the insurance is subsidized (it is unclear how MIGA premiums in general relate to “actuarially fair” premiums), insurance will not be valuable unless the investors are risk averse. Risk-neutral investors will not benefit from insurance sold in conventional insurance markets, which has negative expected value given that premiums must cover the expected value of the risk plus the overhead costs of insurers.

Second, although coverage is available for a range of contingencies, as noted, the availability of coverage does not extend to all of the commitments typically embodied in IIAs. Violations of non-discrimination principles or of the minimum standard of treatment, for example, are not among the coverage options listed by MIGA.

Third, although insurance products can compensate investors for certain inefficient risks, they do not directly reduce them. IIAs induce host countries to internalize the costs of inefficient risk and thus provide an incentive to curtail them. To be sure, insurance markets may have an indirect effect on host country incentives. High insurance premiums may damage a host country’s reputation among investors. Moreover, when host countries have an interest in development lending, MIGA may be able to leverage its position in the World Bank to link development lending to better treatment of foreign investors.

Accordingly, insurance markets have a constructive role to play in addressing inefficient risk for investors. But they do not offer a complete substitute for commitments in IIAs, and do not have as much potential to restrain inefficient behavior by host countries.

Government, Department of Foreign Affairs and Trade, Australia’s Bilateral Investment Treaties, at <https://dfat.gov.au/trade/investment/Pages/australias-bilateral-investment-treaties.aspx>. Indonesia has terminated a number of its BITs and embarked on an effort to renegotiate them, sparked in part by a response to case brought by Churchill Mining. See Leon E. Trakman & Kunal Sharma, *Why Is Indonesia Terminating Its Bilateral Investment Treaties?*, E. ASIA F. (Sept. 20, 2014).

⁵² Multilateral Investment Guarantee Agency, *Investment Guarantee Guide* (July 2015), available at https://www.miga.org/documents/IGG_English_final.pdf.

C. Combining the International Externality and Risk Reduction Accounts

The analysis in Section A assumed that the capital-importing country possesses leverage to extract surplus from foreign investors with specialized capital. That assumption leads to the possibility that capital-importing countries will engage in a variety of inefficient policies due to the existence of international externalities, a problem that IIAs can potentially solve. The analysis in Section B made a different assumption—that capital-importing countries internalize the costs of any policies that reduce the returns to foreign investment through an increase in the cost of imported capital. Agency costs, risk aversion, imperfect information, and time inconsistency problems can still create inefficient risks for investors to the detriment of host countries, however, a problem that IIAs can also potentially solve.

How does the analysis change if we allow for both scenarios simultaneously—that is, if we suppose that host countries have the leverage to extract surplus and create international externalities, but may concurrently create inefficient risks? One can decompose this question into sub-questions: First, will IIAs eliminate all sources of inefficiency? And second, if a host country does not give up its ability to extract surplus from foreign investors, might they prefer to retain the flexibility to create inefficient risks?

The answer to the first sub-question is that an ideal IIA will eliminate all sources of inefficiency that lower the welfare of counterparties, including both the inefficiencies associated with investment protectionism and the costs of inefficient risks. There is no incentive for the parties to leave any source of mutual gains on the table if they can eliminate inefficiencies and divide the surplus between themselves.

It is possible, however, that a country with the capacity to extract substantial surplus from foreign investors may not be willing to give up that ability even if it produces global inefficiencies. The benefits of doing so in an IIA may not be sufficiently reciprocal if investment flows are highly asymmetric between the counterparties, as noted, and arrangements for overcoming such asymmetry, such as issue linkage,⁵³ may be difficult to orchestrate. It is also possible that parties to IIAs will lack monopsony power in relation to their counterparty, in which case they will have no need to address it. More will be said about this issue in Section D.

The answer to the second sub-question is generally no. Even if a host country wishes to retain the ability to extract surplus from foreign investors, it is in the host country's interest to do so with as little economic waste as possible. An IIA that does not altogether disable a host country from exploiting its monopsony leverage over foreign investors, therefore, may still tend to protect investors against inferior mechanisms for surplus extraction that impose unnecessary costs that dissipate surplus and leave less to extract.

To illustrate, suppose that a host country can extract profits from investors in two ways—by charging a fee for an investment license, or by engaging in socially inefficient regulation when foreign investors absorb a substantial share of the costs. The host country can do better under the first option, which effects a monetary transfer from the investor to the host, rather than the second, which reduces the available surplus by wasting resources on socially inefficient regulation.

Under these conditions, we would expect the host country *unilaterally* to eschew the inefficient mechanism for surplus extraction in order to extract more with the efficient

⁵³ See Section I.A.3 *supra*.

mechanism. But might it also make a commitment in an IIA to behave that way even if it receives no reciprocal commitment of any value? The answer may be yes if the commitment to eschew wasteful regulation is also valuable to reduce inefficient risks that the host country must internalize. Then, the host country commitment yields a benefit in the class of cases where the risk reduction rationale for IIAs applies and is costless in the class of cases where the international externality problem arises because the host country retains another (superior) instrument for extracting surplus. We might thus expect to observe IIAs in which the host country retains the right to charge a fee or impose a tax for market access, for example, yet nevertheless agrees to steer clear of policies that burden established foreign investors with socially wasteful obligations.

D. The Relative Importance of the International Externality and Risk Reduction Accounts in Practice

Sections A and B identify two distinct possible functions of IIAs—commitments to address the inefficiencies associated with international externalities, and commitments to eliminate inefficient risks to investors that increase the cost of imported capital. Do existing IIAs tell us anything about which function is more important, either economically or politically?

One indicator of the degree to which an IIA constrains the international externalities associated with host country monopsony is whether it precludes discrimination between domestic and foreign investors regarding the initial opportunity to make an investment. Put differently, an important indicator is whether the IIA provides national treatment for the *establishment* of investments. In its absence, host countries retain the ability to exploit their monopsony power through up-front fees and related methods, even if they are limited in what they can do to extract profits once an investment is established.

The UN Conference on Trade and Development (UNCTAD) sponsors a treaty mapping project covering IIAs, although it is incomplete and omits many of the investment provisions found in free trade agreements as well as many newer agreements such as those of Brazil. According to the current UNCTAD database, however, only 6.5 percent of the included IIAs contain a national treatment obligation applicable to the *pre-establishment* phase of investments. In contrast, 85 percent of the included IIAs require national treatment *subsequent* to the making of an investment. The situation is markedly different for BITs involving the United States, all of which require national treatment in both pre- and post-establishment periods.⁵⁴ These results suggest that non-discriminatory market access per se is not a central objective for the bulk of IIAs not involving the United States, and hence that instruments for the exploitation of monopsony *pre-establishment* are not constrained in most IIAs.

A further indicator in this regard involves the number of formal ISDS and state-to-state dispute settlement disputes. As noted earlier, market access commitments to future investors will tend to be enforced by state-to-state dispute settlement while commitments to established investors are candidates for ISDS. ISDS cases now number in the many hundreds, while formal state-to-state dispute settlement cases are uncommon.⁵⁵ The fact that disputes

⁵⁴ The current results are accessible at <http://investmentpolicyhub.unctad.org/IIA/mappedContent>.

⁵⁵ UNCTAD reports 942 known cases as of year-end 2018. See UNCTAD, Investment Policy Hub, at <https://investmentpolicy.unctad.org/news/hub/1611/20190528-fact-sheet-surge-in-isds-cases-continues-in-2018>. There is no comparable database on state-to-state claims, but academic writing suggests that such claims are quite

regarding restrictions on initial market access are few also hints that the discipline of international externalities may be less significant in practice, although one must be extremely careful in drawing inferences from the number of disputes alone.⁵⁶

A different picture emerges, however, from a recent WTO study of the investment provisions in preferential trade agreements (PTAs).⁵⁷ From a sample of 230 PTAs, the authors identified 111 agreements with investment provisions. The analysis finds that “88 percent of the PTAs surveyed with an investment chapter provide for national treatment in the pre-establishment/acquisition phase of the investment, thus requiring the host state to remove all discriminatory market access barriers and allow foreign investors to invest on the same terms as domestic investors.”⁵⁸ Thus, the dominant practice in modern PTAs is to ensure non-discriminatory market access for investors *if* the agreement addresses investment at all. These protections may be found in many of the most significant PTAs among major players, such as the North American Free Trade Agreement (NAFTA) (and the proposed United States–Mexico–Canada Agreement (USMCA)), the EU–Canada Comprehensive Economic Trade Agreement (CETA), the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP), and of course the EU.⁵⁹

Brazil’s recent approach to investment agreements also presents an interesting case. It has negotiated “Cooperation and Investment Facilitation Agreements” (CIFAs) in lieu of traditional BITs with certain countries. The Brazilian approach deemphasizes protection for existing investments but *does* include national treatment obligations with respect to the establishment of investments.⁶⁰

What might explain this overall pattern of attention to pre-establishment market access policies? The answer is no doubt complicated, but several factors come to mind. First, many IIAs involve at least one small country party, and in many cases two. The value to a counterparty of market access for investment in a small country can be limited if the number of consumers is small and the resources available for production in that country are limited. Small countries may then have little leverage to extract surplus from foreign investors with specialized capital and have no reason to address that issue in their IIAs.

Second, even if one or both counterparties has non-trivial monopsony power in relation to some form of specialized capital, the other party may not care. We would not expect to see market

few in number. See Anthea Roberts, *State-to-State Investment Treaty Arbitration: A Hybrid Theory of Interdependent Rights and Shared Interpretive Authority*, 55 HARV. INT’L L.J. 1 (2014); Nathalie Bernasconi-Osterwalder, *State-State Dispute Settlement in Investment Treaties* (International Institute for Sustainable Development, Oct. 2014).

⁵⁶ There are a variety of reasons why the number of formal state-to-state dispute settlement disputes may not be a reliable indicator here—compliance rates with market access commitments could be quite high, for example, obviating the need for formal disputes.

⁵⁷ This category includes both free trade agreements and customs unions.

⁵⁸ Jo-Ann Crawford & Barbara Kotschwar, *Investment Provisions in Preferential Trade Agreements: Evolution and Current Trends*, 17 (WTO Staff Working Paper ERSD-2018-14, Dec. 14, 2018).

⁵⁹ Under the Treaty on the Functioning of the European Union, pre-establishment “national treatment” is not formally included but the principle is implicit in various non-discrimination, right of establishment and free movement provisions. See Consolidated Version of the Treaty on the Functioning of the European Union, Arts. 18, 49, 63, OJ C 326/47 (Oct. 26, 2012), available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3AA12012E%2FTXT>.

⁶⁰ Institute for Sustainable Development, *Side-By-Side Comparison of the Brazil–Mozambique and Brazil–Angola Cooperation and Facilitation of Investment Agreements* (2015), available at <https://www.iisd.org/sites/default/files/publications/comparison-cooperation-investment-facilitation-agreements.pdf>. See Article 11(2) in each treaty.

access issues emphasized unless at least one of the counterparties is a significant capital exporter to the other. Indeed, market access commitments in IIAs are only important between countries where one or both is an important supplier of the relevant *specialized* capital to the other.

Third, IIAs are frequently based on “model” agreements put forward by the larger counterparty. If the “model” contains a pre-establishment national treatment obligation, it may pervade the resulting IIAs regardless of the economic circumstances found in a particular counterparty.

Finally, as noted, nations with significant capacity to extract surplus from foreign investors will not cede it in exchange for nothing. Some form of reciprocal benefit is needed. This can arise because of significant investment flows in both directions, as between two sizeable developed countries, or when issue linkage is available to provide the capital importer with some benefit outside the investment context in exchange for giving up monopsony power.

Combining these points, many BITs may involve parties with little monopsony power or little specialized capital affected by monopsony power in a counterparty. PTAs, by contrast, may tend to involve more substantial economies, and in any event offer the opportunity for issue linkage involving the non-investment aspects of the PTA. The inclusion of pre-establishment national treatment clauses in Brazil’s CFIA may well reflect the growing importance of Brazil as a capital exporter.⁶¹ And the fact that both U.S. and Brazilian agreements tend to include pre-establishment national treatment clauses, even with small country counterparties, may reflect the role of their respective “model” agreements in negotiations.

These observations suggest an opportunity for some interesting empirical research concerning the prevalence of market access provisions in IIAs, an exercise that is beyond the scope of this project. The important point here is that the market access function of IIAs is potentially important and is embedded in a growing number of modern IIAs.

Finally, as argued extensively above, the international externalities associated with host country monopsony power in relation to specialized foreign capital do not merely produce distortions in initial market access but can result in distortions in post-establishment policies such as regulation and taxation. This danger arises any time the costs of such policies are borne in part by foreign investors whose welfare is not given equal weight with the welfare of domestic actors. Even if an IIA does not ensure pre-establishment market access, therefore, some of its provisions may act as a useful constraint on post-establishment policies that may otherwise be distorted by international externalities. And as noted, host countries may be willing to make such commitments even without reciprocity if these post-establishment measures are inferior ways of extracting investor surplus.

E. Alternative Accounts of IIAs that Challenge the Efficiency Rationales

The preceding sections advance two theoretical rationales for IIAs, and subsequent sections will argue that IIAs in practice serve both functions to an important extent. Before proceeding, however, we pause briefly to consider the empirical evidence on the economic effects of IIAs and some alternative accounts that exist in the literature.

Several empirical scholars have undertaken to study the effects of IIAs on the flow of investment into treaty counterparties. The studies include econometric studies of the effects of IIAs

⁶¹ See *Brazil Direct Investment Abroad*, CEIC, at <https://www.ceicdata.com/en/indicator/brazil/direct-investment-abroad>.

on investment as well as surveys of business executives about the importance of investment treaties on their investment decisions. The results are mixed, with some studies finding no effect of IIAs on investment flows, some finding large positive effects, some finding that business executives care little about the existence of investment treaties, and some finding that significant numbers of executives care a lot about treaty protection.⁶² About all one can say based on the now rather extensive literature is that the impact of IIAs on investment is “contested.”⁶³

Various scholars have seized on the empirical evidence suggesting weak effects from IIAs to put forward alternative accounts of their emergence. One strand of literature emphasizes the “bounded rationality” of treaty signatories, especially developing countries, to suggest that they were fooled into signing treaties that they expected to bring substantial new investment but did not.⁶⁴ Another U.S.-centric analysis suggests that IIAs arise primarily between the United States and its political allies, and are entered to strengthen strategic alliances rather than to promote investment.⁶⁵ A third account casts developing countries as participants in a global “prisoner’s dilemma,” where each developing country expects to benefit from an IIA by attracting capital away from its developing country competitors, yet when all developing countries sign IIAs on that basis, the net effect on inbound investment is negligible.⁶⁶

All of these accounts are intriguing but also have their difficulties. The notion that countries have been fooled into signing useless IIAs is challenged by the fact that thousands of agreements have been signed over the course of more than a half century. Bounded rationality is one thing, but inability to learn from the past is another. True, a few countries (such as Ecuador) have concluded that their IIAs imposed more costs than benefits, but outright withdrawal from the IIA system is rare (India and Indonesia, for example, appear more interested in renegotiation than withdrawal).⁶⁷ If the “bounded rationality” account has wide purchase, why have countries not withdrawn in droves once the counterproductive nature of these treaties became apparent?⁶⁸

The notion that IIAs are just about “politics” suffers from the lack of a clear explanation as to how they have any political significance if they have no economic impact. Even if the United States is more likely to sign IIAs with allies rather than enemies, as the literature

⁶² An extensive collection of papers on the subject may be found in *THE EFFECT OF TREATIES ON FOREIGN DIRECT INVESTMENT: BILATERAL INVESTMENT TREATIES, DOUBLE TAXATION TREATIES, AND INVESTMENT FLOWS* (Karl P. Sauvant & Lisa E. Sachs eds., 2009). See also the survey in BONNITCHA, *supra* note 2, at 105–09. A relatively recent study arguing that IIAs can significantly increase investment flows and enhance the credit and good governance ratings of signatories is Jeswald W. Salacuse, *Of Handcuffs and Signals: Investment Treaties and Capital Flows to Developing Countries*, 58 HARV. J. INT’L L. 127 (2017).

⁶³ See Joost Pauwelyn, *Rational Design or Accidental Evolution: The Emergence of International Investment Law*, in *THE FOUNDATIONS OF INTERNATIONAL INVESTMENT LAW: BRINGING THEORY INTO PRACTICE* (Zachary Douglas, Joost Pauwelyn & Jorge E. Viñuales eds., 2014).

⁶⁴ See, e.g., POULSEN, *supra* note 6.

⁶⁵ See Adam Chilton, *The Political Motivations of the United States Bilateral Investment Treaty Program*, 23 REV. INT’L POL. ECON. 614 (2016).

⁶⁶ See Guzman, *supra* note 6; Zachary Elkins, Andrew Guzman & Beth Simmons, *Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960–2000*, 60 INT’L ORG. 811 (2006).

⁶⁷ See sources cited in note 51 *supra*.

⁶⁸ We do not mean to suggest that all IIAs are initially “optimal” in every respect and cannot be improved. Indeed, as discussed later, one observes considerable evolution in the terms of IIAs over time. Renegotiation is also observed with some regularity. See Timothy Meyer & Tae Jung Park, *Renegotiating International Investment Law*, 21 J. INT’L ECON. L. 655 (2018).

suggests, that phenomenon is entirely consistent with the proposition that IIAs are expected to stimulate investment (or lower the cost of imported capital) and benefit the economies of counterparties, which would naturally discourage the United States from signing IIAs with its enemies.

The prisoner's dilemma account is subject to the objection that prisoners face no dilemma when they can communicate and coordinate their behavior. Developing countries can and do act collectively through organizations like UNCTAD and it is unclear why they would not agree among themselves to resist competitive pressures to sign harmful IIAs. Likewise, in the classic prisoner's dilemma, when a prisoner confesses, they have no ability to take back their confession later. But when countries sign IIAs, if they come to regret it because the agreements are nothing but a counterproductive competition for capital with others, why do they not coordinate on collective withdrawal from IIAs? We have observed some withdrawals to be sure, but they seem to be unilateral responses to adverse experience with IIAs rather than coordinated moves among competitors for imported capital.

Finally, all of these accounts proceed from the highly contested premise that IIAs have little or no economic impact on investment flows or the cost of capital. Concededly, a number of empirical analyses reach this conclusion (and a number of others do not), but all of the empirics suffer from a host of conceptual and statistical challenges,⁶⁹ and it is little surprise that existing studies vary in their findings. Rather than focusing on the mixed econometric record, the best evidence that IIAs benefit the signatories, one might argue, comes from the dramatic growth in the number IIAs over a roughly sixty-year period, including the increasing integration of investment commitments into modern PTAs, the limited number of withdrawals, and the fact that even the recent IIA terminations by major economies such as India and Indonesia seem a prelude to renegotiation than a rejection of the IIA approach altogether. It is difficult to square these simple facts with the notion that IIAs are useless or counterproductive.

II. IMPLICATIONS FOR THE DESIGN AND STRUCTURE OF IIAs

Although IIAs are not uniform in their provisions, the vast majority contain a handful of basic obligations. This section reviews the central substantive commitments found in most IIAs and affords a "bird's-eye view" of their central rationale, leaving to Section III the task of addressing certain controversial aspects of their interpretation. We do not consider procedural or dispute resolution provisions in this section, as the rationale for private rights of action before neutral international arbitrators (ISDS)—a feature found in most IIAs—was explored at some length in Section I. A central claim in this section is that many of the common substantive obligations in IIAs serve *both* to constrain (at least some) policies that create international externalities and other policies that create inefficient risk even in the absence of such externalities.

A. National Treatment Requirements

National treatment obligations are ubiquitous in IIAs and are present in some form in 85 percent of the IIAs covered by the UNCTAD mapping project as noted earlier, although a considerably smaller percentage contain pre-establishment national treatment obligations.

⁶⁹ See sources cited in note 62 *supra*.

Even if an IIA does not ensure market access through pre-establishment national treatment, however, it does not follow that the post-establishment national treatment obligation has no effect on distortions associated with international externalities. As suggested in Section I.C, a host country might enact a socially wasteful regulatory policy that it deems nationally beneficial only because some of the compliance costs are borne by foreign investors. Such a policy might take the form of discriminatory regulations that apply exclusively to foreign companies. The post-establishment national treatment principle precludes such policies. If the host country nevertheless wishes to exploit its monopsony leverage, it will prefer to do so by using more efficient up-front taxes or fees that do not create regulatory waste.⁷⁰

The risk reduction function of national treatment obligations is also readily apparent. Discrimination against foreign investors or their investments is precisely what one would predict from an environment in which the welfare of foreigners is discounted, either because of the agency cost issues noted earlier (such as judicial bias) or the time inconsistency problem. Violations of the national treatment obligation thus serve as a marker for the materialization of certain inefficient risks that host countries can benefit from eliminating. Money damages for the harm done to established foreign investors eliminate the need for them to price these risks into required returns *ex ante* and discourage host countries (and their agents) from imposing the attendant costs *ex post*. They also insure risk-averse investors against losses and lower their risk premium in capital markets. Finally, with respect to the national treatment obligation and every other commitment discussed in this section, the willingness of the host country to make the commitment may signal that the host country is a comparatively safe venue for investment in the face of information asymmetries.

B. *Most-Favored-Nation Requirements*

Most-favored-nation requirements are also routinely found in IIAs. The UNCTAD mapping project shows that some form of most-favored-nation obligation exists in 99 percent of the included IIAs, although less than 10 percent contain pre-establishment most-favored-nation clauses. An example of the latter may be found in NAFTA and the U.S. Model BIT, which mandate “treatment no less favorable” to investors and investments of counterparties than the treatment afforded, “in like circumstances,” to investors or investments of any other party or non-party with respect to “the establishment, acquisition, expansion, management, conduct, operation, and sale or other disposition of investments in its territory.”⁷¹

Most-favored-nation obligations too can relate to both the international externality issue and the inefficient risk issue. Their role with regard to the former is analogous to the role of most-favored-nation obligations in trade agreements, which assure counterparties to such agreements that importing nations will not offer a “better deal” to some third country, a prospect that introduces the danger of “concession erosion.” To see the logic, imagine a sequence of negotiations over market access, initially between Country A (the capital importer) and Country B (a capital exporter), followed later by a market access negotiation between

⁷⁰ This observation is a corollary of the fact that protectionism through fiscal instruments such as tariffs is in general less inefficient than “regulatory protectionism.” See Alan O. Sykes, *Regulatory Protectionism and the Law of International Trade*, 66 U. CHI. L. REV. 1 (1999).

⁷¹ North American Free Trade Agreement, Art. 1103, 32 ILM 289 (*entry into force* Jan. 1, 1994) [hereinafter NAFTA]; U.S. Model BIT, Art. 4.

Country A and Country C (another capital exporter). Absent the most-favored-nation commitment, the second negotiation might result in more favorable market access for Country C than for Country B, which will impose an external cost on Country B when its investors become competitively disadvantaged. The inefficiency will then manifest itself through an uneconomic expansion of the investments of Country C's investors and an uneconomic contraction of the investments by Country B's investors. It may also create an unwillingness of nations to enter IIAs at all.⁷² A pre-establishment, most-favored-nation obligation curtails this problem. And similar to the situation with national treatment, post-establishment most-favored-nation clauses can also police certain inefficiencies associated with host country monopsony power, such as a tendency to shift regulatory compliance burdens onto disfavored foreign investors and away from favored foreign investors.

To appreciate the risk reduction function of most-favored-nation requirements, consider the possibility that a foreign investor has no domestic investors to serve as reliable comparators under the national treatment principle. Discrimination against a foreign investor relative to other foreign investors may then afford alternative evidence of behavior by the host country that results from agency costs or *ex post* opportunism in the face of time inconsistency issues.

C. Minimum Standard of Treatment (Including "Fair and Equitable Treatment")

IIAs often contain a vaguely worded obligation that requires host countries to afford foreign investors a minimum degree of protection, often termed "fair and equitable treatment," which may exceed that required by the national treatment and most-favored-nation requirements. The UNCTAD mapping project reports that about 95 percent of surveyed IIAs contain some form of this clause although the exact wording varies considerably.

NAFTA (and USMCA) and the U.S. Model BIT, for example, embody this obligation in the "minimum standard of treatment."⁷³ The Model BIT version, which provides some clarification in relation to the more open-ended NAFTA language, provides: "Each Party shall accord to covered investments treatment in accordance with customary international law, including fair and equitable treatment and full protection and security." It further underscores that the minimum standard is a reference to the "customary international law minimum standard of treatment for aliens," and that the references to "fair and equitable treatment" and "full protection and security" do not create any greater obligations. "Fair and equitable treatment' *includes* the obligation not to deny justice in criminal, civil, or administrative adjudicatory proceedings in accordance with the principle of due process embodied in the principal legal systems of the world" (emphasis added). "[F]ull protection and security' requires each Party to provide the level of police protection required under customary international law."⁷⁴

Fair and equitable treatment is the subject of considerable ISDS litigation, in substantial part due to its vagueness, and more will be said about it in Section III. But to appreciate

⁷² The problem is analogous to the well-known problem of "trade diversion" that arises when one country receives a tariff preference from an importing nation. *See, e.g.*, KYLE BAGWELL & ROBERT W. STAIGER, *THE ECONOMICS OF THE WORLD TRADING SYSTEM*, ch. 5 (2003).

⁷³ NAFTA, *supra* note 71, Art. 1105; U.S. Model BIT, *supra* note 71, Art. 5.

⁷⁴ U.S. Model BIT, *supra* note 71, Art. 5(2). This elaboration of the two concepts is not found in NAFTA and was introduced into the Model BIT in response to some early NAFTA cases that interpreted the minimum standard of treatment more broadly than customary international law implies.

preliminarily the importance of these “minimum standards,” it is useful to juxtapose them against a national treatment obligation that generally coexists in the background. The minimum standard then has two important dimensions. First, it places a floor on the treatment of foreign investors in situations where there are no domestic investors to establish a benchmark standard of treatment under the national treatment principle. It thereby directly protects foreign investors against certain types of familiar inefficient risks, especially in developing countries. We noted earlier the dangers of corruption and judicial bias that may befall established foreign investors operating in markets with weak domestic protections against such misbehavior by agents. These problems are addressed directly by minimum standards under customary law. Such protections also touch on opportunism in the face of time inconsistency problems—it has been argued elsewhere, for example, that foreign firms with substantial sunk investments are especially vulnerable to extortionate demands by corrupt host country officials.⁷⁵

Second, a minimum standard raises the possibility that foreign investors will be entitled to *better* treatment than domestic investors under circumstances where the treatment of the latter group falls below the customary international law standard. This principle can make sense in the face of weak domestic institutions in the host country. Consider the problem of judicial corruption in a developing country. Corruption is a likely source of inefficiency whether it afflicts domestic or foreign investors, and the ideal solution might seem to be a strong domestic anti-corruption regime applicable to all investors uniformly. But if the host country lacks the capacity to implement such a system domestically, a fallback option may be to implement greater protection for foreign investors under an international legal regime. That arrangement can diminish the inefficiencies suffered by the host country even if others remain.⁷⁶

D. Expropriation

Historically, some of the most prominent disputes between investors and host countries have involved expropriation or “nationalization” of operations established by foreign investors. To a degree, IIAs evolved in response to such events, and to a long-running North-South dialogue over the obligations of host countries that expropriate foreign-owned property.⁷⁷ Virtually all IIAs provide some protections against the consequences of expropriation according to the UNCTAD mapping project, and over 95 percent extend the concept to some form of “indirect expropriation” or measures “tantamount to expropriation.” These extensions have been central to many of the most contentious and controversial ISDS cases, especially when they touch on matters such as environmental or health and safety regulation. A more detailed treatment of the attendant issues will appear in Section III, and the discussion here is limited to an initial sketch of the basic obligations and their relation to the international externality and risk reduction rationales for IIAs.

One complexity in dealing with expropriation is that public taking of property, as well as various government policies that might be deemed “indirect” or “tantamount” to expropriation, can be socially constructive. The term “expropriation” is somewhat pejorative, but it

⁷⁵ See Rebecca L. Perlman & Alan O. Sykes, *The Political Economy of the Foreign Corrupt Practices Act: An Exploratory Analysis*, 9 J. LEG. ANALYSIS 153 (2017).

⁷⁶ This observation is subject to the caveat that better protection for foreign investors may give them an effective “subsidy” that can provide a competitive advantage and lead to inefficiently *high* levels of foreign investment.

⁷⁷ See Guzman, *supra* note 6; Bubb & Rose-Ackerman, *supra* note 6.

may amount to nothing more than familiar eminent domain actions that facilitate the construction of highways, the creation of national parks, and so on. And if “expropriation” can at times encompass changes in regulatory policy, tax policy, etc., it becomes even more apparent that “expropriation” is not necessarily objectionable.

Accordingly, IIAs do not prohibit expropriation in general, but instead require compensation. Typical IIAs provide that no party may “expropriate” or “nationalize” an investment covered by the treaty except in a manner that complies with four conditions: (1) it must be done for a “public purpose”; (2) it must be “non-discriminatory”; (3) it must be done in accordance with due process of law and the minimum standard of treatment; and (4) must be accompanied by proper compensation, measured roughly by the fair market value of what has been expropriated. Thus, there is a prohibition on expropriation *except* when conditions (1)–(3) are met, in which case it is permissible but compensable.

Thorny issues arise in implementing such constraints on expropriation, however, because compensation for changes in government policy that impair the value of investments may well be undesirable, whether those policies entail physical takings, changes in taxation, regulatory changes, or something else. It is well known, for example, that compensating property owners for government takings of property may create inefficient incentives to develop the property ahead of a possible taking—in essence, a moral hazard.⁷⁸ The same point may be made about compensation for other kinds of changes in policy. The simplest version of the argument rests on an assumption that changes in government policy only occur if they are socially efficient. Then, it is desirable for investors to take account of how such changes may impair their investments, and to invest less in response. Compensation leads the investors to behave as if returns are certain and thus to overinvestment.⁷⁹

Of course, if one abandons the assumption that governments always behave efficiently regardless of the law governing compensation for “expropriation,” other margins are in play.⁸⁰ If government officials can realize benefits from expropriative behavior but externalize the costs to investors, they will expropriate too often and a compensation requirement may be constructive.⁸¹ More generally, government officials may see neither the full benefits nor the full costs of the policy changes that they enact. The effect of a compensation requirement on the efficiency of government choices is then entirely unclear. This observation begets fears over the possibility that compensation requirements may “chill” socially desirable policy changes. The issues become still more complex if one contemplates compensation requirements that do not always entail full fair market value, and that instead allow the amount of compensation to vary over a range from zero to fair market value or even higher.⁸² In short, the efficiency issues associated with “expropriation” and attendant compensation requirements are a morass.

Against this confusing backdrop, the compensation requirement under IIAs in practice turns on the scope of the term “expropriation.” A critical question then arises as to whether the actions deemed to constitute expropriation involve government behavior for which a compensation requirement is plausibly constructive. Section III will discuss this issue in

⁷⁸ See Blume & Rubinfeld, *supra* note 5.

⁷⁹ See Kaplow, *supra* note 5.

⁸⁰ For an extended discussion, see STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW, ch. 5 (2004).

⁸¹ See RICHARD EPSTEIN, TAKINGS: PRIVATE PROPERTY UNDER THE POWER OF EMINENT DOMAIN (1985).

⁸² See Horn and Tangerås, *supra* note 2, for a formal model along these lines.

more depth, and how the concept of expropriation has evolved in ISDS litigation and in recent treaty modifications.

Putting those issues aside for the moment, however, one can interpret constraints on expropriation under IIAs as relating in part to the inefficiencies associated with international externalities. A probability of expropriation following the establishment of an investment then offers a way to extract surplus from foreign investors that will not be passed along in full in the cost of imported capital. The result can be a tendency to engage in expropriative behavior too often from a global efficiency standpoint.

Sensible constraints on expropriation also map nicely into the risk reduction objective of IIAs. Expropriation claims arise after sunk investments are taken physically by government action or reduced in value substantially by other changes in government policy. They involve paradigm examples of the kind of conduct that investors may fear because of the time inconsistency problem. They also relate to host country conduct that may be difficult to predict, and constraints on expropriation may enable host countries to signal that the risk of expropriation is low under conditions of asymmetric information.

This analysis again suggests a possible rationale for treating foreign investors “better” than domestic investors. The case for compensating “expropriation” is stronger, other things being equal, when government officials neglect to take the interests of adversely affected investors into account. A recurring theme in this Article is that foreign investors often lack the political efficacy and influence of domestic investors, raising the possibility that a compensation requirement limited to, or more favorable toward, foreign investors may have some policy logic.

E. Monetary Transfers

One of the narrower obligations found in most IIAs restricts host countries from interfering with monetary transfers relating to investments. The UNCTAD mapping project reports that 99 percent of the included IIAs contain some provision on the transfer of funds.

As with the obligations discussed above, these principles have a ready interpretation in relation to both the international externality and risk reduction functions of IIAs. Section I noted that restrictions on monetary transfers could enable host countries effectively to tax foreign investors and capture surplus from foreign suppliers of specialized capital. Likewise, in the absence of rules against such practices, investors would need to price into their required returns the risk associated with the possibility that these practices might emerge at some future date. Host countries that internalize such costs will benefit from promises not to engage in them as long as the practices are not a source of joint efficiencies.

In that regard, it is difficult to imagine a general efficiency rationale for restrictions on monetary transfers, although in certain special instances they might have some sound basis. Perhaps transfers constitute fraudulent conveyance in bankruptcy, for example, facilitate money laundering or assist in the evasion of court judgments. Capital transfer restrictions may also promote macroeconomic stability in financial crises. Accordingly, the prohibition on monetary transfer restrictions is often accompanied by some exceptions relating to the non-discriminatory and good faith enforcement of measures such as bankruptcy laws, criminal laws, and those relating to compliance with judicial proceedings,⁸³ as well as balance of

⁸³ NAFTA, *supra* note 71, Art. 1109(4); U.S. Model BIT, *supra* note 71, Art. 7(4).

payments matters. But absent an applicable exception, transfer restrictions are likely to be inefficient and other mechanisms for exploiting host country monopsony are likely to dominate.

III. CURRENT ISSUES IN IIA INTERPRETATION AND EVOLUTION

Section II offered a broad overview of the central substantive obligations found in the great majority of IIAs and suggested how they map into the international externality and risk reduction rationales for IIAs. This section proceeds to address some prominent controversies that have arisen in ISDS litigation, and how treaty language has evolved over time in response. The goal is to highlight intriguing issues and suggest how the economic account of the rationale for IIAs may aid our thinking about them. Parts of the discussion raise more questions than answers and suggest directions for further research.

A. “Jurisdictional” Issues

Before an arbitral tribunal can adjudicate an investor-state dispute, the tribunal must have “jurisdiction” over the claim. Jurisdiction rests on a foundation of “consent” on the part of both parties to the dispute. By the time that an investor wishes to file a dispute, however, the host state may have little interest in consenting to jurisdiction, especially if the investor’s claim appears meritorious. Accordingly, states consent to the jurisdiction of arbitral tribunals in advance, before a dispute arises. Advance consent to jurisdiction can be conferred by a contract between the investor and the host state. The focus here is on IIAs, however, which typically contain a general consent to arbitral jurisdiction over disputes between “investors” and a party to an IIA (a host state). Questions can arise regarding who qualifies as an “investor” and what counts as an act of the host state. Likewise, consent to jurisdiction is limited to disputes concerning “investments,” a concept that raises many interpretive issues.

1. *Investors and Host States*

Perhaps the most common controversy over the complainant’s qualification as an “investor” concerns the use of corporate shells to gain the protection of an IIA. In *Tokios Tokelés v. Ukraine*,⁸⁴ for example, a Ukrainian publisher nominally “owned” by a Lithuanian company brought a claim under the Lithuania-Ukraine BIT. The tribunal accepted jurisdiction, even though the Lithuanian company was apparently 99 percent owned by Ukrainian nationals. *Saluka Investments v. Czech Republic*⁸⁵ involved a dispute between a Czech bank and the Czech government under the Netherlands-Czech Republic BIT. The Czech bank had been purchased by a Japanese financial firm and its shares transferred to Saluka, a Netherlands corporation, created for the sole purpose of holding those shares in the Netherlands to gain the protection of the BIT. Once again, the tribunal accepted jurisdiction.

In both of these cases, the decision rested on the definition of “investor” in the BIT, which simply required a duly constituted entity in a counterparty state. The *Saluka* tribunal further

⁸⁴ ICSID Case No. ARB/02/18, Decision on Jurisdiction (Apr. 29, 2004), paras. 21–70.

⁸⁵ Partial Award, at 46–50 (UNCITRAL Mar. 17, 2006).

noted that the Czech government was on notice of the arrangement with Saluka when the investment was made.

The outcome in *Tokelés* seems clearly at odds with the theory of IIAs developed herein. Domestic investors gained the protection of an IIA simply by setting up a corporate shell in a treaty counterparty to hold their domestic investments. An extension of IIA commitments to domestic investors cannot be squared with an account that emphasizes international externalities and the extraction of surplus from foreign investors, or the need to protect foreign investors from inefficient risk to lower the cost of imported capital.

Saluka presents somewhat more subtle issues. It allows an investor from a country (Japan) that is not a party to a BIT with the host country (the Czech Republic) to gain the protection of a BIT between the host country and a third country (the Netherlands) by creating a corporate shell in that third country to hold the investment. If one emphasizes the risk-reduction function of IIAs, this outcome is not subject to any great objection, as it is unclear why Dutch investors and Japanese investors are not both in need of protection from inefficient risks, or why the Czech Republic would not benefit from extending such protection to investors from both countries to the extent that it internalizes the costs of such risks.

The international externalities account of IIAs and its focus on distortions associated with monopsony, however, suggests a potential concern with the outcome in *Saluka*—is it problematic for Japanese investors to “free ride” on the BIT between the Netherlands and the Czech Republic? If Czech investors would value market access commitments from Japan or other treaty commitments that constrain the ability of Japan to extract surplus by virtue of a monopsony position, the ability of Japanese investors to free ride might undermine the incentive for Japan to make reciprocal commitments to Czech investors.

It is instructive to consider the response to the *Saluka* issue in other treaties. The 2012 U.S. Model BIT offers one illustration in Article 17(2), which provides:

A Party may deny the benefits of this Treaty to an investor of the other Party that is an enterprise of such other Party and to investments of that investor if the enterprise has no substantial business activities in the territory of the other Party and persons of a non-Party, or of the denying Party, own or control the enterprise.⁸⁶

These “denial of benefits” clauses, in one form or another, are found in slightly less than 10 percent of IIAs covered by the UNCTAD mapping project and are ubiquitous in IIAs to which the United States is a party. Similar language is also found in CETA.⁸⁷

Pursuant to such a clause, when the investor is a corporate shell controlled by investors from either a third country or the host country, the host country can deny the investor the benefits of the IIA. The potential exclusion of benefits to investors from third countries suggests that at least some treaty negotiators are concerned about free riding and wish to curtail it. Further, the fact that denial of benefits provisions are routine in U.S. IIAs and found in CETA fits nicely with the fact that U.S. agreements and CETA contain pre-establishment national

⁸⁶ See also NAFTA, *supra* note 71, Art. 1113(2). The NAFTA treatment of third-country investors is the same, but the Model BIT added the reference to host country investors.

⁸⁷ EU-Canada Comprehensive Economic Trade Agreement, Art. 8.16, OJ L 11/23 (Jan. 14, 2017) [hereinafter CETA].

treatment obligations. As explained above, the free rider issue is more concerning when the objectives of the IIA include market access.

In addition to who qualifies as an “investor,” arbitrators must also at times confront the question of what measures and actors trigger the responsibility of the “host state.” Does it matter what type of activity the entity performs—must it be a “governmental function” in some sense? Is the extent of government ownership or control the key issue? What if the government owns, say, 20 percent of the outstanding stock in an otherwise private company?

One leading decision, *Salini v. Morocco*,⁸⁸ involved a dispute between a foreign construction firm and a highway construction authority constituted as a public-private partnership in the host state. The tribunal held⁸⁹ that the test for whether the host country is responsible for the entity in question involves consideration of two factors: a structural factor (ownership and control), and a functional factor (whether its activity is governmental), both issues resolved in the claimant’s favor in that case.

A focus on state ownership or control is a natural one. Like other aspects of international law, IIAs concern measures taken by the parties to such agreements, and do not seek to regulate private sector activity involving non-parties. The ability of the state to influence the measures at issue in litigation is appropriately central to the inquiry, and the degree of state ownership and control is thus crucial.

The further requirement in *Salini* that the entity be engaged in a government “function,” however, is open to some question. Consider a state-owned airline or telecommunications company, for example, that favors domestically owned producers of input products, or reneges on commitments to foreign investors. Such state-owned enterprises may not be engaged in a governmental “function.” The intuition behind *Salini*’s exclusion of such actors from IIA coverage perhaps rests on the notion that when governments undertake activities that are commonly in the private sector, they should be subject to no greater legal disciplines than private sector actors performing the same activities. The potential flaw in this view is that when governments control decisions, whether they relate to a “government function” or not, they are subject to interest group pressures that are less likely to afflict decision making by private firms, and more likely to be contaminated by parochial considerations that discount the well-being of foreigners.

In this regard, the test for government responsibility under the U.S. Model BIT is different from that in *Salini*, albeit a bit muddled. Article 2(2) provides that a party’s obligation shall apply “to a state enterprise or other person when it exercises any regulatory, administrative, or other governmental authority delegated to it by that Party.” A “state enterprise” is defined as “an enterprise owned, or controlled through ownership interests, by a Party.” The upshot is that authority “delegated” by a government to a state-controlled enterprise or “other person” is within the scope of obligations if it involves “regulatory, administrative or governmental” action. Although ambiguities remain, the test here seems at least somewhat broader than the “governmental function” requirement in *Salini*.

⁸⁸ ICSID Case No. ARB/00/4, Decision on Jurisdiction (July 31, 2001).

⁸⁹ See *id.*, paras. 28–35. In this instance, the decision was a merits ruling rather than a jurisdictional ruling, as the claimant had properly filed against the state of Morocco.

2. *Investments*

Virtually all economic transactions can be said to entail a degree of “investment.” The simplest sale of goods transaction, for example, often entails an investment of time and energy in the search for an appropriate buyer or seller. Transactions governed by contracts require investment in creating the terms of the contract. Many transactions will also involve various relationship specific investments if an aspect of the transaction requires customization. The concept of an “investment” under IIAs is considerably narrower than these observations imply, however, and the question of how an optimal IIA will define covered “investments” is complex.

a. Why Are Some Transactions Not “Investments?”

Consider an exporter located in Country A selling to buyers located in Country B, both countries parties to an IIA. Suppose further that Country B has made a commitment to Country A in a trade agreement to regulate imported products in accordance with the national treatment principle (a commitment found in most trade agreements—see, for example, General Agreement on Tariffs and Trade (GATT) Article III). Finally, suppose that Country B violates this national treatment commitment and imposes discriminatory regulations on Country A’s exporter. What remedy should the aggrieved exporter in Country A have?

The trade agreement between A and B likely affords Country A an opportunity to bring the matter to dispute resolution. Trade agreements rely on state-to-state dispute settlement and not ISDS, however, and the exporter will have no private right of action. But what if the IIA between countries A and B allows ISDS and also contains a national treatment obligation—is the exporter then entitled to a monetary remedy through ISDS? The answer would be “yes” if the exporter in Country A qualified as an “investor” holding an “investment” in Country B, but almost certainly it will not. In one fashion or another, the IIA will define terms to exclude the exporter’s claim.

Consider how the issue is addressed in CETA Article 8.1:

Investment means every kind of asset that an investor owns or controls, directly or indirectly, that has the characteristics of an investment, which includes a certain duration and other characteristics such as the commitment of capital or other resources, the expectation of gain or profit, or the assumption of risk. Forms that an investment may take include . . . claims to money or claims to performance under a contract. For greater certainty, claims to money does not include . . . claims to money that arise solely from commercial contracts for the sale of goods or services by a natural person or enterprise in the territory of a Party to a natural person or enterprise in the territory of the other Party.

A footnote to Article 1 of the U.S. Model BIT makes a similar point:

Some forms of debt, such as bonds, debentures, and long-term notes, are more likely to have the characteristics of an investment, while other forms of debt, such as claims to payment that are immediately due and result from the sale of goods or services, are less likely to have such characteristics.

With reference to the hypothetical exporter above from Country A, any loss of profits on future transactions falls outside the definition of “investment,” as future transactions are not an “asset” that the investor “owns.” And to the degree that the national treatment violation impairs the value of existing contractual arrangements, these are no more than “claims to money” on cross-border sales transactions, excluded from the investment definitions.

Why exclude the exporter from coverage and limit it to the state-to-state dispute settlement available under the relevant trade agreement? One answer is that importing countries have no interest in keeping exporters happy *per se*. Instead, they make market access “concessions” benefiting foreign exporters, foregoing the opportunity to exploit monopsony power, for the purpose of obtaining reciprocal concessions from trading partners. Their interest lies in keeping counterparty governments happy so that the reciprocal concessions are not withdrawn. State-to-state dispute settlement gives counterparties a right to enforce obligations whenever they care about them, and additional private rights of action for exporters create greater litigation exposure unnecessarily.⁹⁰

In contrast, IIAs exist in considerable part (or in some cases primarily) to curtail inefficient risks for investors that capital-importing countries internalize through the cost of imported capital. When a capital-importing country bears these costs, it has an interest in affording private rights of action to reduce inefficient risk and put investors at ease as argued in Section I.

In addition, as argued in Section I, the inefficient risks that concern foreign investors arise from the existence of sunk capital investments. Sunk investments make investors more vulnerable to opportunism associated with time inconsistency issues, more vulnerable to arbitrary actions by biased agents, and so on. It follows that the value of IIA protections for an “investor” will tend to greater, the greater the extent to which the investor has sunk capital at risk.

This line of thinking helps to explain why short-term contract claims, such as claims for payment of money “immediately due” under a contract for the sale of goods and services, may not qualify as protected “investments” even though the transaction occurs entirely within the host country. Although parties to such arrangements may have incurred some sunk costs, the magnitude is likely to be modest, and the absence of a lengthy time period between performance and payment of the typical contract for sale will diminish the risk of changes in host country policies that would impair value.

At the other end of the continuum, the paradigm examples of protected “investments” under IIAs commonly embody substantial sunk costs that can only be recouped over an extended period of time, and that are more vulnerable to changes in host country policies. The creation of an enterprise in the host country, investments in real estate, contracts to manage in-country enterprises, investments in regulatory licensing, and investments in the creation of intellectual property,⁹¹ commonly put substantial capital at risk for extended periods.

In between these relatively “easy cases” lie innumerable kinds of assets that expose asset holders to important risks, but that are usually less vulnerable to host country opportunism than, for example, large in-country investments of physical capital. A complete treatment is

⁹⁰ See Sykes, *Public Versus Private Enforcement*, *supra* note 7; Ossa, Staiger & Sykes, *supra* note 44.

⁹¹ See, for example, the lists of covered “investments in CPTPP, Article 9.1 and CETA, Article 8.1.”

beyond the scope of this discussion, and we now turn to consideration of two examples—portfolio investment and sovereign debt instruments.

b. Portfolio Investment

Portfolio investment is typically defined as investment that does not entail any degree of control over the operations of the enterprise in which the funds are invested. It is often quite liquid, in the sense that ready markets often exist for the sale of portfolio assets, and the risks of portfolio investment are generally diversifiable. Portfolio investment itself often entails little in the way of sunk costs (perhaps a modest brokerage commission), although the underlying assets of firms held by portfolio investors may involve considerable sunk investments.

Should IIAs cover portfolio investment? All IIAs to which the United States is party cover it, along with CETA, and CPTPP. According to the UNCTAD mapping project, however, a small number of IIAs (around 1 percent) exclude it.⁹²

The argument for including portfolio investment within the scope of an IIA seems obvious. Many enterprises that are vulnerable to inefficient risks will be owned in substantial part by or owe debt to portfolio investors. To exclude all compensation to such investors for the materialization of inefficient risks would defeat the goal of reducing the cost of imported capital. Yet, there are some difficult issues in the background associated with portfolio investment.

The first concerns the question of who holds the right to seek compensation for breach of an IIA—the enterprise or its owners? Surely it cannot be both, lest the result be a double recovery. This issue is best handled by including portfolio investment within the definition of protected “investments,” however, and using damages principles to prevent double recovery if both the enterprise and its owners/creditors are in a position to bring claims.

A harder issue concerns strategic changes in ownership structure to take advantage of IIAs. Consider an enterprise that is initially owned and financed entirely by domestic nationals in the host country. In response to some measure that impairs a commitment under an IIA, suppose that the domestic owners of the enterprise sell some of their equity to foreign investors whose nationality brings them within the IIA. This arrangement raises the same general concern as the *Tokelés* decision noted earlier, involving the creation of a foreign shell corporation to hold a domestic investment, effectively extending the protection of an IIA to domestic nationals of the host country. Alternatively, suppose that after some measure that implicates IIA obligations, foreign owners of the enterprise not entitled to the protection of any IIA sell some equity to investors in a country covered by an IIA. Here, the effect is analogous to the effects of the restructuring done in *Saluka*, discussed earlier.

A complete exclusion of portfolio investment from the coverage of IIAs, however, seems an excessive response to such problems. If efficient financing of a venture requires portfolio investment, its exclusion from coverage could easily increase cost of capital and swamp any benefits from addressing strategic equity transactions. A better solution would be to police such transactions through other legal doctrines. As an illustration, following the enactment of tobacco plain packaging legislation in Australia, Phillip Morris transferred its interests in its Australian operations to a Hong Kong entity, which allowed it to challenge the plain packaging measures under the Hong Kong–Australia BIT. The tribunal dismissed the case,

⁹² See the compendium available at <http://investmentpolicyhub.unctad.org/IIA/mappedContent#iaInnerMenu>.

however, ruling that a transaction entered solely for the purpose of taking advantage of the protection of an IIA after a dispute has arisen (or was clearly foreseen) constitutes an abuse of rights under customary law that deprives the tribunal of jurisdiction.⁹³

c. Sovereign Debt

Consider the more controversial case of sovereign debt, a matter on which there is more variation in treaty coverage and much commentary.⁹⁴ The definition of “investment” in the U.S. Model BIT is broad enough to encompass sovereign bond issues (see item c in the illustrative list of Article 1). NAFTA, in contrast, presents a somewhat different picture. Its definition of investment includes “debt securities” in an “enterprise” or “loans” to an “enterprise,” but excludes debt securities and loans to a “state enterprise.” More generally, the UNCTAD mapping project reports that some IIAs explicitly exclude coverage of sovereign debt.⁹⁵ The new CPTPP includes sovereign debt within the definition of investment but provides in Annex 9-G that mere default is not a breach of the IIA—the claimant must show that non-payment violates an independent obligation such as national treatment, most-favored-nation, the minimum standard of treatment, or the expropriation provisions. It further provides that negotiated restructurings are not actionable unless they violate national treatment or most-favored-nation commitments. Against this variation in treaty language, a number of arbitral tribunals have wrestled with purported ambiguity over the coverage of sovereign debt in particular IIAs, and have reached different conclusions.⁹⁶

Variation in the treatment of sovereign debt instruments is not terribly surprising, as several considerations suggest that including them under IIAs may be unnecessary. First, investors in sovereign debt instruments have not made substantial investments in physical facilities or other assets vulnerable to hold up opportunism. Second, sovereign debt instruments are not subject to many of the regulatory and tax policy risks that can impair returns to investments by business enterprises. Third, sovereign debt is usually held by a wide range of investors, domestic and foreign, and it is harder for debtor states to engage in behavior that targets foreign investors in particular. Fourth, sovereign debt issuers are commonly engaged in the issuance of new debt, or the refinancing of old debt, and their need to return to the capital markets gives them a strong reputational interest in avoiding misbehavior. Fifth, the greatest risk associated with sovereign debt instruments—default—typically flows from macroeconomic crises, and the likelihood of such crises seems unlikely to be reduced much by IIA commitments. In other words, such commitments likely do little or nothing to reduce any “inefficient risk” of sovereign default. Finally, and rather trivially, it is not clear that money judgments pursuant to IIA commitments are worth much against governments that are unable to pay their debts.

⁹³ Phillip Morris Asia Ltd v. Australia, PCA Case No. 2012-12, Award on Jurisdiction and Admissibility, at 166–85 (UNCITRAL Dec. 17, 2015).

⁹⁴ See, e.g., Michael Waibel, *Opening Pandora's Box: Sovereign Bonds in International Arbitration*, 101 AJIL 711 (2007); Pietro Ortolani, *Are Bondholders Investors? Sovereign Debt and Investment Arbitration After Poštová*, 30 LEIDEN J. INT'L L. 383 (2017).

⁹⁵ See UNCTAD, *supra* note 92.

⁹⁶ See *Poštová Banka, a.s. and Istrokapital SE v. Greece*, ICSID Case No. ARB/13/08, Award, paras. 248–350 (Apr. 9, 2015) (sovereign debt not covered); *Abaclat v. Argentina*, ICSID Case No. ARB/07/5, Decision on Jurisdiction and Admissibility, paras. 352–57 (Aug. 4, 2011) (public securities covered).

All that said, one can imagine some arguments for including sovereign debt under IIAs as well. If the debt instrument is issued by a subsidiary government or state enterprise, for example, IIA protection can place the central government in the position of guarantor. In situations of imperfect information, a central government guarantee may be of considerable value to investors who are uncertain about the creditworthiness of the subsidiary government entity, although contracts may suffice for this purpose when central government guarantees are desirable. Another possibility is that in the absence of IIA protection, debtor states might discriminate against foreign debt holders in the event of a solvency crisis. This concern perhaps explains the “watered down” coverage of sovereign debt default under the CPTPP, wherein default alone is not covered, only default in conjunction with the violation of another commitment (such as national treatment).

d. The “Salini Factors”

Among the more controversial jurisdictional decisions is *Salini v. Morocco*,⁹⁷ involving unpaid money on a construction contract between a Moroccan entity and Italian construction firms. The applicable BIT defined “investment” to include “rights to any contractual benefit having an economic value” and “any right of an economic nature conferred by law or contract.” The tribunal concurred that the contract rights at issue in the case were included within this definition, but further determined that the definition of “investment” within the BIT alone was not dispositive of its jurisdiction. Instead, jurisdiction must also be appropriate in relation to the terms of the International Centre for Settlement of Investment Disputes (ICSID) convention, which provides: “The jurisdiction of [ICSID] shall extend to any legal dispute arising directly out of an investment . . . which the parties to the dispute consent in writing to submit”

Acknowledging that the Convention does not define the term “investment,” the tribunal nevertheless concluded based on earlier academic commentary that an “investment” has certain inherent characteristics including “(asset) contributions, a certain duration of performance of the contract and a participation in the risks of the transaction.” It also pointed to the preamble of the ICSID Convention, which recites that ICSID was established “(c)onsidering the need for international cooperation and development, and the role of private international investment therein,” as a basis for concluding that an “investment” for purposes of ICSID jurisdiction must make a “contribution to the economic development of the host state.”⁹⁸ Although the tribunal accepted jurisdiction over the matter at issue in the case, these four “*Salini* factors” (asset contribution, duration, risk, and contribution to development) have been invoked by other tribunals to scrutinize investor claims and deny jurisdiction even if the matter at issue falls within the definition of “investment” in the underlying IIA.⁹⁹

⁹⁷ ICSID Case No. ARB/00/4, Decision on Jurisdiction (July 23, 2001).

⁹⁸ *Id.*, paras. 50–58.

⁹⁹ See *Mitchell v. Congo*, ICSID Case No. ARB/99/7, Decision on the Application for Annulment of the Award (Feb. 9, 2004) (not shown that activity of claimant’s law firm contributed to development of the host state); *Malaysian Historical Salvors v. Malaysia*, ICSID Case No. ARB/05/10, Decision on Jurisdiction (May 17, 2007) (alleged breach of marine salvage contract not actionable because insufficiently related to development).

Non-ICSID tribunals have also occasionally suggested that the term “investment” has an objective meaning apart from its definition in the IIA that limits arbitral jurisdiction.¹⁰⁰

The first three *Salini* factors emerged in academic commentary about earlier disputes in which garden variety claims over unpaid sales contracts were put forward for “investment” arbitration, including a case where the secretary general of ICSID refused to register such a matter for arbitration.¹⁰¹ Such issues are resolved in more modern IIAs by tighter definitions of “investment” that exclude them from coverage, such as the definition set out above in the U.S. Model BIT. Accordingly, the issues submitted to arbitration by investors in recent years tend to satisfy easily the first three factors.

The fourth *Salini* factor—contribution to development—is more problematic. All economic activity that contributes to the national income of the host country contributes in some sense to “development.” Likewise, any measure that eliminates inefficient risk and lowers the cost of imported capital contributes to national income. The line between economic activity that contributes to development and (lawful¹⁰²) economic activity that does not is thus a murky one if it exists at all. The qualifications of investment arbitrators to assess the importance of an activity to “development” may also be seriously questioned. It is much better, one might argue, to let the parties to IIAs write definitions to exclude matters that they do not wish to be subject to arbitration.

It is no surprise, therefore, and a welcome development, that serious attention to the fourth *Salini* factor is waning. One tribunal has even gone so far as to opine: “[T]he criterion of contribution to economic development has been discredited and has not been adopted recently by any tribunal. It is generally considered that this criterion is unworkable owing to its subjective nature.”¹⁰³

3. Umbrella Clauses

As the preceding sections make clear, jurisdiction requires an “investor” who makes a claim in relation to a covered “investment.” The claim must also involve some substantive obligation embodied in the IIA that has allegedly been violated. Accordingly, if the host state breaches a contractual obligation to the investor, the tribunal will have no jurisdiction unless the IIA somehow makes such a breach into a violation of a treaty obligation. Some contractual breaches, such as breach of a contractual most-favored-nation commitment, might also breach a standard treaty obligation. But many other types of contractual obligations do not touch on treaty obligations in the absence of an applicable “umbrella clause.”

According to the UNCTAD mapping project, umbrella clauses appear in over 40 percent of the surveyed IIAs. Their wording varies considerably. As one example, consider the clause at issue in *Noble Ventures v. Romania*¹⁰⁴ from the U.S.–Romania BIT: “Each party shall observe

¹⁰⁰ See *Romak S.A. v. Uzbekistan*, PCA Case No. AA280, Award, paras. 180, 207 (UNCITRAL Nov. 26, 2009). An extensive discussion of *Salini* and its progeny may be found in KRISTA NADAKAVUKAREN SCHEFER, INTERNATIONAL INVESTMENT LAW: TEXT, CASES AND MATERIALS 84–112 (2016).

¹⁰¹ See *Salini*, *supra* note 88, para. 52.

¹⁰² Unlawful investments are generally excluded from coverage under IIAs. See SCHEFER, *supra* note 100, at 125–34.

¹⁰³ *Deutsche Bank v. Sri Lanka*, ICSID Case ARB/09/2, Award, paras. 306–07 (Oct. 31, 2012).

¹⁰⁴ ICSID ARB/01/11, Award, paras. 46–62 (Oct. 12, 2005). Contrast *SGS v. Pakistan*, ICSID Case No. ARB/01/13, Decision of the Tribunal on Objections to Jurisdiction (Aug. 6, 2003).

any obligation it may have entered into with respect to investments.” The Tribunal held that this language converts contractual obligations into treaty obligations and afforded the Tribunal jurisdiction over contract claims.

There are substantial pros and cons associated with the inclusion of umbrella clauses in IIAs. On the one hand, breach of contract by a host state can impose heavy losses on investors with large sunk investments. Domestic remedies may be inadequate due to delay, corruption, or bias. The risk of a major breach of contract, with the remedy limited to the domestic legal system, may be fully as worrisome to the investor as concerns about expropriation, discrimination, restrictions on monetary transfers, or other measures that may be explicitly addressed by the IIA. An umbrella clause then serves as a “gap filler” to protect investors from serious opportunism that the IIA might not otherwise cover.

On the other hand, breaches of contract may involve relatively minor issues. If the investor can take any sort of breach to international arbitration, the investor may be able to impose substantial litigation costs (and perhaps reputational costs) on the host state that are out of proportion to the significance of the breach. This situation may provide the investor with bargaining leverage to extract large settlements from weak claims and impose costs on host states that considerably exceeds the value to them in reducing the cost of imported capital.

It is unsurprising, therefore, that the use of umbrella clauses is controversial and that treaty signatories exhibit great variability in their receptiveness to including them. A possible intermediate strategy has emerged in some more recent IIAs, whereby the IIA specifies that certain “investment agreements” are subject to arbitration. The CPTPP, for example, allows investors to go to arbitration following alleged breach of certain contracts with the host state on which the investor “relies” in making a covered investment, and that relate to the development of natural resources controlled by the host state, public utility services and infrastructure construction.¹⁰⁵ Such a “positive list” approach to the coverage of contractual breaches allows treaty counterparties to elevate contractual obligations to treaty obligations only in cases where doing so appears most important to investor security.

B. Non-discrimination Rules

As suggested in Section II, national treatment and most-favored-nation obligations are ubiquitous in IIAs. These non-discrimination principles lie at the core of the commitments that curtail international externalities, and violations of these principles will often suggest host country behavior that creates inefficient risks. But the application of these obligations in practice raises some important interpretive questions. This section considers a few of the issues that have arisen.

1. The Meaning of Discrimination and “Like Circumstances”

The most fundamental task in the application of any non-discrimination principle is the determination whether a complainant and the putative comparator are in fact “comparable,” and that observed differences in treatment can meaningfully be called “discrimination.” In international trade agreements, for example, the question is often whether foreign producers

¹⁰⁵ See Comprehensive and Progressive Agreement for Trans-Pacific Partnership, Arts. 1 (“investment agreement” defined), 9.19 (*entry into force* Dec. 30, 2018) [hereinafter CPTPP].

of “like products” are being treated less favorably.¹⁰⁶ Under IIAs, a common phrasing asks whether foreign investors in “like circumstances”¹⁰⁷ or “like situations”¹⁰⁸ are being treated less favorably. A related issue concerns the rationale for differential treatment—is the law concerned with the intent of a measure, or only its effect? Three important arbitral decisions regarding the national treatment obligation highlight the challenges associated with these issues.

*GAMI Investments v. Mexico*¹⁰⁹ was a NAFTA dispute concerning the expropriation of sugar mills by the Mexican government. The sugar market was depressed, and Mexico determined that it would expropriate half the mills in the country to reduce production capacity and promote market recovery. The expropriations were indeed followed by a recovery in the sugar market, but the benefits inured only to the mills that had remained in private hands. GAMI was an American company that held a minority interest (14 percent) in a company that owned a group of five Mexican sugar mills, all of which were expropriated. The national treatment claim in the case centered on the fact that a purportedly comparable company, owning the same number of mills and having no foreign shareholders, had none of its mills expropriated. The arbitral tribunal could not point to any compelling evidence that the comparator mills were financially healthier or otherwise better positioned to remain in business than the expropriated mills, but it nevertheless ruled for Mexico. It found that GAMI failed to prove that the challenged expropriations “resulted from or have any connection to GAMI’s [ownership] participation,” and that “a reason exists for the measure that is not itself discriminatory.”¹¹⁰ GAMI failed to prove that its circumstances “were demonstrably so ‘like’ those of non-expropriated mill owners that it was wrong to treat [it] differently.”¹¹¹

The situation in *GAMI* presents a difficult problem. One can interpret the decision as resting on the general principle that the complainant has the burden of proof to establish a violation and could not show that the challenged measures were undertaken *because of* the investor’s nationality, and of course Mexico asserted that they were not. But the circumstances were at least suspicious. And given the lack of extensive discovery tools in ISDS litigation, it may be unrealistic to expect foreign investors to provide clear evidence that they were singled out because of their nationality. In part for this reason, dispute tribunals in other contexts (such as the WTO) focus not on the *intent* behind a measure alleged to produce discrimination, but on its *effect*. The factual record as disclosed in the *GAMI* award does not allow one to assess whether the overall effect of Mexico’s expropriation program fell disproportionately on foreigners.

A sole focus on effect without any concern for intent, however, has its own downside. Even if a measure has a disproportionate effect on foreign investors, that effect may be coincidental, and may not implicate the sort of inefficient targeting of foreigners that IIAs seek to police. And as the *GAMI* tribunal also noted, there is a cost to opening the door too widely to national treatment claims whenever a company harmed by a measure has some degree of

¹⁰⁶ See General Agreement on Tariffs and Trade, Art. III(2), (4), Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, 1867 UNTS 187, 33 ILM 1153 (Apr. 15, 1994).

¹⁰⁷ See, e.g., CPTPP, *supra* note 104, Art. 9.4(1).

¹⁰⁸ See, e.g., CETA, *supra* note 87, Art. 8.6(1).

¹⁰⁹ Final Award (UNCITRAL Nov. 15, 2004).

¹¹⁰ *Id.*, paras. 114–15.

¹¹¹ *Id.*, para. 114.

foreign ownership: “It is not conceivable that a Mexican corporation becomes entitled to the anti-discrimination protections of international law by virtue of the sole fact that a foreigner buys a share of it.”¹¹² The case clearly raises the concern noted earlier about strategic corporate transactions designed to gain the protection of an IIA for domestic investors.

The general lesson from cases like *GAMI* is that violations of non-discrimination principles are often non-transparent, and host countries will invariably claim that their actions rest on legitimate differences among investors. As with trade law, the law of IIAs can be expected to evolve away from requiring proof of discriminatory intent, but that simply shifts the battleground to the issue of “like circumstances.” The host country will argue that differential treatment is not based on nationality but on some other factor that makes the complainant “unlike” the putative comparator. Close cases will surely arise, often turning on the burden of proof, and the extent of foreign ownership in claimant companies may prove an important factor.

Subsequent litigation relating to the Mexican sugar industry concerned a tax on high fructose corn syrup (HFCS) products such as carbonated beverages. *Cargill v. Mexico* is illustrative.¹¹³ The HFCS tax was imposed following a NAFTA dispute over Mexican access to the U.S. sugar market, which the United States refused to arbitrate. In the hope of coercing the United States to change its position, Mexico imposed the tax on HFCS products, many imported from the United States and distributed by Mexican subsidiaries of U.S. companies. The tax exempted sugar-containing products, the bulk of which were made in Mexico. U.S. investors brought claims under the investment chapter of NAFTA alleging, inter alia, that the HFCS tax violated the national treatment obligation under NAFTA. The claim required the investors to argue that as producers of HFCS-containing beverages, for example, they were in “like circumstances” to the Mexican producers of sugar-containing beverages. The tribunals found that such beverages were in a competitive relationship and, indeed, the evident purpose of the tax was to punish the United States by disadvantaging HFCS products. Mexico’s argument that the industries were not in “like circumstances” because the Mexican sugar industry was depressed was rejected: “Mexico did not claim that it took the measure simply to allow sugar producers to capture the sweetener market for soft drinks. Its rationale for the measure was to bring pressure on the United States government to live up to its NAFTA obligations.”¹¹⁴

The decision in *Cargill*, insofar as it rests on the competitive relationship between the products of the two groups of investors, seems sound. The tax was both designed and intended to injure the foreign competitors of domestic investors, singling them out for disadvantageous treatment and curtailing their profits from Mexican investments.

The tribunal’s distinction between measures intended to punish the United States and measures intended to bolster a depressed local industry, however, is questionable. The panel hints that if Mexico taken its action out of a purely protectionist desire to assist a local industry facing hard times, it might have found an absence of “like circumstances.” But such a view offers a license for investment protectionism, whether pre- or post-establishment, and would threaten to undermine both the international externality and risk reduction

¹¹² *Id.*, para. 115.

¹¹³ See, e.g., *Cargill, Inc. v. United Mexican States*, ICSID Case No. ARB(AF)/05/2, Award (Sept. 18, 2009).

¹¹⁴ *Id.*, paras. 208–10.

rationales for the national treatment principle. If treaty counterparties wish to create an exception for measures to protect depressed industries, it makes far more sense to do so expressly (as is done by the safeguards provisions of GATT Article XIX).

As a further illustration of the issues that arise in interpreting “like circumstances,” consider *Occidental Exploration and Production Co. v. Ecuador*,¹¹⁵ a case brought under the U.S.–Ecuador BIT. Occidental developed oilfield properties in Ecuador and exported petroleum products. After many years of litigating the issue in Ecuadoran courts, it was ultimately denied the opportunity to obtain value added tax (VAT) refunds for taxes paid on locally sourced and imported goods used in the manufacture of goods for exportation. Among other claims, Occidental argued that domestically owned Ecuadoran companies, including those involved in producing floral products, mining products, and seafood, were granted VAT refunds in relation to purchases of goods used to manufacture exports. Ecuador argued in response that all oil producers, including its national oil company PETROECUADOR, were treated equally for VAT refund purposes, and that the companies in the other industries identified by Occidental were not “like” Occidental. The tribunal ruled for Occidental on the national treatment claim, however, in effect holding that the obligation to treat foreign and domestic investors equally applied to the entire class of firms engaged in export trade.¹¹⁶

The economic issues in *Occidental* are also difficult. The U.S.–Ecuador BIT included pre-establishment national treatment. But more favorable tax rules for domestic firms engaged in floral, mining, and seafood production do not plausibly injure oil producers given the lack of substitutability between floral, mining, and seafood products on the one hand, and petroleum products on the other. And the fact that Occidental and PETROECUADOR were treated equally seems to rule out any issue of investment protectionism within the oil industry.

Yet, Occidental’s initial investments were made at a time when Ecuadoran tax authorities granted VAT refunds to oil exporters, and only later did they reverse themselves on the existence of authority for such refunds under Ecuadoran law. Oilfield exploration and development plainly entails large sunk investments, and a substantial and unfavorable change in tax laws subsequent to such investments does raise concerns about a manifestation of the time inconsistency problem. Perhaps this worry was implicitly behind the tribunal’s willingness to stretch the concept of “likeness” so broadly.

The national treatment obligation is a crude instrument for addressing this concern in *Occidental*, however, because the broad conception of “likeness” necessary to rule for the claimant has collateral effects, constraining policy differentiation across industries in situations where there is little or no concern about the impairment of substantial sunk investments or tipping the scales on market access. A better approach to policing changes in tax and other laws that harm established foreign investors may be to focus on the fair and equitable treatment principle (discussed in Section C below), asking whether policy changes frustrate expectations on which investors reasonably relied in making their investments. It is noteworthy that the *Occidental* tribunal also found a breach of the fair and equitable treatment obligation on that basis, and thus might have ruled for Occidental without any need to stretch the national treatment concept of “likeness.”

¹¹⁵ LCIA Case No. UN3467, Final Award (UNCITRAL July 1, 2004).

¹¹⁶ *Id.*, paras. 167–79.

2. *Substance Versus Procedure and the Most-Favored-Nation Obligation*

One of the most controversial ISDS decisions is *Maffezini v. Spain*,¹¹⁷ brought under the Spain–Argentina BIT. The claimant invested in chemical production in Spain and ended up in a dispute that alleged violations of the BIT. But the Spain–Argentina BIT required the investor to litigate in Spanish courts for eighteen months before turning to international arbitration. Another BIT between Chile and Spain had no such requirement, however, and the Spain–Argentina BIT had a most-favored-nation obligation.¹¹⁸ Maffezini claimed that by virtue of the most-favored-nation commitment, he was entitled to go to international arbitration immediately, and the tribunal ruled in his favor. Other tribunals have reached the opposite conclusion,¹¹⁹ although to be sure the precise wording of the most-favored-nation clauses at issue differ somewhat among the various cases.

The core issue in *Maffezini* is whether the most-favored-nation obligation should extend to procedural aspects of ISDS litigation or is limited to substantive policies of the host country that affect foreign investors. One can make arguments either way based on the economic considerations developed earlier.

Dispute settlement procedures are important to investors, and differences in the options available to foreign investors from different source countries can affect their competitive position in relation to each other. Where the objective of the IIA includes maintaining competitive equality among foreign investors of varying nationalities in the face of international externalities that affect their relative position, uneven treatment of competing investors in the dispute settlement process presents a potential problem.

But this problem should not be exaggerated. The number of ISDS cases in a given industry and a given host country is surely small—most investments do not end up in litigation, let alone litigation under IIAs. The competitive edge for investors from a particular source country created by their access to a more advantageous ISDS dispute settlement process seems likely to be small in light of the low probability of ISDS *ex ante*.

Moreover, if one focuses on the risk reduction function of most-favored-nation commitments, it is difficult to argue that holding foreign investors to the dispute settlement provisions of a treaty negotiated by their home governments constitutes an inefficient risk. It is not the sort of behavior that might represent a manifestation of the time inconsistency problem or the agency cost concerns developed in Section I. Likewise, it does not reflect problems of asymmetric information or host country deception and is not a source of risk for risk-averse investors. Thus, the exclusion of procedural matters from the coverage of the most-favored-nation obligation does not seem to diminish the risk reduction function of the most-favored-nation obligation in any way.

¹¹⁷ ICSID Case No. ARB/97/7, Award (Nov. 13, 2000).

¹¹⁸ The most-favored-nation clause appears in conjunction with the fair and equitable treatment obligation in the Spain–Argentina BIT and reads as follows: “In all matters governed by this Agreement, such (fair and equitable) treatment shall not be less favorable than that accorded by each Party to the investments made in its territory by investors of a third country.” Article 4 (2). Note the absence of any reference to “like circumstances,” as in NAFTA and the U.S. Model BIT.

¹¹⁹ See *Plama v. Bulgaria*, ICSID Case No. ARB/03/24, Decision on Jurisdiction (Feb. 8, 2005). On the heterogeneity of most-favored-nation clauses generally see Simon Batifort & J. Benton Heath, *The New Debate on the Interpretation of Most-Favored-Nation Clauses in Investment Treaties: Putting the Brakes on Multilateralization*, 111 AJIL 873 (2017).

If this analysis is right, one might predict a drift away from the *Maffezini* outcome in treaty practice. That drift is apparent in the U.S. Model BIT, which requires “treatment no less favorable” in “like circumstances.” Under this phrasing of the most-favored-nation obligation, and depending on the scope of the “likeness” concept, a complainant might have to identify another existing foreign investor in a competitive relationship with the claimant who had received more favorable treatment in the form of access to superior dispute resolution options. Such a showing might not be possible much of the time.

To eliminate any ambiguity, however, some newer IIAs eliminate the possibility of procedural claims based on the most-favored-nation principle. The new CPTPP, for example, adds the following provision to its most-favored-nation obligation in Article 9.5(3): “For greater certainty, the treatment referred to in this Article does not encompass international dispute resolution procedures or mechanisms” It thus forecloses any possibility for a claimant to use the most-favored-nation commitment to open the door to alternative dispute resolution procedures that may be found in another IIA involving a CPTPP member. CETA contains language to the same effect in Article 8.7(4). The analysis here suggests that this drift in treaty practice is not surprising.

C. *Minimum Standard/Fair and Equitable Treatment*

The UNCTAD mapping project reports that over 90 percent of surveyed IIAs (including all involving the United States) contain some version of a minimum standard of treatment/fair and equitable treatment obligation as discussed in Section II. The exact phrasing of the obligation varies, but it inevitably exhibits considerable vagueness. That fact likely explains why some IIAs do not include these obligations at all,¹²⁰ and surely has much to do with the extensive litigation over their precise content through the years. A comprehensive survey of the issues that have arisen in litigation is far beyond the scope of this Article, and we limit the discussion to two general sets of developments. Both relate to the discussion in Section I regarding a key reason why ISDS may fail to promote the joint interests of treaty counterparties—inefficient underlying substantive obligations, or obligations so vague that arbitrators cannot reliably distinguish efficient and inefficient behavior.

The first issue concerns the scope of arbitral discretion. Does a “fair and equitable treatment” obligation empower arbitral tribunals to impose their own conceptions of what is “fair” and “equitable?” Parties to IIAs quickly became wary of giving arbitrators such authority, sparked in part by some early NAFTA decisions in which tribunals were perceived to have concocted excessively pro-investor standards.¹²¹ The result was a treaty “interpretation” issued by the three-party NAFTA Commission which stated, *inter alia*: “The concepts of ‘fair and equitable treatment’ and ‘full protection and security’ do not require treatment in addition to or beyond that which is required by the customary international law minimum standard of treatment of aliens.”¹²² Article 5(2) of the U.S. Model BIT incorporates this clarification and goes on to provide that the minimum standards of treatment “do not create additional substantive rights” that go beyond the customary law requirements for the treatment of

¹²⁰ See, for example, the Australia-China BIT.

¹²¹ See especially *Metalclad v. Mexico*, ICSID Case No. ARB(AF)/97/1, Award (Aug. 30, 2000); Pope & Talbot v. Canada, Award on the Merits of Phase 2 (UNCITRAL Apr. 10, 2001).

¹²² See http://www.sice.oas.org/tpd/nafta/Commission/CH11understanding_e.asp.

aliens.” UNCTAD reports that nearly 15–20 percent of the surveyed IIAs have some language along these lines, tying the fair and equitable treatment concept to principles of international law such as the customary minimum standard of treatment. Purely open-ended fair and equitable treatment standards, without some additional language to contain their substantive sweep, nevertheless survive in about 80 percent of surveyed IIAs.

Even when fair and equitable treatment is anchored in customary international law, however, the content of customary law is hardly non-controversial. International law scholars, who are the source of judgments about what has become “customary,” often disagree. A consensus exists that basic due process protections and the benefits of police protection are part of the customary obligations to aliens, but what else?

This query leads to the second set of developments to be discussed, involving an important strand of case law that focuses on measures by the host country that frustrate an investor’s “reasonable expectations.” Properly circumscribed, this interpretation of the fair and equitable treatment principle can respond effectively to a key source of inefficient risk noted in Section I—the possibility that host countries can exploit conditions of asymmetric information to mislead investors about future contingencies. But which expectations are “reasonable” and deserve protection? Surely not all changes in policy that disadvantage investors can be deemed to frustrate “reasonable” expectations. Investors would not expect that new information about hazards, for example, will not result in a properly measured regulatory response. And as noted in Section II, it is undesirable in general for investors to be insured against harm due to changes in government policy. Investors will be led to overinvest in reliance on the status quo, when optimal investment incentives require them to take full account of future contingencies that may diminish the value of an investment or render it worthless. As long as governments do not mislead investors about the likelihood of policy change, or conceal information that they could easily disclose, there is little basis for awarding compensation when investors are disappointed by a change in policy that violates no other treaty obligation.

A number of well-known ISDS cases concerning fair and equitable treatment do involve allegations of government deceit or at least careless misrepresentation. An early case was *Metalclad v. Mexico*.¹²³ A U.S. investor in a Mexican landfill project was confronted with environmental approval issues at multiple levels of government. According to the investor, in an account that the tribunal largely accepted, various officials misled or misinformed the investor about the environmental approval requirements that would apply to the project. The investor was led to believe that it had the necessary approvals in hand, resulting in a large capital investment that was completely lost after local environmental objections blocked the project at the local level and the area was eventually turned into a protected cactus reserve.

More recently, *Bilcon v. Canada*¹²⁴ concerned the development of a quarry in Canada. The investor argued, and the tribunal broadly accepted, that Canadian officials had encouraged substantial initial expenditures on the project and provided optimistic information about the regulatory approval process. When environmental opposition developed after the project was well underway, opponents of the project orchestrated an environmental review to block the project that deviated from standard practice under Canadian law, did not afford the investor a

¹²³ Note 118, *supra*.

¹²⁴ PCA Case No. 2009-04, Award on Jurisdiction and Liability (UNCITRAL Mar. 17, 2015).

fair hearing, and failed to provide the investor with notice as to the standards that would be used to adjudicate the regulatory issues.

The facts in these cases are disputed, and the point here is not to endorse the claimants' account of them or the tribunals' factual conclusions. But *if* one accepts the claimants' narratives, they suggest clear examples of situations in which foreign investors, disadvantaged in dealing with government officials and lacking the political influence that domestic investors often possess, were deliberately or carelessly misled by officials in the host country to overestimate considerably the likelihood of a successful undertaking, and in turn led to make substantial irreversible expenditures in reliance. The host countries were penalized by the tribunals not for undertaking reasonable environmental regulation, but for deceiving investors about the regulatory process. Such deception is akin to a denial of due process and is a source of inefficient risk that IIAs can usefully eliminate.

Such decisions have nevertheless proven controversial, and various commentators have suggested that the fair and equitable treatment principle may interfere with the ability of host states to undertake sensible regulation or create "regulatory chill."¹²⁵ Concern for this problem has sparked further changes in treaty practice.

The CPTPP, for example, adds the following paragraphs to its minimum standard of treatment in Article 9.6:¹²⁶

For greater certainty, the mere fact that a Party takes or fails to take an action that may be inconsistent with an investor's expectations does not constitute a breach of this Article, even if there is loss or damage to the covered investment as a result.

Hence, a "mere" affront to "expectations" of the investor is not actionable in the CPTPP, but the door is left open to claims that the host country has misled investors and induced undue reliance on circumstances that do not materialize.

A further innovation involves the inclusion of a list of possible ways that a host country may breach a fair and equitable treatment obligation. UNCTAD reports that over one hundred surveyed IIAs now contain such an approach. CETA, for example, provides in Article 8.10:

A Party breaches the obligation of fair and equitable treatment referenced in paragraph 1 if a measure or series of measures constitutes:

- (a) denial of justice in criminal, civil or administrative proceedings;
- (b) fundamental breach of due process, including a fundamental breach of transparency, in judicial and administrative proceedings;
- (c) manifest arbitrariness;
- (d) targeted discrimination on manifestly wrongful grounds, such as gender, race or religious belief;
- (e) abusive treatment of investors, such as coercion, duress and harassment

¹²⁵ See, e.g., Been & Beauvais, *supra* note 4; BONNITCHA, *supra* note 2, § 3.7; UNCTAD, Fair and Equitable Treatment: A Sequel (2012), available at https://unctad.org/en/Docs/unctaddiaaia2011d5_en.pdf (discussing various treaty language options for protecting the "right to regulate").

¹²⁶ See also CETA, *supra* note 87, Art. 8.9(2).

Thus, if an investor is misled by the host country to his detriment, it must frame its claim as “denial of justice,” “fundamental breach of due process,” “manifest arbitrariness,” or “abusive treatment,” all of which suggest a fairly high standard for the investor to meet. But Article 8.10 goes on to say:

When applying the above fair and equitable treatment obligation, the Tribunal may take into account whether a Party made a specific representation to an investor to induce a covered investment, that created a legitimate expectation, and upon which the investor relied in deciding to make or maintain the covered investment, but that the Party subsequently frustrated.

The approach in CETA accords broadly with the account of fair and equitable treatment suggested herein for cases where investors are disappointed by changes in government policy. It offers no protection against policy evolution per se but does afford protection for cases in which the host country possesses private information and deliberately or carelessly misleads foreign investors about future contingencies.

In sum, despite controversies over fair and equitable treatment and related “minimum standards,” no inherent reason exists why these obligations should undermine legitimate regulatory measures or chill socially constructive policy evolution. Basic protections in accordance with settled customary law, coupled with a remedy for situations in which investors have reasonably relied on improper representations by host country officials, do the necessary work to protect investors while allowing regulators reasonable policy space. Treaty practice has already moved significantly toward an embodiment of these principles. Any “regulatory chill” associated with fair and equitable treatment and the early decisions in relation to it can be abated as treaties and tribunals embrace this evolution and regulators come to understand its implications.

D. Expropriation

Investor claims involving host country regulatory policies often focus on the fair and equitable treatment obligation, but a number of claimants have asserted that regulatory changes amount to “expropriation.” We conclude with a brief discussion of the expropriation issue under IIAs, with special attention to cases involving regulatory matters.

The seizure of investor property by governments may be termed “direct expropriation,” which is typically compensable. IIAs almost always address the phenomenon of “indirect expropriation” as well (over 95 percent of surveyed IIAs according to UNCTAD), although the test for what constitutes “indirect expropriation” is open to controversy. Several factors are relevant.¹²⁷ The first factor is the extent to which the value of the investor’s assets have been impaired. The investment must be largely wiped out according to many decisions—policy that simply reduces the value of an investment, even by quite a substantial percentage, is unlikely to be deemed an expropriation.¹²⁸ If the host country government receives no direct economic benefit in the form of property or monetary gain, expropriation is also less likely to

¹²⁷ See L. Yves Fortier & Steven L. Drymer, *Indirect Expropriation and the Law of International Investment: I Know It When I See It, or Caveat Investor*, 19 ICSID REV. FOR. INV. L.J. 293 (2004).

¹²⁸ See, e.g., *Glamis Gold v. United States*, Award (UNCITRAL June 8, 2009) (reduction in value from \$49 million to \$20 million insufficient for expropriation).

be found.¹²⁹ The “legitimate expectations” of the investor are also a consideration here as under the fair and equitable treatment analysis, as well as the “public purpose” behind the purported expropriation.¹³⁰

All of these factors tend to weigh against the investor who makes a claim of “expropriation” by regulation. Regulation typically imposes costs on investors that stop well short of wiping out the value of an investment. Regulation does not entail a transfer of any property or money to the government. As long as the regulatory objective seems legitimate and reasonable, and the government has not misled the investor about its intentions to induce an investment, a proper public purpose exists, and the investor will typically be said to have no basis for an expectation that regulatory policy would not change. Accordingly, modern tribunals commonly suggest that non-discriminatory regulation for a public purpose, taken in accordance with due process, is not indirect expropriation and likely not compensable on any other theory.¹³¹

Treaty language has moved in this direction as well. Annex 9-B of the CPTPP, derived from the U.S. Model BIT, provides: “Except in rare circumstances, non-discriminatory regulatory actions by a Party that are designed and applied to protect legitimate public welfare objectives, such as public health and safety and the environment, do not constitute indirect expropriations, except in rare circumstances.”

As suggested in Section II, the economic wisdom of these principles is open to some question. Even in cases of direct expropriation, it is well known that compensation may lead to inefficiencies, such as overinvestment in improvements on real property. If expropriation is socially efficient, landowners should curtail investment in improvements to take account of the likelihood that land may be expropriated for an alternative use.

Likewise, an absence of compensation for changes in regulatory policy is not *necessarily* efficient. Horn and Tangerås,¹³² for example, study a simple formal model in which a host country may introduce regulation in response to new information about a production externality. Production is undertaken by foreign-owned firms, and the host country does not take the loss of profit for those firms into account when deciding whether to regulate (a variant of the basic international externality considered in Section I.A). The result is overregulation against a global efficiency benchmark, as well as underinvestment from a global standpoint because investors anticipate that their investments will at times be inefficiently impaired by host country regulation. A compensation requirement induces the host country to “internalize” the lost profit to foreign investors and eliminates the over-regulation problem.¹³³

Models such as that of Horn and Tangerås do not allow for undesirable “regulatory chill” associated with compensation requirements, however, because host country governments are assumed to maximize their national welfare. When they internalize externalities to foreigners

¹²⁹ See, e.g., *SD Myers v. Canada*, Partial Award (UNCITRAL Nov. 13, 2000).

¹³⁰ Fortier & Drymer, *supra* note 127.

¹³¹ See, e.g., *Phillip Morris v. Uruguay*, ICSID Case No. ARB/10/7, Award (July 8, 2016); *Methanex v. United States*, Final Award of the Tribunal on Jurisdiction and Merits (UNCITRAL Aug. 3, 2005).

¹³² Note 2, *supra*.

¹³³ Full compensation would lead to overinvestment (due to moral hazard stemming from “insurance” against regulation), however, and so an optimal compensation mechanism would deliver compensation only under certain contingencies (ideally, only when regulation is globally inefficient). If that condition is not verifiable, various second-best options emerge in their model. See also Aisbett & Bonnitcha, *supra* note 2 (distinguishing regulatory changes attributable to time inconsistency issues from those driven by “new information”).

they then maximize global welfare and behave efficiently. The modern critique of ISDS discussed in Section I.B, however, implicitly rejects such assumptions about government decision making in favor of a belief that compensation requirements, particularly in the area of regulatory policy, will discourage governments from regulating when they should.

The critics seem to believe that the benefits of regulation inure to the citizenry at large, while the costs of a compensation requirement fall on governments that are already politically constrained in their ability to impose new taxes, and that may face electoral revolt if they try. Implicitly, taxpayers are unwilling to pay more even to facilitate policies that yield them net benefits. The general absence of compensation for regulatory takings under domestic law in many developed countries further reinforces the notion that compensation is a mistake, and that a broader compensation requirement under IIAs would unduly favor foreigners over domestic investors. Whatever the soundness of this view, it clearly predominates in the current political environment and both arbitral tribunals and modern treaty practice embrace it.

Accordingly, the critique of IIAs (and ISDS) based on a purported threat to sound social regulation from claims of expropriation also seems overblown, at least in relation to more recent developments with respect to the meaning of “indirect expropriation.” Whether the legal issue rests on fair and equitable treatment or indirect expropriation, the pendulum has swung heavily toward a presumption that investor losses due to changes in regulatory policies are rarely compensable absent clear discrimination, a denial of due process, or substantially misleading representations to investors by government officials.¹³⁴ As this understanding informs the thinking of regulators over time, fears of unwarranted liability exposure under IIAs should abate along with the problem of “regulatory chill.”

CONCLUSION

Treaty negotiators need a clear conception of the costs and benefits of the options before them if treaty evolution is to proceed constructively. Likewise, as much as arbitrators may at times pretend to the contrary, a purely textual analysis rarely suffices for a satisfactory interpretation of treaty obligations that are fuzzy and incomplete, and the interpretive exercise often cries out for an appreciation of the functional justification for ambiguous provisions. Careful reflection on the economic underpinnings of IIAs can importantly advance the process of treaty negotiation, drafting and interpretation to the benefit of all concerned.

To that end, this Article has developed two complementary accounts of the economic function of IIAs—the control of international externalities that involve the extraction of surplus from foreign investors with specialized capital, and the elimination of inefficient risks that cause uneconomic increases in the cost of capital to capital-importing countries. It has suggested how these twin objectives can explain the core substantive commitments found in IIAs. It also applied the economic understanding of IIAs to numerous legal issues that have become prominent in ISDS litigation, offering at times an explanation and at times a critique of the jurisprudence, as well as some observations about the rationale for trends in treaty obligations over time.

¹³⁴ This analysis accords with the defense of ISDS in the regulatory arena put forward most forcefully by Charles Brower. See, e.g., Charles N. Brower & Sadie Blanchard, *What's in a Meme? The Truth About Investor-State Arbitration: Why It Need, Not, and Must Not, Be Repossessed by States*, 52 COLUM. J. TRANSNAT'L. L. 689 (2014) (arguing that ISDS does not impair regulatory sovereignty).