Multilateral Investment Agreement in a Political Equilibrium*

Dieter Urban†

October 6, 2006

Abstract

This study provides a theoretical explanation, first, why some less developed countries (LDCs) have complained about the OECD negotiations of a multilateral investment agreement (MAI) in 1998 although they were free to join or opt out. Second, it explains why we observe instead an explosion of bilateral investment treaties (BITs). The explanation rests on an FDI model with three distortions: there is a time-inconsistency problem of extracting rents from FDI, there is an underprovision of public goods in LDCs, and there is a lobbying distortion in political decision making that is initially unobservable to foreign investors which causes political risk. The negotiation of MAI by a club exerts a negative information externality on non-members. A regime of BITs undermines the club agreement and unravels the information-asymmetry problem. However, an appropriately designed MAI is world-welfare superior compared to a regime of BITs by alleviating the lobbying distortion.

JEL Reference: F13, F23, O19.

Keywords: foreign direct investment, multilateral agreements, adverse selection, bilateral investment treaties.

*The author would like to thank Hartmut Egger, Martin Kolmar, Georg Tillmann†, Alessandro Turrini, Andreas Waldkirch and participants at the CESifo summer school in Venice, the EEA conference in Vienna, the ETSG conference in Vienna, and the Jahrestagung of the Verein für Socialpolitik in Bayreuth for comments. The usual disclaimer applies.

†Institute for International Economic Theory, FB03, Johannes Gutenberg-University Mainz, 55099 Mainz, Germany. Tel: ++49 (0) 6131 39 25140. Email: dieter.urban@uni-mainz.de.

1 Introduction

An important part of globalization is the activity of multinational enterprises (MNEs). The explosion of such activity in recent years had an impact not only on the efficiency of world wide production sharing but also on relative factor incomes. The resulting conflict of interest across owners of different factors may stimulate policy reactions to globalization both on a national and an international level where some governments favor liberalization and others regulation of foreign direct investment (FDI).

In fact, there was a trial to implement a multilateral agreement of investment (MAI) to liberalize FDI. In the year 1995, negotiations for a MAI were initiated by the OECD among experts and a draft proposal emerged. When the draft became known to the public, it received criticism by several interest and opinion groups within advanced countries (NGOs, trade unions) and a fierce opposition by some developing countries. Eventually, this agreement failed on an OECD summit in 1998, but there was at the same time a surprising explosion in bilateral investment treaties (BITs) throughout the 90ies.

The purpose of this study is first to explain, why some less developed countries (LDCs) have complained about the OECD negotiations although they were free to join or opt out. Second, why do we observe instead an explosion of BITs? And finally, is there a role for a multilateral investment agreement to improve world welfare if appropriately designed?

I answer these questions in a model, where FDI is driven by relative factor endowments and political risk. Political risk arises from change in government and incomplete information of MNEs on the new governments' attitude towards FDI before experiencing the new policy. Because MNEs base their location decision in this stage on some average perception, they loose from businesses in illiberal countries but gain from those in liberal ones. Hence, resources are transferred indirectly from countries that are liberal towards FDI to those that are not. This creates the incentive for liberal countries to create a multilateral investment agreement that enforces a minimum standard of FDI liberalization. Since all countries in the world are free to join or opt out, their membership choice reveals information on their type and reduces the indirect subsidy by causing investment diversion

towards members.

While this can explain the protest of some outsiders to the negotiation, it also explains why the agreement is unstable, because the same conflict of interest applies among negotiating countries. Hence, a regime will emerge where many agreements are formed at different degrees of FDI liberalization taking into account differences among countries in their FDI-related policy goals. One may view this regime as one of BITs. Still, there is some potential for a multilateral agreement to improve world welfare compared to a regime of BITs. An appropriately designed multilateral agreement bends policies of a self-interested government towards the one that optimizes host-country welfare even if countries are not members, because the threat of investment diversion disciplines the government when acting in its own interest.

My conclusions are limited in two respects. When explaining the failure of MAI in 1998, the protest of LDCs was not the only reason - and maybe not the decisive one. There was considerable disagreement among the OECD countries for a number of reasons, as well. I disregard these considerations and focus on the LDC stake in MAI negotiations.[1] Moreover, I show the emergence of a BIT regime. However, this model is not thought to be a comprehensive explanation for why many different investment treaties emerge rather than an explanation for why one multilateral treaty does not sustain.

There are few studies on multilateral investment agreements. Markusen (2001) discusses the LDC decision to accede a MAI. Commitment enhances credibility that promotes FDI. Instead, the loss of discretion when implementing MAI rules can remove a threat of host countries against MNEs and worsen its bargaining position in MNE rent sharing. Turrini and Urban (2001) take up the argument that MAI membership results in a loss of bargaining power of LDCs vis a vis MNEs. The benefit from MAI membership results, however, from signalling by accession to MAI a liberal FDI policy and thereby attracting additional investment. In addition, the endogenous membership choice gives rise to multiple equilibria where either all countries join, no MAI is formed, or some countries form a MAI. It is shown that countries which stay outside of MAI actually loose from the formation of a liberal investment club relative to a world without MAI, because FDI is diverted

[1] See Hoekman and Saggi (2000), p. 184ff, for a comprehensive list of reasons.

from them towards the club.[2] Che and Willmann (2006) focus on the dispute settlement procedure of MAI to generate a self-enforcing contract. MAI verifies any expropriation of MNE assets by a member country and triggers a coordinated withdrawal of capital by all MNEs. As a consequence, MAI can self-enforce more profitable capital flows relative to a world without MAI. The drawback is that the world interest rate rises which hurts poor countries the most.

I deviate from the previous studies by endogenizing the choice of partners with whom to form agreements and the choice of strictness of the agreement. With these extensions, I can explain why MAI negotiation failed and resulted into a scattered system of BITs. Moreover, I give a political economy microfoundation to Turrini and Urban (2001) of what determines the choice of MNE rent extraction by host countries. Contrary to Che and Willmann (2006), enforcement is exogenous and the focus is on explaining the implications of the most-favored-nation and non-discrimination principle rather than of the dispute settlement procedure (DSP).[3]

The rest of this study is organized as follows. In the next section, I introduce a baseline model of FDI, add a political equilibrium, discuss the implications of a multilateral agreement for the equilibrium, and eventually characterize the equilibrium for an exogenous MAI rule. In section 3, I endogenize the formation of negotiation groups and show how a MAI of a given negotiation group collapses and a regime of BITs assembles. In section 4, I suggest a MAI design that is world-welfare superior and causes only modest resistance by governments. A discussion of the model robustness is followed in section 5, and a conclusion is found in section 6.

[2] Our paper is also somewhat related to a theory of regional trade and investment agreements by Fernandez-Arias and Spiegel (1997). However, FDI is considered the movement of physical capital rather than the exertion of control rights across borders, and no discussion is devoted to multilateral agreements, how they form, whether they are stable, and whether they yield additional welfare gains.

[3] There is also an informal discussion on MAI by Hoekman and Saggi (2000), and Ferrarini (2003). Hoekman and Saggi (2000) see in MAI an instrument of FDI promotion that constrains countries' control over MNEs. When joining MAI, governments forgo the possibility to shift rents from MNEs to host countries in exchange for using MAI as a commitment device or signal to enhance credibility of an FDI promotion policy. In addition, FDI influences the factor income distribution and affects thereby the political equilibrium in the host country. Ferrarini (2003) extends the discussion from the OECD MAI draft to a comprehensive list of investment measures in the multilateral organization of the world economy.

2 Model Set-up

In broad terms, I describe a multi-country economy of potential host countries to Northern FDI with endogenous government policy. To be more precise, there is a mass one of host countries indexed by i. I abstract from Northern home countries to FDI, since I am interested in the LDC stake in MAI. To keep the model tractable, I set up the model such that each host country can be analyzed in isolation, since each country acts like a small open economy. This is consistent with a single country having measure zero mass in world income and no influence on world prices.

I define multinational enterprises (MNEs) as firms for which the nationality of corporate control is different from the nationality of the plant location. Typically, differences in control involve differences in technology both with respect to total factor productivity and factor intensities. FDI flows are then flows of technology that result from MNE location decisions.[4]

FDI inflows are driven by two factors: political risk and factor cost. I capture two stylized facts with these two driving forces. The first is the Lucas puzzle, i.e. why capital does not flow from rich to poor countries.[5] And the second is the observation that some LDCs such as Mexico, Eastern Europe, China, India, etc., have actually achieved such an influx of foreign capital and managed to trigger a catch-up process. Low factor cost ensure the attractiveness of LDCs for FDI, but FDI inflows are undesired by the government in some political equilibria despite its welfare benefits to the host country. In such a case governments impose expropriation risk on MNEs and thereby expel FDI. The FDI model is described in more detail in section 2.1.

I assume that factor intensities of MNEs and local firms do not coincide and factor price equalization does not hold. Hence, FDI inflows change the distribution of factor income. If there is no mechanism in the political system such that winners automatically

[4] An alternative view is to consider FDI as flow of physical capital. Che and Willmann (2006) adopt the alternative view in their analysis of MAI. Empirically, the activity of MNEs and capital flows are positively but far from perfectly correlated. Braconier et al. (2005) describes the correlation in their sample of bilateral FDI around the world.

[5] Reinhart and Rogoff (2004) has recently argued contrary to Lucas (1990) that political risk is the core cause of the lack of capital flows from rich to poor countries.

compensate losers of FDI inflows, then groups of different factor owners will have conflicting interests in attracting or inhibiting FDI. The outcome of the resulting political struggle will shape government decisions on national policies and at a multilateral negotiation. At the same time, the political decision maker has a national policy instrument - the rent extraction through taxation, expropriation, bribery, and many other informal measures that affect FDI profits - to steer the amount of FDI inflows compatible with the political equilibrium of the host country. I denote this rent extraction rate by β.

Yet, political equilibria are unstable, and governments change either as the result of an election outcome in a democracy or a coup d'etat in an authoritarian regime. New governments encompass a new composition of power among social groups which changes at the same time the government preference of how liberal or illiberal it is towards FDI. The degree of liberality is captured by a preference parameter χ which is assumed to be a uniformly distributed random variable on the range $[0,\bar{\chi}]$. I will show in section 2.2 that the preference for liberality, χ, is one determinant of the rent extraction rate, β.

However, investors learn the type of the new government mostly at a time lag by own experience and the one of their competitors. Even if they learn the type immediately, it will take time to adjust their investments. To capture this idea, I assume that MNEs are caught by surprise by a new government policy β in a first period, $t = 1$, but can re-locate after revelation of the government policy in a second period, $t = 2$. Thereafter a new policy struggle begins and a new government comes into place. However, I assume that there is no way that one government can borrow reputation to the next one. Hence, it is sufficient to analyse a game over the two periods of any one legacy alone. Both the timing of the policy action, β, and the restriction of the game to two periods establish a time-inconsistency problem. Since the new government knows that MNEs cannot react immediately on a change in policy, they are tempted to extract too much rents from MNEs which protect themselves by reducing their activities. Apart from the time-inconsistency problem, political risk involves also an asymmetric information problem, because MNEs do not know in the first period the government preferences, χ.[6]

[6]There exists an ample literature on a theory of expropriation risk of MNEs. Most of the time it is regarded as some sort of a time inconsistency problem between investor and governments. In this tradition

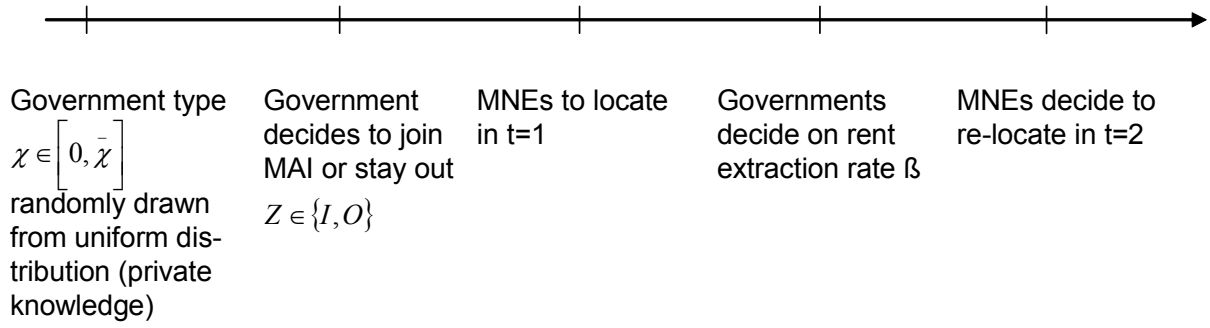

Figure 1: Sequencing of events in section 2

Finally, I think of a MAI membership decision as a credible commitment device to adhere to some given rules which are explained later in section 2.3. Hence, the accession decision, Z, to join MAI, I, or to opt out, O, will take place in the beginning of the game.[7] I can summarize the previous discussion by displaying the resulting timing of the game in Figure 1.

2.1 Modelling FDI

The only purpose of the FDI model is to obtain a relation between FDI inflows and relative factor prices. My results on explaining the protest against MAI by some LDCs and the failure of MAI do not depend on which factor actually gains from FDI inflows. A vertical FDI model of Helpman (1984) would produce such a result, but turns out to be too complex to be integrated into my MAI game. Instead, I resort to a drastic simplification and refrain from many features of MNE models. For example, I define MNEs as one-plant firms with foreign ownership rather than multiplant firms. Other features of MNEs are kept. For example, MNEs are exclusive owners of a superior technology to produce a particular good. Moreover, MNEs are free to choose their location among many countries. Finally, MNEs operate with scale economies.

Each country, i, is endowed with some units of labor L_i and capital K_i. I think of

can be found Eaton and Gersovitz (1984), Raff (1992), Doyle and van Wijnbergen (1994), Thomas and Worall (1994), Schnitzer (1999) and Janeba (2002), among others.

[7] The negotiation of MAI is further endogenized in section 3 when the strictness of the MAI rule is bargained over and the group of negotiating countries is endogenously formed.

capital as internationally immobile factor such as human capital, land, or natural resources. An MNE operates with constant marginal cost. There are A units of labor necessary to produce one additional unit of output, $A < 1$. To start operation, overhead cost have to be born. For simplicity, one unit of a Cobb-Douglas composite of capital and labor is necessary, i.e. $1 = k^{0.5}l^{0.5}$, where k and l are the plant-headquarter demand of an MNE of capital and labor in a host country, respectively. I assume that all MNEs are symmetric and do not require therefore a firm index. The capital share is set to 0.5 only to obtain smooth analytical closed form solutions. However, my results will not depend qualitatively on this choice.

Local firms do not have access to modern technology. I assume as in Murphy et al. (1989) that local firms, i.e. firms owned by nationals, employ a traditional CRS technology, where one unit of labor is transformed into one unit of output. Such a (handicraft) technology is available to anyone in all product markets without market entry or exit barriers.

Consumers are split into workers and capital owners. Only workers demand MNE products. Otherwise, consumers are in all countries identical. There is a unit mass of products, of which an endogenous fraction is supplied by MNEs. For the sake of concreteness, the period-utility function U_{it} of workers in country i at time t takes the form

$$U_{it} = \int_0^1 \ln c_{it}(j)dj$$

with the consumption of good j by consumers in country i at time t, $c_{it}(j)$. The utility function of capital owners is analogous but covers only goods that are tailor-made rather than MNE mass-fabrics. For simplicity, I abstract from time preference.

Trade cost are absent. Factors are intersectorally, but not internationally mobile. Wages are chosen to be the numeraire (in one country). I assume that there is no complete specialization. Hence, product prices in sectors without MNEs will be one due to international trade and wages will be one in all countries due to perfect competition on the labor market. Product prices in sectors with MNEs will also be (marginally below) one, because MNEs will employ a limit pricing strategy. If they chose a price larger than

one, local firms would enter and the MNE would forgo strictly positive profits. Since the monopoly price is larger than one, the MNE would also give up on profit opportunities if the product price was set strictly below one. As a result, whenever an MNE exists in a product market, it will serve the entire market. Local firms will only be present in markets where no MNE exists.

Turning to optimal MNE choices, the relative factor demand will obey $l_{itZ} = r_{itZ}k_{itZ}$, where r_{itZ} is the rental rate of country i at time period t. Moreover, the index Z, $Z \in \{N, I, O\}$, indicates whether I consider a world without a MAI (N), or a country in a world with MAI that opts out of MAI (O), or a country in a world with MAI that opts in (I). Capital market clearing requires $\phi_{itZ} \cdot k_{itZ} = K_i$, where ϕ_{itZ} is the mass of sectors with MNE firms in country i - called FDI inflows for short.

The equilibrium rental rate, r_{itZ}, is found to be from relative factor demand, capital market clearing, and the production function for plant headquarter services

$$r_{itZ} = \phi_{itZ}^2 K_i^{-2}. \tag{1}$$

Hence, FDI inflows exert upward pressure on rental rates. While this is not directly obvious from vertical FDI theory, it corresponds with empirical stylized facts. Quite important for the perception of politicians that negotiated the MAI in the late 90ies was the experience of Ireland in the early 90ies. This country was able to trigger a spectacular catch-up by an FDI promotion strategy including a generous tax break on MNE profits repatriated from Ireland. Barry and Bradley (1996) report that inward FDI was both human-capital and physical-capital intensive. At the same time, relative high-skilled wages and real-estate rental rates rose sharply with incoming FDI flows in line with my model. Similar relative factor price changes are found in Feenstra and Hanson (1997) for Mexico, after NAFTA was founded. NAFTA contained FDI provisions quite similar to the ones in the MAI draft.[8]

Turning to the location decision of MNEs, I assume that there is free entry and exit and risk neutrality of MNEs. Then, there is entry of MNEs in host countries until

[8] Markusen (2002), p. 133, assumes also in his knowledge-capital model that all MNE activities are more high-skill intensive than the local economy consistent with my model.

expected profits are zero.[9] In period 2, the rent-extraction rate is known and the zero-profit condition is given by:

$$l_{i2Z} + r_{i2Z} \cdot k_{i2Z} = [1 - \beta_{iZ}] (1 - A) Y, \tag{2}$$

where I recall that price and wage are one, $(1 - A)$ is the operating-profit mark-up, Y demand for one MNE good, and β_{iZ} incorporates a rate of rent extraction of countries from MNE operating profits that is time-invariant in accordance with Figure 1.[10] I think of rent extraction as all kind of policy measures that are intended to reduce MNE profits. These may be both direct measures like taxation, fines, or expropriation, and indirect measures like the requirement to engage in joint ventures with local partners, relaxation of intellectual property right protection, or bribery. Rent extraction generates (some) government revenue which is spent on public goods that are produced by the CRS technology in sectors without MNEs. Still in period 2, the zero profit condition (2) implies together with (1), the relative factor demand, and the capital market clearing condition:

$$\phi_{i2Z} = 0.5 K_i [1 - \beta_{iZ}] Y \equiv \phi (\beta_{iZ}) . \tag{3}$$

FDI inflows depend on two country-specific factors - the capital endowment of a country K_i and the rent extraction rate β_{iZ}. Solving for the rental rate in period 2 finally from (1)-(3) yields:

$$r_{i2Z} = 0.25 (1 - \beta_{iZ})^2 Y^2 \equiv r (\beta_{iZ}) . \tag{4}$$

Since FDI inflows increase the capital rewards and the government can steer by choice of the FDI regime the FDI inflows, the degree of FDI liberalization appears as an indirect policy instrument to redistribute factor incomes within a country.

In period 1, the firm does not know the rent extraction rate β_{iZ}. In a world without

[9] I allow for a mixed strategy in the location decision. By the law of large numbers, the probability of one MNE to locate in country i during period t is equal to the mass of MNEs, ϕ_{it}, locating there.

[10] It would be easy to add a Northern headquarter fixed cost to render our definition of a vertical FDI firm comparable to Helpman (1984). It would obviously remain inconsequential for the model as long as this cost is constant.

MAI, a risk neutral firm simply forms the unconditional expected value on it which I denote in slight misuse of standard notation by $E[\beta_{iN}|N]$. Instead, in a world with MAI, the firm exploits the information on the MAI membership decision $M \in \{I, O\}$ in forming a conditional expectation on β_{iM} denoted by $E[\beta_{iM}|M]$. Since the zero-profit condition (2) is linear in its random variable, equations (2)-(4) hold analogously by replacing β_{iZ} with its corresponding conditional expected value.

To close the model, I assume for simplicity that world demand for MNE goods stems only from workers. Workers symmetrically distribute their income towards a mass one of products (since all prices are one), whereof a fraction Φ_{tZ}, which is equal to the aggregate of ϕ_{itZ} over all countries, is produced by MNEs. Hence, world demand that falls onto supply of any one MNE is just $Y = \int L_i di$, where I recall that wages, prices, and the total mass of goods are all one. This assumption serves to sharpen my results and to render analysis tractable. There won't be wrong conclusions drawn from my model as long as there is not expected a major increase of world income through the creation of MAI. This is plausible, since MNEs make up only a small part of economic activity even in highly developed countries.[11] It is further assumed that $\Phi_{tZ} < 1$ in equilibrium.[12]

[11] A related argument has been used in Krugman (2000) showing that international trade of the North-South type makes up an almost negligible part of OECD GDP. This is even more so true for North-South FDI despite its growth. Exact numbers are hard to come by, but the following figure may give an idea of the magnitude of FDI activity. There are a mere 2.7 million employees of US affiliates in emerging market economies and LDCs in 1996 (UNCTAD, 1999, p. 450), which is quite small compared to the US labor force.

[12] To see that the model set-up is indeed consistent, note that world CRS good supply is given by

$$\int L_i di - \left(\int L_i di\right) \Phi_{tZ} A - 0.5 \left(\int \phi_i (1 - \beta_i) di\right) (1 - A) \left(\int L_i di\right).$$

The first term is total world labor supply of which is deducted the labor employed in MNE production, i.e. labor coefficient, A, times the share of MNEs, Φ_{tZ}, times output of one MNE, $Y = \int L_i di$. From total world labor supply is also deducted the labor demand for MNE plant- headquarter services, which must be equal to fixed labor cost divided by the wage which is one. However, fixed labor cost of all MNEs is given by total fixed cost which must be equal to operating profits, $\left(\int \phi_i (1 - \beta_i) di\right) (1 - A) \left(\int L_i di\right)$, over all MNEs according to the zero-profit condition, times the labor share in fixed cost, 0.5, which is derived from the Cobb-Douglas production function with factor cost share 0.5.

Instead, the world demand for CRS goods is given by demand of workers, i.e. income share on CRS goods, $(1 - \Phi_{tZ})$, times income from labor, $\int L_i di$, plus demand of capital owners, i.e. income share one times capital income $0.5 \left(\int \phi_i (1 - \beta_i) di\right) (1 - A) \left(\int L_i di\right)$, which in turn is derived analogue to the labor cost component in fixed cost, plus public goods demand, $\left(\int \phi_i \beta_i di\right) (1 - A) \left(\int L_i di\right)$, derived from the expression for government revenue. It is straightforward to show, that demand and supply indeed equate each other. Exports of CRS goods of each country are obtained from subtracting CRS demand of any country from its supply for a given ϕ_{itZ}. Clearly, countries with a large share in ϕ_{itZ} export MNE

In the next section, I discuss what may determine the political equilibrium of factor income redistribution.

2.2 Modelling the political equilibrium

The rent extraction rate β_{iZ} is determined by governments. I assume that a government of a country maximizes a country welfare term that is biased towards some interest group. Grossman and Helpman (1994) consider a lobbying model where specific-factor owners pay campaign contributions to the government to convince her to shape government policy in the lobby-groups' interest. I assume that capital owners are lobbying for policies that increase rental rates which is consistent with the observation that LDCs tend to have an income distribution that is more unequal than in developed countries. Taking rich people as capital owners, lobbying by capital owners will produce a larger relative rental-wage rate. I will discuss this assumption further in section 5.2. Such a set up can be written in reduced form as a maximization problem of the following government objective function $W_{iZ}(\chi, \beta_{iZ})$ with respect to the rent extraction rate β_{iZ} (see Grossman and Helpman, 1996, and Ethier, 1998)[13]:

$$\begin{aligned} W_{iZ}(\chi, \beta_{iZ}) &= 2L + \chi r_{i1Z} K_i + \chi r(\beta_{iZ}) K_i \\ &\quad + u\beta_{iZ} Y \phi_{i1Z} + u\beta_{iZ} Y \phi(\beta_{iZ}). \end{aligned} \tag{5}$$

The government objective function depends on factor incomes and on tax revenue in both periods.[14] Importantly, factor income of a country rises with inflows of FDI. This provides an incentive to liberalize FDI.[15]

products and import CRS products and vice versa. Finally, labor market clearing follows from Walras' law.

[13] A study that applies a lobbying framework on policy determinants of factor prices to a small open economy with internationally mobile factors is Facchini and Willmann (2005). Grossman and Helpman (1996) analyzes trade protection in a lobbying model in the presence of MNEs.

[14] We ignore discounting for simplicity.

[15] This result is in contrast to Brecher and Diaz-Alejandro (1977) which shows that capital inflows may reduce host country welfare in the presence of a small tariff. The difference in effects arises from the different definitions of FDI. While Brecher and Diaz-Alejandro (1977) considers FDI as an inflow of physical capital, I consider FDI as an inflow of technology that rises the demand for capital rather than

Since firms locate first, then governments choose rent extraction, and finally firms re-locate, there arises a time inconsistency problem. Because MNEs cannot evade higher rent extraction in the first period, this will create an upward bias in rent extraction when compared to the welfare-optimal one. This time inconsistency problem is standard in modelling political risk.[16] Less standard is the lobbying term χ, $\chi \geq 1$, that increases the weight of capital income in the government objective function and will bias the rent extraction rate downward. Strictly speaking, the time inconsistency problem does not involve risk, since the optimal government behavior can be foreseen. I introduce political risk by assuming that MNEs have incomplete knowledge on the lobbying term χ which is assumed to be a random variable. To simplify the analysis, I assume that χ is uniformly distributed on the interval $[1, \bar{\chi}]$, where $\bar{\chi}$ is some upper bound such that $1 < 2u < \bar{\chi} < \infty$.[17] I distinguish in my notation the random variable χ from its realization of country i, χ_i. Political risk arises thus from the unpredictability of government change when there are different types which are more or less favorable to FDI liberalization. If $\chi = 1$, then equation (5) represents country welfare.

While it is widely acknowledged that there are host country benefits from FDI, LDCs claim, however, control over foreign investors. For example, Ganesan, a former Indian commerce secretary to the government, points out that it " *... becomes necessary for developing countries to employ an appropriate mix of incentives and performance requirements for FDI to achieve specific developmental objectives.*" (Ganesan, 1998, p. 5)

To build into the model a specific LDC development objective with respect to FDI, I allow for an underprovision of public goods that is typical for many developing countries. Such an underprovision may arise from the lack of an efficient tax authority, for example, in countries with corrupt bureaucracies. Moreover, there are scale economies of control

its supply. My result can easily be shown to generalize to general functional forms and different sector settings. It does not hinge on fixing the wage at one.

[16] Janeba (2002) has pointed out that there must be a mutual time inconsistency problem. Countries cannot commit to certain policies and firms cannot commit to stay in a country. Otherwise, an upfront subsidy would solve the time inconsistency problem and no international agreement was necessary. An upfront subsidy will not be an equilibrium, if a firm cannot commit to invest after receiving the subsidy. I exclude an upfront-subsidy solution to the time-inconsistency problem exogenously.

[17] The second inequality serves only to avoid cumbersome notation but does not substantially affect results.

to avoid tax evasions. Small local firms may not be taxed because fixed control cost are larger than the tax revenue from a firm if production is small.[18] For these reasons, an LDC government may have a budget constraint on public goods such that marginal utility of public goods is larger than of private goods ($u > 1$). For simplicity, I assume the marginal utility of public goods to be a constant mark-up over private marginal utility.

The optimal choice of the rent extraction rate in a world without MAI, β_{iN}, is found by inserting the period 1 analogue to equation (3):[19]

$$\beta_{iN} = \begin{cases} \beta\left(\chi, E\left[\beta_{iN}\left|N\right.\right]\right) & \text{if } \chi < \chi_{Nu} \\ 0 & \text{else} \end{cases} \tag{6}$$

where

$$\beta\left(\chi, E\left[\beta_{iN}\left|N\right.\right]\right) \equiv \underbrace{\frac{u-1}{2u-1}}_{\substack{\text{distortion from} \\ \text{underprovision} \\ \text{of public goods} \\ (+)}} \underbrace{-\left[\frac{u-1}{2u-1} - \frac{u-\chi}{2u-\chi}\right]}_{\substack{\text{lobbying} \\ \text{distortion} \\ (-)}} + \underbrace{\frac{u\left(1-E\left[\beta_{iN}\left|N\right.\right]\right)}{2u-\chi}}_{\substack{\text{time} \\ \text{inconsistency} \\ \text{distortion} \\ (+)}}, \tag{7}$$

and $\chi_{Nu} \equiv u\left(2 - E\left[\beta_{iN}\left|N\right.\right]\right)$. There is a simple interpretation of equation (7). Without any distortions, the country-welfare optimum is zero rent extraction, since factor allocation is efficient. If there is a larger marginal utility of public than of private goods, then a country is willing to accept lower factor incomes in exchange for a larger government revenue and better public good provision. Hence, some strictly positive rent extraction is optimal. Indeed, the optimal value to country-welfare optimization is $(u-1)/(2u-1) \geq 0$. Its deviation from the term $[u-\chi]/[2u-\chi]$ in equation (7) represents the distortion through lobbying. Since the capital lobbyists favor lower rent extraction to promote FDI and push up rental rates, the second distortion is negative. The third term on the right-

[18] See Burgess and Stern (1993) for this reasoning, p. 775 and p. 799.

[19] Consistent with my argument on a lack of government financing through local taxes, I assume that MNE subsidies are not feasible.

hand side is positive and represents the time-inconsistency distortion of β_{iN} from the optimum of the government. Interestingly, the optimal rent extraction rate depends only on the country characteristic χ but not on endowment differences. The following Lemma characterizes the equation (7) further.

Lemma 1 *The function* $\beta(\chi, E[\beta_{iN}|N])$ *is strictly monotonically decreasing in* χ *on the range* $1 \leq \chi < \chi_{Nu}$*; and it holds* $0 \leq \beta_{iN} < 1$ *for all countries* i*.*

Proof: See Appendix 1.

I will need these properties of β_{iN} to analyze further optimal firm behavior and the MAI entry choice of countries.

2.3 Modelling MAI

Before I continue the analysis, I define MAI. There are many provisions in the negotiated MAI agreement in the various versions from 1995 to 1998.[20] However, economically relevant seems to be how the various provisions restrict the action space of governments to extract rents from MNEs.[21] In particular, provisions like the most-favored-nation clause and the national-treatment clause reduce incentives of governments to extract rents, since agents with the weakest political support (i.e. foreigners without votes in national parliaments) are protected by those with the strongest (i.e. national firms or foreign firms supported by powerful governments). Moreover, there were provisions in the proposal that guaranteed liberalization of FDI in a broad sense.

For the sake of concreteness, I assume that MAI members agree on a common standard of maximum rent extraction B such that

$$\beta_{iI} \leq B. \tag{8}$$

Such an asymmetric rule that constrains some countries but not others is a likely outcome if an exclusive club negotiates it and is bound to agree unanimously among club members.

[20] The reference to the MAI draft is OECD (1998).

[21] A similar formalization of FDI provisions in NAFTA is found in Fernandez-Arias and Spiegel (1997).

In fact, it is easiest to agree unanimously to a rule that constrains only countries that do not participate in the negotiation. There may still be a benefit to the negotiators from this rule if the MAI causes a positive externality to the negotiators (through investment diversion).

I assume that such a rule is self-enforcing without providing a formal argument. Self-enforcement may follow from the DSP that was part of the MAI provisions and may act in a similar way as the one for the WTO.[22] Alternatively, the self-enforcement may arise from combined trade and investment accords. See Fernandez-Arias and Spiegel (1997) for a regional trade and investment agreement in a three-country world where a North-South agreement is self-enforcing, because the South punishes by taxing FDI and the North punishes by levying a tariff on imports in a model with outsourcing.

In practice, the MAI draft of 1998 included an even stronger mechanism that ensures enforceability. The MAI was planned to become legally binding by incorporating it into national law and opening the doors of national courts to MNEs to sue governments for their policies and regulatory takings that conflict with the MAI provisions.[23]

2.4 Defining equilibrium of subgame with exogenous MAI rule

I am now ready to define a perfect Bayesian equilibrium of the stages over MAI membership choice, optimal rent extraction by governments, and firm location in both periods. Such an equilibrium will be given if (i) MNEs enter or exit a country with a probability ϕ_{i2Z} until the zero-expected-profit condition holds conditional on knowing the government type χ and its MAI choice, Z, in period 2; (ii) governments maximize their objective function conditional on their membership choice, Z, and on expectation formation of MNEs; (iii) MNEs apply Bayes' rule in forming beliefs on the government type whenever possible; (iv) MNEs enter or exit a country with probability ϕ_{i1Z} until the zero-expected-profit condition holds in period 1 based on their beliefs on government types and the observed MAI choice, Z; (v) governments optimize their objective function when deciding on MAI membership

[22] Bagwell and Staiger (1999) shows how the dispute settlement procedure of GATT is self-enforcing.

[23] Ethier (1998) takes the enforcement issue also as exogenous in a theory of regionalism and provides additional reasons that apply in our case, too.

given other governments' actions, firm expectation formation, and their own policy of rent extraction after membership choice.

Moreover, I define a partial MAI equilibrium of this game at a given MAI rule (8) as an equilibrium where some countries decide to enter MAI and others opt out. Instead, a complete MAI is an equilibrium where all countries join MAI. I am mainly interested in the analysis of partial MAI.

Optimal firm behavior is already discussed in section 2.1 apart from MNEs' expectation formation. To characterize expectation formation, however, the optimal government choices have to be analyzed. The countries that opt into MAI choose the optimal rent extraction rate in analogy to the case without MAI in equation (6) by $\beta_{iI} = \beta\left(\chi, E\left[\beta_{iI}\left|I\right.\right]\right)$ whenever the MAI rule (8) is not binding ($\beta\left(\chi, E\left[\beta_{iI}\left|I\right.\right]\right) \leq B$). The case, when the rule is binding will be discussed in section 2.6. The optimal rent extraction rate choice of a country i when opting out is given by $\beta_{iO} = \beta\left(\chi, E\left[\beta_{iO}\left|O\right.\right]\right)$. The expectation on rent extraction depends on how countries decide to enter MAI or opt out once a particular MAI exists. This is analyzed in the next section.

2.5 Self-selection of countries

The decision to enter a MAI of given strictness B is made based on the government objective function under the two alternative choices. First, the government objective function when entering MAI shall be denoted $W_{iI}\left(\chi, \beta_{iI}\right)$ and the objective function when opting out by $W_{iO}\left(\chi, \beta_{iO}\right)$. By the Nash-conjecture, opponent actions are taken as given and thus ϕ_{i1Z} is exogenous to the MAI membership choice. Under this condition, I find a self-sorting order in Lemma 2.

Lemma 2 *A country i will prefer to opt out (in) if $\chi_i < \chi^*$ ($\chi_i > \chi^*$), given a threshold value χ^*, $\chi^* > 1$, such that $W_{iI}\left(\chi^*, \beta_{iI}\right) = W_{iO}\left(\chi^*, \beta_{iO}\right)$. Formally, $W_{iI}\left(\chi_i, \beta_{iI}\right) \lesseqgtr W_{iO}\left(\chi_i, \beta_{iO}\right)$ if and only if $\chi_i \lesseqgtr \chi^*$.*

Proof: See Appendix 2.

Intuitively, each country faces a trade-off. When increasing the rent extraction rate, the government revenue increases and FDI inflows decrease which reduces total factor income. The larger is the weight χ_i the smaller is chosen a β_{iZ} to find the optimal trade-off. When comparing the government objective functions of two countries i and i' with marginally different lobbying terms χ_i and $\chi_{i'}$ with $\chi_i > \chi_{i'}$, then i will choose a marginally lower rent extraction rate β. Since the β-choice of both countries is optimal for them, the impact of changing β optimally when changing χ has only a second-order effect on the objective function by the envelope theorem. The only remaining first-order effect is the direct effect of a change in the valuation χ of capital income. Capital income in MAI is, however, larger than outside MAI, because rent extraction inside is restricted while it is not outside. Hence, MNEs expect lower rent extraction from members, and devote more FDI there which causes a country to have a higher rental rate when inside MAI than when outside. Finally, the larger rental income benefits more those countries with larger valuation of capital income. Hence, MAI membership is the more desired the larger is the valuation term χ.

I can show that this mechanism holds for general functional forms and different ways of modelling FDI as long as MAI consists of a maximum rent extraction rule for members, and is credible. Yet, this self-sorting order will be a crucial mechanics for this model.

2.6 Forming expectations on rent extraction

When deciding in period 1 on the MNE location, firms observe the decision of countries to join MAI or opt out. Hence, they form expectations on the rent extraction rate conditional on this information.

Before I can form expectations, I need to characterize how countries i determine their rent extraction rate and membership decision in the presence of a MAI rule (8) when it is binding ($\beta(\chi_i, E[\beta_{iI}|I]) > B$). For this purpose, I define a threshold $\tilde{\chi}$ such that $\beta\left(\tilde{\chi}, E[\beta_{iI}|I]\right) \equiv B$, and obtain the next Lemma.

Lemma 3 *(i) There is a group of countries i with $\chi^* < \chi_i < \tilde{\chi}$ such that these coun-*

tries enter MAI, although the MAI rule (8) is binding if a χ^*, $\chi^* > 1$, *exists such that* $W_{iI}(\chi^*, \beta_{iI}) = W_{iO}(\chi^*, \beta_{iO})$.
(ii) For a country i^* *with a corresponding* $\chi_{i^*} = \chi^*$, $\chi^* > 1$, *such that* $W_{iI}(\chi^*, \beta_{iI}) = W_{iO}(\chi^*, \beta_{iO})$, *there must be* $\beta(\chi^*, E[\beta_{iO}|O]) > B$.

Proof: See Appendix 3.

Summing up the insights on the optimal rent extraction rate from Lemmas 2 and 3, I can describe the country choice of the rent extraction rate β_{iM} in a world with MAI under the assumption that country i^* with lobbying parameter $\chi_{i^*} = \chi^*$ is indifferent of its choice to join MAI or opt out as:

$$\begin{aligned} \beta_{iM} &= \begin{cases} \beta(\chi_i, E[\beta_{iO}|O]) & \text{if } \chi_i \leq \chi^* \\ B & \text{if } \chi^* \leq \chi_i < \tilde{\chi} \\ \beta(\chi_i, E[\beta_{iI}|O]) & \text{if } \tilde{\chi} \leq \chi_i < \chi_{Iu} \\ 0 & \text{else} \end{cases} \\ &\equiv \beta_M(\chi_i) \end{aligned} \tag{9}$$

where $\chi_{Iu} = u(2 - E[\beta_{iI}|I])$. First of all, we observe that β_{iM} does neither depend on world income Y nor on any country factor endowments. Moreover, it is only through χ_i depending on i. It is straight forward to see that

$$\frac{\partial \beta_M(\chi)}{\partial \chi} \leq 0 \qquad \text{for } \chi \neq \chi^* \tag{10}$$

analogous to Lemma 1. The larger is the weight of the government objective function on capital income the less inclined is this government to extract rents from foreign firms, reduce the FDI inflows, and suffer from lower capital income.

I assume that there exists a country i^* with a corresponding $\chi_{i^*} = \chi^*$ such that this country is indifferent of entering MAI or opting out. Then, the conditional expectation of the rent extraction rate β dependent on the MAI decision is given as follows:

$$E\left[\beta_{iO}\left|O\right.\right]=\frac{\int_{1}^{\chi^{*}}\beta_{M}\left(\chi\right)d\chi}{\chi^{*}-1}, \tag{11}$$

where I apply Lemma 2, the assumption on a uniform distribution of the lobbying term χ, and equation (9). Fortunately, a closed form solution can be obtained for the expected value $E\left[\beta_{iO}\left|O\right.\right]$ which is given in Appendix 4. It depends only on one endogenous variable - the partition of countries χ^{*} - and must be falling in it, because the expansion of χ^{*} leads to an addition of countries outside MAI with a rent extraction rate lower than the average among outsiders.

The expected value of the rent extraction rate inside MAI is a bit more complicated to determine. It is given from Lemmas 2 and 3 as:

$$E\left[\beta_{iI}\left|I\right.\right]=\frac{\int_{\tilde{\chi}}^{\chi_{u}}\beta_{M}\left(\chi\right)d\chi+\left[\tilde{\chi}-\chi^{*}\right]B}{\bar{\chi}-\chi^{*}}. \tag{12}$$

Also in this case, I can obtain a closed form solution which is given again in Appendix 4. Accordingly, the expected value $E\left[\beta_{iI}\left|I\right.\right]$ depends only on the endogenous variable χ^{*}, and falls in it for the same reason as before.

Finally, I obtain the optimal choice of the rent extraction rate when the indifferent country i^{*} is opting out:

$$\beta_{i^{*}O}=1-u\frac{E\left[\beta_{iO}\left|O\right.\right]}{2u-\chi^{*}}\equiv\beta_{O}\left(\chi^{*}\right). \tag{13}$$

Also this rent extraction rate depends only on one endogenous variable, χ^{*}. Concerning the exogenous variables, the expected values of the rent extraction rates depend on the marginal utility of public goods u, the strictness of MAI B, and the distribution parameter $\bar{\chi}$. They do not depend on world income or factor endowments.

From equations (9), (10), (11) and (12), I can infer the following order of rent extraction rates

$$E\left[\beta_{iI}\left|I\right.\right]\leq B<\beta_{O}\left(\chi^{*}\right)\leq E\left[\beta_{iO}\left|O\right.\right] \tag{14}$$

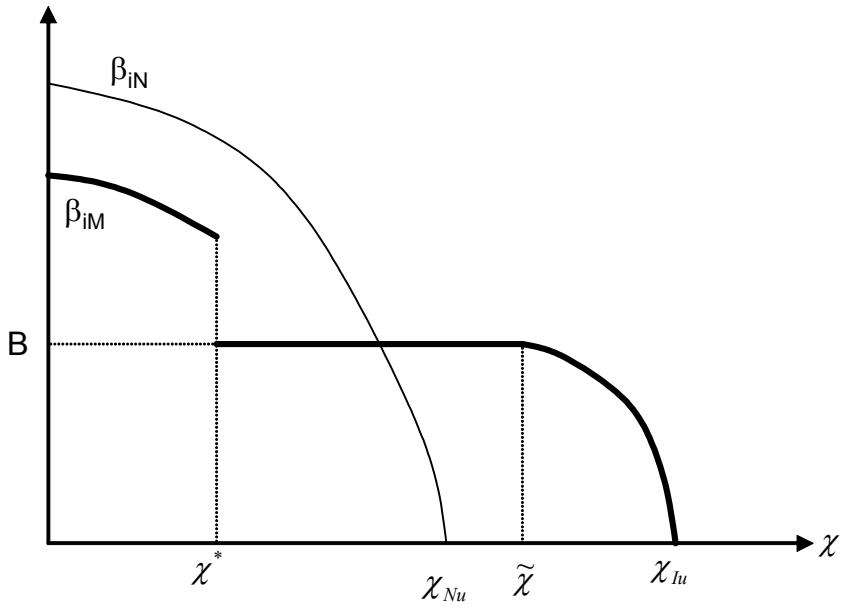

Figure 2: Optimal rent extraction rate choice

expected by MNEs for all possible interior values of χ^*.

Finally, I find from the period 1 analogue to (3), and (14) that

$$\phi_{i1I} > \phi_{i1N} > \phi_{i1O} \tag{15}$$

for all possible interior values of χ^*. Inequality (15) expresses an *investment diversion* channel in the model. Since countries that join MAI are self-selected to be the ones that extract the least rents from MNEs, MAI membership reveals information to the MNEs and increases FDI inflows relative to both a world without MAI and to a situation when a country opts out. Likewise, when opting out, there is a signal of bad quality to MNEs, and FDI inflows are lower both relative to a situation when joining MAI or when there does not exist a MAI. Equation (9) reveals then that MAI countries get an extra incentive to extract rents, since a larger rent-extraction base enlarges the time-inconsistency problem. Conversely, outsiders have a reduced incentive to extract rents, since the rent-extraction base is diminished and the time-inconsistency problem reduced.

I can summarize the discussion in this section by drawing a graph of the chosen rent extraction rate in dependence of the underlying lobbying parameter χ_i (formalized

in equation (9)). The thick line in Figure 2 describes the rent extraction rate in a world with partial MAI. If a country with lobbying parameter χ^* exists that is indifferent to join, then a jump occurs which is derived in Lemma 3. Since, when opting out of MAI, FDI inflows are reduced due to the disadvantageous signal of non-membership, the rent extraction rate must be substantially higher than the maximum standard to compensate with higher revenues both for the loss in capital income and the reduction in the rent extraction base. The thin line, instead, depicts the rent extraction rate that would have prevailed in a world without MAI (formalized in equation (6)).

2.7 Characterizing equilibrium for a given MAI

Next, I characterize whether a partial MAI exists. A country i is indifferent on the MAI membership decision if $W_{iI}(\chi^*, \beta_{iI})$ equals $W_{iO}(\chi^*, \beta_{iO})$. I can write this indifference condition as follows:

$$\begin{aligned} \tilde{W}(\chi^*) &\equiv w(E[\beta_{iO}|O], \beta_O(\chi^*), \chi^*) + w(\beta_O(\chi^*), \beta_O(\chi^*), \chi^*) \\ &\quad - w(E[\beta_{iI}|I], B, \chi^*) - w(B, B, \chi^*) \\ &= 0, \end{aligned} \tag{16}$$

where $w(x, y, \chi) \equiv 0.25\chi(1-x)^2 + 0.5u(1-x)y$. The function $w(x, y, \chi)$ is the government objective value in a regime and a year of a country with a lobbying parameter, χ, a government choice of the rent extraction rate, y, and a given expectation of MNEs on the rent extraction rate, x. Interestingly, when expectations are correct ($x = y$), the function has a peak in x for a given χ at $(u-\chi)/(2u-\chi)$. Only if $u > 1$, the peak will be in the positive range and the desired rent extraction rate is positive. Otherwise, a rent extraction rate of zero is optimal for any indifferent country i^*. If $\chi = 1$, the function $w(x, y, \chi)$ captures country welfare in a period.

Equation (16) is an implicit function of one endogenous variable χ^*, because $E[\beta_{iO}|O]$, and $E[\beta_{iI}|I]$ depend according to (11) and (12) on no other endogenous variable but χ^*.

All other endogenous variables of the subgame can be solved for, once a χ^* exists and is known.

Importantly, the indifference condition is homogeneous of degree zero in both world income and country factor endowments. Hence, the indifference function $\tilde{W}(\chi^*)$ does not depend on a country index i. Henceforth, the country index i can be dropped and country differences are fully accounted for by keeping track of χ. This property hinges on the assumption of a Cobb-Douglas functional form of the fixed cost component of MNEs. Only the strictness of the MAI participation rule, B, the distortion from the underprovision of public goods, u, and properties of the distribution function of the pressure group distortion term, $\bar{\chi}$, matter for the equilibrium.

Proposition 1 *There exists a perfect Bayesian equilibrium where all countries with* $\chi_i > \chi^*$ *join and all countries with* $\chi_i < \chi^*$ *opt out of MAI for some value* $\chi^*, 1 < \chi^* < \bar{\chi}$, *if*

$$2w\left(\frac{2u-1}{3u-1}, \frac{2u-1}{3u-1}, 1\right) - w\left(B, B, 1\right) - w\left(\frac{(2u-1)B}{\bar{\chi}-1+\ln(1-B)}, B, 1\right) > 0. \qquad (17)$$

Proof: See appendix 5.

Since the condition (17) in Proposition 1 is a bit hard to understand economically, I provide two corollaries with simple intuitions.

Corollary 1: (i) *If* $u = 1$ *, then the only equilibrium is one where all countries join MAI.*

Proof: See Appendix 6.

Corollary 1 has a simple economic meaning. If there is no underprovision of public goods ($u = 1$), then the only Pareto-optimal allocation of factors is the one without government intervention. The only two distortions in the economy left are the time-inconsistency problem and the lobbying distortion. The time inconsistency problem is fully solved by the MAI rule, and the lobbying distortion is towards zero, i.e. the free

market solution. Hence, all countries decide to join MAI and the MAI preferred by all countries is the strictest one possible, i.e. $B = 0$. Since there were LDCs objecting to MAI negotiations, the scenario without underprovision of public goods does not fit this empirical observation and I continue to analyze the case $u > 1$.

Another benchmark on the condition (17) is found with respect to the strictness of MAI, B. I define a particular benchmark of B such that

$$B^* = \frac{u-1}{2u-1} \tag{18}$$

and characterize a MAI when such a benchmark happens to be negotiated.

Corollary 2: *If $B = B^*$, then the only equilibrium is one where all countries join MAI.*

Proof: See appendix 7.

The intuition of Corollary 2 is also straight forward. The benchmark B^* is the value of the rent extraction rate that is chosen by a social planner in a world without information asymmetries. Since the lobbying distortion pushes the desired rent extraction rate below B^*, the time inconsistency distortion alone causes a rent extraction rate above B^*. However, MAI rule B^* provides to all countries a commitment device to solve the time-inconsistency problem by joining MAI at no cost. Hence, opting out of MAI makes no longer sense. Overall, a weak MAI (large B) is not sharply binding and every country joins MAI to avoid losses from signalling high rent-extraction rates in case of staying out.

Next, I compare the incentives of governments to protest against MAI negotiations. Every government objects negotiation if it expects a loss in a world with MAI compared to one without.

Proposition 2 *Governments of all countries that do not join MAI lose compared to a world without MAI; governments of all countries with $\chi_i > \tilde{\chi}$ gain. At least some governments of countries that join MAI with χ_i such that $\chi^* > \chi_i > \tilde{\chi}$ lose.*

Proof: See appendix 8.

This proposition can explain why some countries object other countries to negotiate a MAI even though they are neither forced into nor excluded from membership. A partial MAI, i.e. a MAI where some countries join and others opt out, exerts a negative information externality on non-members; outsiders signal that they are inclined to extract large rents from MNEs. The resulting investment diversion harms governments that do not decide to join. It need to be kept in mind, however, that a loss for a government does not necessarily imply a welfare loss of the country, since government objectives are distorted by lobbying groups.[24]

Proposition 2 can explain the protest storm of some LDCs against the negotiation of MAI by the club of the OECD countries, although they were both free to opt in or out. According to my explanation, they were fearing the information externality that may arise from the decision to opt out. The protest comes from governments that are ex post but not ex ante contra free-market spirited. Again, a quote by the former Commerce Secretary to Government of India supports this model feature: "*Selective and judicious government intervention is therefore widely considered necessary to support or protect domestic industry and technology creation ... Adequate freedom and flexibility to pursue their own policies towards FDI and foreign technology is therefore regarded by developing countries as a matter of fundamental importance ...*" (Ganesan, 1998, p. 5)

3 Endogenous MAI Formation

I extend now the game by two additional stages to endogenize which countries start negotiating agreements among themselves. I superimpose on top of the previous stages the choice of countries with which other countries to start negotiation and the choice of each negotiation group of how strict the MAI rule is going to be. The new timing is given in Figure 3.

[24]The result in proposition 2 mirrors the one in Turrini and Urban (2001). However, I derive the result analytically in a world of partial MAI while the previous study derived this result for a complete MAI.

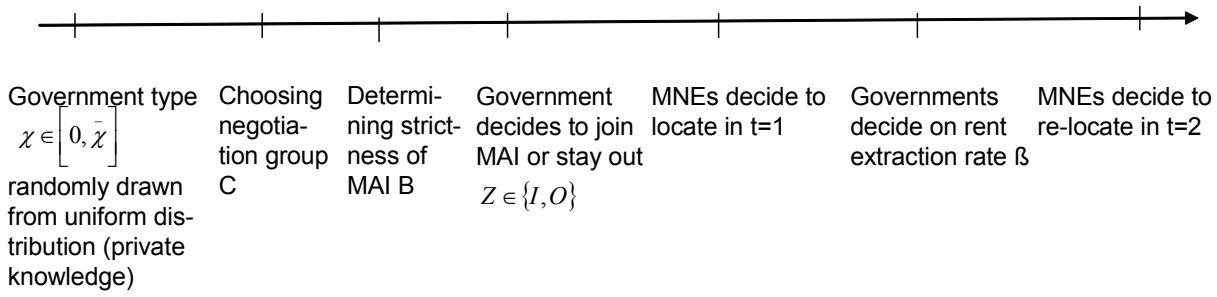

Figure 3: Sequencing of events in section 3

3.1 Negotiating MAI

In this section, I turn to the stage when MAI is negotiated. I assume that there exists a club C of a countable number of countries that starts exclusively negotiating an agreement.[25] In particular, this group chooses the threshold B. The club is assumed to have more favorable political-risk characteristics than the world as a whole.

All countries are then free to opt in or out after the agreement is written. I assume also that MAI takes the form of a rule $\beta \leq B$. Then the strictness of MAI, i.e. the threshold value B, can be found from a simple Nash bargaining solution where the Nash product is defined as

$$\prod_{c \in C} \left[W_{cI}\left(\chi_c, \beta_{iM}, B\right) - W_{cN}\left(\chi_c, \beta_{iM}\right)\right]^{\frac{1}{|C|}}, \tag{19}$$

$|C|$ is the number of group members, and the government objective function of successful negotiation $W_{cI}\left(\chi_c, \beta_{iM}, B\right)$ obtains an additional argument B, since the strictness of MAI is now allowed to vary. In addition applies the participation constraint due to the assumption of unanimity among negotiators

$$\left[W_{cI}\left(\chi_c, \beta_{iM}, B\right) - W_{cN}\left(\chi_c, \beta_{iM}\right)\right] \geq 0 \tag{20}$$

for all countries $c \in C$. I denote with χ_s the country with the smallest weight on capital

[25] By the chosen set-up, Northern home country interests are consistently excluded from shaping the agreement, because MNE profits are zero in any case. Any Northern home country will thus have interests similar to host countries. This feature clearly falls short of reality. Section 5 argues, however, why this particular model feature does not upset the model mechanics.

income in the negotiation group C, i.e. $\chi_s = \min_{c \in C} \{\chi_c\}$. Likewise, I denote the country with the largest weight χ_l, i.e. $\chi_l = \max_{c \in C} \{\chi_c\}$. Then the constraint (20) is not binding for any country unless it is binding for country χ_s, since χ_s is the country that is first hit by a welfare loss according to Proposition 2 when the agreement gets too strict.

When maximizing the Nash product (19) with respect to MAI strictness B under the participation constraint (20), one obtains the following first order condition

$$\sum_{c \in C} \frac{\partial W_{cI}(\chi_c, \beta_{iM}, B)}{\partial B} \frac{1}{W_{cI}(\chi_c, \beta_{iM}, B) - W_{cN}(\chi_c, \beta_{iM})} = 0. \tag{21}$$

The negotiated strictness of the agreement is a weighted average over the individually optimal choices given by the first oder condition $\frac{\partial W_{cI}(\chi_c, \beta_{iM}, B)}{\partial B} = 0$ the solution of which is denoted by $B(c)$. The first order condition (21) implies that the participation constraint (20) is never binding. As soon as B gets too strict for country χ_s it obtains an infinite weighting factor on its own first-order condition which ensures that B does not deviate too much from its own preferred choice. Hence, there must always exist a solution to the first order condition (21) such that the negotiated strictness B is between the one desired by the club member with the largest weight and the one with the lowest weight, i.e. $B(l) < B < B(s)$. Finally, B is unique, because the function $W_{cI}(\chi_c, \beta_{iM}, B)$ is single-peaked.

I collect the results in the following proposition.

Proposition 3 *There exists a unique mapping of a club with a member set C and an agreement threshold B, $B = B(C)$. The strictness of this agreement is strictly bounded by the minimum and maximum of the individually optimal choices, i.e. $B(l) < B(C) < B(s)$.*

Proposition 3 is useful to characterize the equilibrium of the entire game.

3.2 Prevalence of Bilateral Investment Agreements

I turn now to the first stage of the game when many different groups can potentially negotiate agreements. I assume that any country can negotiate and join any number of

agreements. Credibility of an agreement requires at least one partner country ("natural contract enforcer", Ethier, 1998) from the North, although the North continues to stay outside the model. While there is no endogenous choice by Northern countries, it is plausible that they have an interest in protecting their MNEs by the strictest agreeable FDI rules applied to potential host countries. Hence, there is a Northern supply of commitment devices to any Southern host-country demand for it at the desired degree of strictness. If a host country joins several agreements, then the rules of the strictest agreement apply to all MNEs located in a host country.[26] Moreover, each host country can be analyzed in isolation by construction of the model. Hence, the previous analysis of MAI in section 2 carries over to the extension when many agreements can be negotiated by many different groups of countries, since only the strictest agreement to which a host country subscribes is economically relevant under these assumptions. Finally, we assume marginal costs to negotiate an agreement.

It turns out that the negotiation result of an exogenously given heterogeneous group is not yielding a stable outcome. Not only is there protest against an agreement from some outsiders that obviously lose from the negotiation according to Proposition 2. There is also a disinterest of some countries within the negotiation group that find the negotiated contract too weak. This incentive leads to the next proposition.

Proposition 4 *(i) To any negotiation group of countries C that is heterogeneous in the lobbying parameter χ, there exists a strict sub-group of C that prefers a MAI negotiated within the sub-group.*
(ii) If there is Northern supply of commitment technology to form an investment agreement at any degree of strictness, then the only perfect Bayesian equilibrium is one where each country negotiates agreements together with countries of the same type χ.

[26] A host country attracts MNEs of those Northern home countries that are granted the best protection under its strictest agreement. One can think of MNEs engaging into "treaty shopping" by setting up holding companies in the home country that obtains preferential treatment by the host country of its foreign affiliate. See Weichenrieder (2006) for empirical evidence by investigating ownership chains.

(iii) The negotiation outcome will be

$$B(\chi) = \max\left[\frac{u-\chi}{2u-\chi}, 0\right]. \tag{22}$$

Proof: See appendix 9.

The intuition behind this proposition is simple. Consider again the country with the largest lobbying term χ_l in the negotiation group C. By forming a MAI which is stricter than the average over all members, this country can reveal more information on its type and thereby attract additional FDI without constraining its rent extraction rate choice, since country χ_l's optimal choice of the constraint B is a strictly lower value than the one negotiated by the large group. Hence, the most liberal country of a group has always an incentive to form a sub-group and to negotiate a stricter MAI. If every government does that, then there will result a world where there is an investment agreement for each type χ (BIT regime).

This theoretical result matches quite well with two empirical observations: First, after the group of negotiating countries was enlarged to include 8 LDCs on top of the OECD member countries to alleviate the protest storm of LDCs and NGOs, the latest versions of the agreement appeared full of exemptions of FDI liberalization, while the first versions were attempting a very general liberalization of FDI. When this contract suggestion was so weak that it was no longer binding for most countries, the main FDI source countries lost interest in it.[27]

Second, while a multilateral agreement failed in 1998, there was a spread of BITs at varying strictness throughout the 90ies. The number of BITs quadrupled during the 90ies to well above 1600 in the year 1998[28].

While the BIT regime is the one that is preferred by any government, it is not world-welfare maximizing, since government objectives are distorted by lobbying. I turn next to a MAI that is world-welfare superior compared to a BIT regime.

[27] This reasoning for the failure of MAI is found in Hoekman and Saggi (2000, p. 185): *"In the end, OECD countries were only able to agree on a package that was less far-reaching than what is often found in the bilateral investment agreements between high-income and developing countries ..., reducing the interest of the business community to push for the agreement."*

[28] See UNCTAD (1998), figure III.3, p. 83.

4 A World-Welfare-Superior MAI-Design

In this section, I suggest a MAI design that is welfare-superior compared to a regime of BITs, although I do not provide an optimal mechanism. Such a world-welfare-superior agreement can be designed as follows.

Proposition 5 *Suppose an agreement exists such that*

$$\beta_{iI} \leq \frac{u-1}{2u-1} \equiv B^* \tag{23}$$

for all members, and bilateral agreements outside this arrangement do not exist, then this arrangement is strictly world–welfare superior compared to a regime with BITs.

Proof: See appendix 10.

Again there is a simple intuition behind this result. First note that all countries join MAI under rule (23), as was found in Corollary 2. Then, the chosen rent extraction rate of any country i if strictly positive is

$$\beta_{iI} = \max\left[\frac{u-\chi_i}{2u-\chi_i} + \frac{u\left(1-E[\beta_{iI}\,|I\,]\right)}{[2u-\chi_i]}, B^*\right] \tag{24}$$

which is between the one that maximizes the government-objective function, i.e. $\frac{u-\chi_i}{2u-\chi_i}$, and the one that maximizes the country-welfare function, i.e. B^*, in period 2. One qualification is necessary for period 1, because then the expected rent-extraction rate is not equal to the actual one. In fact, MNEs are too pessimistic about their business opportunities in countries with low rent extraction and too optimistic in countries with large rent extraction. However, this effect averages out when calculating world-welfare. All together, there must be a world-welfare gain from such an arrangement, since the country-welfare function is single-peaked.

The design exploits two effects: first the BIT regime is one that solves completely the time-inconsistency problem but biases the rent extraction rate downward through lobbying of capital owners. A MAI rule such as (23) does not rule out the time-inconsistency bias. On the contrary, it exploits this bias to rise the voluntarily chosen rent extraction

rate in the direction of the world-welfare optimal one. At the same time, the cap at B^* ensures that the time-inconsistency bias is not too strong to rise the rent extraction rate beyond what is world-welfare optimal. Moreover, the MAI rule (23) ensures that all countries indeed voluntarily join MAI (Corollary 2) such that the cap applies indeed to all countries and prevents thus a time-inconsistency bias that is too strong for some countries.

Two remarks on Proposition 5 follow. First of all, the appeal of MAI design (23) is its property to alleviate a national political-economy distortion without touching on the sovereignty of this country. The first best choice of a government whose preferences are distorted relative to world welfare is to find a partner that guarantees its desired policy. This strategy is only available, however, if some other country with a commitment technology is willing to co-operate. Once this is not the case, such a government enters voluntarily a MAI agreement with design (23) and its own time-inconsistency problem drives its policy in the direction of the world-welfare optimal choice.

Second, there may still be some countries that lose compared to a BIT regime, although the MAI design (23) is world-welfare superior. To see why, note that the MAI rule does not completely resolve the information asymmetry. Countries that choose low rent extraction indirectly subsidize countries that extract more rents, because MNEs have positive ex-post profits in the first group and negative ones in the second. The first group of countries can avoid the subsidy by entering the BIT regime where all informational asymmetries are resolved and country welfare is improved at the expense of world welfare. Taking this argument at face value, there is a reason for a "grand bargain"[29], i.e. the possibility to use investment liberalization as a bargaining chip in exchange for other trade issues to create side-payment instruments and compensate losers by winners.

5 Model Discussion

In this section I discuss the robustness of my results. I address in turn the FDI model, the political equilibrium within host countries, and the way MAI is implemented into the model.

[29] Hoekman and Saggi (2000) argue for such a "grand bargain".

5.1 Robustness of FDI model

The FDI model is clearly rudimentary and should be best thought of as a reduced form. Crucial for the model is that FDI inflows change relative factor incomes and that factor income as a whole increases through FDI inflows. As pointed out, this hinges crucially on viewing FDI as flow of technology rather than of physical capital. It also hinges on a pure theory of vertical FDI. Markusen (2002) has shown in a model of horizontal and vertical FDI that investment liberalization may indeed be welfare-deteriorating at certain factor endowment combinations. Apart from these restrictions, the FDI model generalizes to general functional forms, as well as different sector settings. More seriously, I have ignored MNE profits that are arguably at the heart of highly-developed-country interests in fostering a MAI.[30]

5.2 Robustness of political equilibrium

I have described a rather specific political equilibrium. The lobbying distortion arises exclusively from capital owners. However, trade unions may also exert lobbying power. Moreover, I have ignored monopoly rents of local firms that may be destroyed by FDI inflows thus turning local-firm owners against FDI liberalization. Last but not least, governments may have a genuine interest in maximizing government revenue at the expense of country welfare.

As a result of all those political distortions in government decision making, there will be a policy that biases rent extraction upward. Hence, there will always be a group of countries that does not join MAI if this upward bias is sufficiently strong. Clearly, outsider governments of MAI still lose and there is a strong incentive of governments that wish to extract the least rents from MNEs to form bilateral agreements and signal their type. Finally, the suggested MAI rule B^* is world-welfare superior, since the rent-extraction rate will be above B^* even when resolving the time-inconsistency problem in a BIT regime. Hence, those countries that enter MAI are just kept at the country-welfare

[30] There is a discussion of the implications of an asymmetric distribution of outward FDI in the presence of MNE monopoly rents in Turrini and Urban (2001).

optimum B^* and those countries that opt out, face investment diversion that erodes their rent-extraction-rate base and induces them to choose less rent extraction. However, then the political-economy distortion is alleviated even among countries that do not join MAI.

5.3 Robustness of MAI rule

Next, I address informally the question why a common maximum standard may be the rule that is agreed upon by a club. There are four reasons for using such a rule. First, a club may try to shift a maximum of welfare from non-members to club members. If the club members are not constrained by the rule, it is very suitable to exert an asymmetric impact on countries.

Second, such a rule is easier to agree to than - say - an equality constraint. After all, all countries that negotiate MAI will be able to choose their optimal policy without any constraints once they accept that one country cannot influence the average perception of MNEs on all future members. In particular, this is the case if there are many countries part of the club that negotiates the agreement. Then, there is no hope for any single country to shape the agreement in its own interest alone.

Third, Turrini and Urban (2001) show that the actual shape of how MAI membership reduces the rent extraction capabilities of host countries may vary substantially and yet produce very similar outcomes. It is only important that some countries have a relative advantage above others. However, any negotiation of a club is likely to yield such an outcome in one way or another.

Finally, I have demonstrated that an inequality rule with a particular cap value is world welfare superior. Taking such a rule as a starting point but manipulating it in the negotiators' interest during the negotiation is easier to obscure from the electorate.

6 Concluding Remarks

I have addressed in this study the puzzle why the multilateral investment agreement (MAI) negotiated by the OECD in 1998 was objected by many less developed countries,

and eventually failed while at the same time bilateral investment agreements spread out. My explanation rests on a model where there are three distortions. First, there is a time-inconsistency problem that inclines governments to extract too much rents from FDI. Second, there is an underprovision of public goods which renders at least some rent extraction desirable to finance the provision of public goods. Third, there is a lobbying distortion in the political system by capital owners who favor FDI inflows. Moreover, the lobbying distortion is hidden information to foreign investors giving rise to political risk. MAI is a device to pre-commit to a certain policy of rent extraction if a group of countries agrees on it.

Any arbitrary club of countries has an incentive to distinguish from countries with larger propensity to rent extraction to redirect FDI towards themselves. This exerts a negative information externality on outsiders and can explain the protest of some LDCs against MAI negotiation. At the same time, there is a genuine interest of some club members to deviate and form an even stricter agreement among like-minded countries. A system of bilateral investment agreements emerges that reveals government types to MNEs. While this solves the informational distortion in the economy, the political economy distortion remains present. If a MAI design is implemented that excludes bilateral treaties and sets a cap of the rent extraction rate at the welfare-optimal level, then this alleviates the political economy distortion and increases world welfare, because a MAI retains a modest time-inconsistency bias that induces governments to deviate from their own first-best policy in the direction of the first-best policy for the country. Yet, all countries join voluntarily once these two rules are implemented in the world trading system.

A Appendix

A.1 Appendix 1. Proof of Lemma 1.

First, the interior solution of β_{iN} is zero if for some country r,

$$\chi_r = u\left[1 + \frac{2\phi_{r1N}}{YK_i}\right] \equiv \chi_{Nu} > u, \tag{25}$$

where the inequality follows, since ϕ_{r1N} is strictly positive by the period 1 analogue of (3).

Note, second, that $0 \leq \beta_{iN} \leq 1$ by definition of an extraction rate with strict inequality for some i. Then, I have

$$\phi_{i1N} = 0.5K_i\left(1 - E\left[\beta_{iN}\,|N\right]\right)Y < 0.5K_iY. \tag{26}$$

Next, I find:

$$\begin{aligned} \frac{\partial \beta_{iN}}{\partial \chi_i} &= \begin{cases} -\frac{u\left[1-\frac{\phi_{i1N}}{0.5YK_i}\right]}{(\chi_i - 2u)^2} & \text{if } \chi_i < \chi_{Nu} \\ 0 & \text{else} \end{cases} \\ &\leq 0, \end{aligned} \tag{27}$$

where the inequality follows from (26). Moreover, for a country s such that $\chi_s = 1$, holds:

$$\beta_{sN} = \frac{u + u\frac{\phi_{s1N}}{0.5YK} - 1}{2u - 1} < 1, \tag{28}$$

where the inequality follows from (26). (27) and (28) imply $\beta_{iN} < 1$ for all i.□

A.2 Appendix 2: Proof of Lemma 2.

First, all countries i with a $\beta_{iI} \leq B$ join MAI, since they are not constrained by the rule, but can gain additional FDI inflows by revealing information on their type being liberal to FDI. Next, only countries i with $\beta_{iO} > B$ do not join MAI. Otherwise, they would signal to be of bad quality without exploiting the possibility of violating the MAI rule. This

implies also that $E[\beta_{iO}|O] > B > E[\beta_{iI}|I]$. But then follows from the period 1 analogue of (4) for any country i that

$$r_{i1I} > r_{i1O} \tag{29}$$

and

$$r(\beta_{iO}) < r(B). \tag{30}$$

Differentiating the country indifference condition for any i and given ϕ_{i1Z} yields

$$\frac{\partial [W_{iI}(\chi, \beta_{iI}) - W_{iO}(\chi, \beta_{iO})]}{\partial \chi} = [r_{i1I} + r(B) - r(\beta_{iO}) - r_{i1O}] K_i > 0, \tag{31}$$

where I applied the envelope theorem and the inequality follows from (29) and (30). By assumption, I have

$$W_{iI}(\chi^*, \beta_{iI}) = W_{iO}(\chi^*, \beta_{iO}). \tag{32}$$

The inequality (31) and equation (32) together imply the Lemma 2. □

A.3 Appendix 3. Closed Form Solutions of Expected Rent Extraction Rates.

When plugging (6) into equation (11) and developing the integral I obtain a convenient closed form solution of the expected rent extraction when outside of MAI

$$E[\beta_{iO}|O] = \frac{-u \ln\left(\frac{2u-\chi^*}{2u-1}\right)}{\chi^* - 1 - u \ln\left(\frac{2u-\chi^*}{2u-1}\right)} \tag{33}$$

while the expected rent extraction rate when opting in is only found after tedious calculations:

$$E[\beta_{iI}|I] = \frac{(2u-\chi^*)B}{\bar{\chi} - \chi^* - u \ln(1-B)}. \tag{34}$$

A.4 Appendix 4. Proof of Lemma 3.

I proof (i) first. Assume that there are no constraint countries, i.e. $\chi_i \geq \tilde{\chi}$ enters MAI while $\chi_i < \tilde{\chi}$ opts out, when a country $\tilde{\varepsilon}$ with $\chi_{\tilde{\varepsilon}} = \tilde{\chi}$ is indifferent, i.e.

$$W_{\tilde{\varepsilon}I}\left(\tilde{\chi}, \beta_{\tilde{\varepsilon}I}\right) = W_{\tilde{\varepsilon}O}\left(\tilde{\chi}, \beta_{\tilde{\varepsilon}O}\right). \tag{35}$$

Then, the optimal choice of β_{iM} in a world with MAI is given analogously to (6) by

$$\beta_{iM} = \begin{cases} \frac{\chi_i - u}{\chi_i - 2u} - \frac{1 - E[\beta(i,O)|O]}{\chi_i - 2u} \text{ if } \chi_i < \tilde{\chi} \\ \frac{\chi_i - u}{\chi_i - 2u} - \frac{1 - E[\beta(i,I)|I]}{\chi_i - 2u} \text{ if } \tilde{\chi} < \chi_i < \chi_{Iu} \\ 0 \qquad\qquad\qquad\qquad \text{else} \end{cases} \tag{36}$$

where χ_{Iu} is given by

$$\chi_{Iu} \equiv u\left[2 - E\left[\beta_{iI}\,|I\right]\right].$$

Next, I find that

$$\phi_{\tilde{\varepsilon}1I} > \phi_{\tilde{\varepsilon}1O} \tag{37}$$

by Lemma 2 and the period 1 analogue of (3) and equation (36). From (1) and (37) follows

$$r_{\tilde{\varepsilon}1I} > r_{\tilde{\varepsilon}1O}. \tag{38}$$

Next, I derive from (5) and the assumption that no country is constrained the expressions

$$W\left(\tilde{\varepsilon}, I\right) = \max_{\beta}\left[\chi_{\tilde{\varepsilon}} r_{\tilde{\varepsilon}1I} K_{\tilde{\varepsilon}} + \chi_{\tilde{\varepsilon}} r\left(\beta\right) K_{\tilde{\varepsilon}} + u\beta Y \phi_{\tilde{\varepsilon}1I} + u\beta Y \phi\left(\beta\right)\right] \tag{39}$$

and

$$W\left(\tilde{\varepsilon}, O\right) = \max_{\beta}\left[\chi_{\tilde{\varepsilon}} r_{\tilde{\varepsilon},1,O} K_{\tilde{\varepsilon}} + \chi_{\tilde{\varepsilon}} r\left(\beta\right) K_{\tilde{\varepsilon}} + u\beta Y \phi_{\tilde{\varepsilon}1O} + u\beta Y \phi\left(\beta\right)\right], \tag{40}$$

respectively. From (37)-(40) follows that $W\left(\tilde{\varepsilon}, I\right) > W\left(\tilde{\varepsilon}, O\right)$ which contradicts the assumption (35). Hence, there must exist countries i such that $\chi^* < \chi_i < \tilde{\chi}$.

For part (ii), I assume to the contrary that $\beta_{i^*O} = B$ for the indifferent country i^*

such that

$$W(i^*, O) = W(i^*, I). \tag{41}$$

By Lemma 2, the equations analoguos to (3) and (4) in period 1, and equation (41), I infer that $\phi_{i^*1I} > \phi_{i^*1O}$ and $r(i^*, 1, I) > r(i^*, 1, O)$. But then follows under consideration of (5) that

$$W(i^*, I) - W(i^*, O) = B[\phi_{i^*1I} - \phi_{i^*1O}]Y + r_{i^*1I} - r_{i^*1O} > 0, \tag{42}$$

which contradicts (41). Since by (i), β_{i^*O} is constrained, it cannot be smaller than B. Hence, (ii) follows. □

A.5 Appendix 5. Proof of Proposition 1.

Condition (17) is equivalent to

$$\tilde{W}(1) > 0. \tag{43}$$

Next, I find from the period 1 analogue to (3), and equations (9) and (15) that

$$\beta_{iI} > \beta_{iO} \tag{44}$$

for all possible interior values of χ^*. I conclude from (10) and (44) in turn that there must exist a $\chi^* = \chi'$ with $\chi' < \tilde{\chi}$ such that

$$\beta_{iO} = B \tag{45}$$

at $\chi^* = \chi'$. However, then holds

$$\begin{aligned} \tilde{W}(\chi') &= w(E[\beta_{iO}|O], B, \chi') - w(E[\beta_{iI}|I], B, \chi') \\ &= \chi'[r_{i1O} - r_{i1I}]K_i \\ &\quad + u \cdot B \cdot Y \cdot [\phi_{i1O} - \phi_{i1I}] \\ &< 0, \end{aligned} \tag{46}$$

where the inequality follows from (1) and (15). Also $\tilde{W}(\chi^*) < 0$ for $\chi^* > \chi'$. Hence, $\tilde{W}(\chi^*) = 0$ can only have an interior solution if $1 < \chi^* < \chi'$. Additionally, the function $\tilde{W}(\chi^*)$ is continuous in the relevant range. Then applies together with (43) and (46) the intermediate value theorem.

Having proven the existence of some interior value of χ^*, the Nash conditions can be easily assembled. Countries χ_i have no incentive to deviate from the equilibrium strategy to opt in if and only if $\chi_i \leq \chi^*$ by Lemma 2. Firms have no incentive to deviate from their location decision.[31] Expectation formation in (33) and (34) is consistent with Bayes' rule. To see this, note that the prior belief to be type χ_i is $1/(\bar{\chi} - 1)$, the conditional probability to observe policy I is 1 if $\chi_i > \chi^*$ and 0 else. The unconditional probability of policy I is $(\bar{\chi} - \chi^*)/(\bar{\chi} - 1)$. Hence, the posterior probability to be type χ_i, after policy I is observed, is $1/(\bar{\chi} - \chi^*)$ for $\chi_i > \chi^*$ and 0 else. Likewise, the posterior belief of type χ_i, after policy O is observed, is $1/(\chi^* - 1)$ for $\chi_i < \chi^*$ and zero else. However, expectation formation in (33) and (34) uses exactly these posterior probabilities. Finally, countries have no incentive to deviate from their optimal rent extraction rate choice by construction of the maximization problem and Lemma 3 given expectation formation.□

A.6 Appendix 6. Proof of Corollary 1.

The function $w(h, h, \chi)$ is quadratic in h and has its peak at $h = (u - \chi)/(2u - \chi)$. If $u = 1$, then the peak is in the negative range that is economically irrelevant and I have from (14) the inequality

$$h < B < \beta_{i^*O}. \tag{47}$$

Next, I investigate the function $w(x, y, \chi)$ that has the properties

$$\frac{\partial w(h, h, \chi)}{\partial h} < 0 \tag{48}$$

[31] To be formally precise, MNEs randomize with probability $\phi_{\iota t}$ their choice to locate in country i. At the equilibrium value of ϕ_{it}, expected profits are zero when locating in country i, but also when not locating there. By proposition 8.D.1 in Mas-Colell et al. (1995) this suffices for a mixed-strategy Nash equilibrium of the location subgame in period t. By the law of large numbers there will be exactly ϕ_{it} MNEs in country i.

and

$$\frac{\partial w(x,y,\chi)}{\partial x} < 0. \tag{49}$$

over the entire range $0 \leq h, x, y \leq 1$ and for any $\chi \geqq 1$. From (9), (11), and (12) follows when assuming an interior solution, i.e. when there is at least one country i^* with $\chi_{i^*} = \chi^*$ and

$$\tilde{W}(\chi^*) = 0, \tag{50}$$

that the inequality

$$E[\beta_{iO}|O] \geq \beta_{i^*O} > B \geq E[\beta_{iI}|I] \tag{51}$$

holds. Next, I obtain from the sign of the derivative in (49) the inequalities

$$w(E[\beta_{iO}|O], \beta_{i^*O}, \chi^*) < w(\beta_{i^*O}, \beta_{i^*O}, \chi^*) \tag{52}$$

and

$$w(E[\beta_{iI}|I], B, \chi^*) > w(B, B, \chi^*). \tag{53}$$

Likewise, I exploit the sign of the derivative in (48) and the inequality (47) to conclude that

$$w(\beta_{i^*O}, \beta_{i^*O}, \chi^*) < w(B, B, \chi^*). \tag{54}$$

When assembling the inequalities (52), (53), and (54), I have

$$\begin{aligned} & w(E[\beta_{iO}|O], \beta_{i^*O}, \chi^*) + w(\beta_{i^*O}, \beta_{i^*O}, \chi^*) \\ < \; & 2w(\beta_{i^*O}, \beta_{i^*O}, \chi^*) \\ < \; & 2w(B, B, \chi^*) \\ < \; & w(E[\beta_{iI}|I], B, \chi^*) + w(B, B, \chi^*) \end{aligned} \tag{55}$$

However, inequality (55) implies that

$$\begin{aligned} & \tilde{W}(\chi^*) \\ = \; & w(E[\beta_{iO}|O], \beta_{i^*O}, \chi^*) + w(\beta_{i^*O}, \beta_{i^*O}, \chi^*) - w(E[\beta_{iI}|I], B, \chi^*) - w(B, B, \chi^*) < 0 \end{aligned} \tag{56}$$

for any χ^* which contradicts (50). □

A.7 Appendix 7. Proof of Corollary 2.

Suppose there exists a partial MAI. Then, must hold

$$\tilde{W}(1) > 0. \tag{57}$$

Since from (14), in particular $E[\beta_{iO}|O] \geq \beta_{iO}$, I have that

$$w(\beta_{iO}, \beta_{iO}, 1) \geq w(E[\beta_{iO}|O], \beta_{iO}, 1). \tag{58}$$

Analogously, from (14), in particular $B^* \geq E[\beta_{iI}|I]$, I have that

$$w(B^*, B^*, 1) \geq w(E[\beta_{iI}|I], B^*, 1). \tag{59}$$

Next I note that $w(h, h, 1)$ is quadratic in h and peaks at

$$h = \frac{u-1}{2u-1} \equiv B^*. \tag{60}$$

Hence, (60) implies that

$$w(B^*, B^*, 1) > w(\beta_{iO}, \beta_{iO}, 1) \tag{61}$$

However, (58), (59), and (61) lead to the inequality

$$\begin{aligned} \tilde{W}(1) &= w(\beta_{iO}, \beta_{iO}, 1) + w(E[\beta_{iO}|O], \beta_{iO}, 1) \\ &\quad - w(B^*, B^*, 1) - w(E[\beta_{iI}|I], B^*, 1) \\ &< 0, \end{aligned} \tag{62}$$

which contradicts (57). Hence, there cannot exist a partial MAI, where some countries opt in and others out. Since some countries always join MAI and partial MAI does not exist, all countries must join MAI. □

A.8 Appendix 8. Proof of Proposition 2.

I have analogue to equation (13) that

$$\beta_{iZ} = 1 - u\frac{E\left[\beta_{iZ}\left|Z\right.\right]}{2u - \chi_i} \tag{63}$$

for $Z \in \{N, O\}$. Forming expected values yields

$$E\left[\beta_{iN}\left|N\right.\right] = 1 - \frac{E\left[\beta_{iN}\left|N\right.\right]}{\bar{\chi} - 1}\int_1^{\bar{\chi}} \frac{u}{2u - \chi} d\chi \tag{64}$$

and

$$E\left[\beta_{iO}\left|O\right.\right] = 1 - \frac{E\left[\beta_{iO}\left|O\right.\right]}{\chi^* - 1}\int_1^{\chi^*} \frac{u}{2u - \chi} d\chi, \tag{65}$$

respectively. One can readily see that

$$\frac{1}{\bar{\chi} - 1}\int_1^{\bar{\chi}} \frac{u}{2u - \chi} > \frac{1}{\chi^* - 1}\int_1^{\chi^*} \frac{u}{2u - \chi} d\chi, \tag{66}$$

as $\bar{\chi} > \chi^*$. The equation that corresponds to (3) in period 1, and equations (64), (65), and (66) imply in turn that

$$\phi_{i1N} > \phi_{i1O}. \tag{67}$$

Next, the government objective function can be written in the two regimes of a world without MAI $W\left(i, N\right)$ and a world, where a country opts out of MAI, $W\left(i, O\right)$ as

$$\begin{aligned} & W\left(i, N\right) \\ & = \max_{\beta}\left[\chi_i r\left(\phi_{i1N}\right) K_i + \chi_i r\left(\beta\right) K_i + u\beta Y \phi_{i1N} + u\beta Y \phi\left(\beta\right)\right] \\ \equiv\ & \max_{\beta} W\left(i, \beta, \phi_{i1N}\right) \end{aligned} \tag{68}$$

and

$$\begin{aligned} & W(i,O) \\ = & \max_{\beta}\left[\chi_i r\left(\phi_{i1O}\right)K_i+\chi_i r\left(\beta\right)K_i+u\beta Y\phi_{i1O}+u\beta Y\phi\left(\beta\right)\right] \\ \equiv & \max_{\beta} W\left(i,\beta,\phi_{i1O}\right), \end{aligned} \tag{69}$$

where I note that the function $W(i,\beta,\phi)$ applies in both regimes albeit with different arguments. I recall that the functions $r(.)$, $\phi(.)$ are defined in (3) and (4), respectively. I note that

$$\frac{\partial W(i,\beta,\phi)}{\partial\phi}>0. \tag{70}$$

Next, I infer from equations and inequalities (67)-(70) that

$$W(i,O)<W(i,N) \tag{71}$$

for every country i with $\chi_i<\chi^*$, when a partial MAI exists with indifferent country χ^*. Similarly, I find

$$\phi_{i1N}<\phi_{i1I} \tag{72}$$

and eventually

$$W(i,I)>W(i,N) \tag{73}$$

for all countries i with $\tilde{\chi}\leq\chi_i\leq\bar{\chi}$.

Finally, I argue why some countries i with $\chi^*\leq\chi_i\leq\tilde{\chi}$ must lose in a world with MAI relative to a world without MAI. The country i with $\chi_i=\chi^*$ is indifferent of joining MAI but when it does not join MAI it unambiguously loses according to (71). Hence, it must also lose from a world of MAI when entering MAI.□

A.9 Appendix 9. Proof of Proposition 3.

Proof of part (i): Consider countries c that form a subgroup of all countries C that negotiate a MAI in a club such that $1<\chi_c\leq\chi_l\leq\bar{\chi}$ for $c\in C$ and a threshold value χ_l.

Then, countries c have a government objective value $W(c, C)$ conditional on the group C negotiating an agreement $B(C)$ that can be written as follows:

$$W(c, C) = \max_{\beta} \left[w\left(E\left[\beta_{cI} \left| B(C), I \right. \right], \beta, \chi_c \right) + w(\beta, \beta, \chi_c) \right]. \tag{74}$$

where $E\left[\beta_{cI} \left| B(C), I \right. \right]$ denotes the conditional expected value of the rent extraction rate β when knowing the strictness of MAI $B(C)$ and that the country c is MAI member. Then, there exists an alternative agreement among all countries c' with $\chi_{c'} = \chi_l$ and government objective value

$$W(c', c') = \max_{\beta} \left[w\left(E\left[\beta_{c'I} \left| B(c'), I \right. \right], \beta, \chi_{c'} \right) + w(\beta, \beta, \chi(c')) \right] \tag{75}$$

where $E\left[\beta_{c'I} \left| B(c'), I \right. \right] = \beta$. By construction must hold

$$E\left[\beta_{c'I} \left| B(c'), I \right. \right] < E\left[\beta_{cI} \left| B(C), I \right. \right]. \tag{76}$$

Call $\widehat{\beta}$ the rent extraction rate that maximizes $W(c, C)$, i.e.

$$\widehat{\beta} = \arg\max \left[w\left(E\left[\beta_{cI} \left| B(C), I \right. \right], \beta, \chi_c \right) + w(\beta, \beta, \chi_c) \right]. \tag{77}$$

Inequality (76) together with the property $dw(x, y, \chi)/dx > 0$ implies

$$\begin{aligned} W(c', C) &\equiv w\left(E\left[\beta_{c'I} \left| B(C), I \right. \right], \widehat{\beta}, \chi_{c'} \right) + w\left(\widehat{\beta}, \widehat{\beta}, \chi_{c'} \right) \\ &< w\left(E\left[\beta_{c'I} \left| B(c'), I \right. \right], \widehat{\beta}, \chi_{c'} \right) + w\left(\widehat{\beta}, \widehat{\beta}, \chi_{c'} \right) \\ &< \max_{\beta} \left[w\left(E\left[\beta_{c'I} \left| B(c'), I \right. \right], \beta, \chi_{c'} \right) + w(\beta, \beta, \chi_{c'}) \right] \\ &\equiv W(c', c'). \end{aligned} \tag{78}$$

Hence, there is an incentive for a subgroup of C to deviate from the commonly negotiated contract.

Proof of part (iii): consider a BIT of countries c with $\chi_c = \chi$. Then Proposition 3

and equation (21) yield a negotiated strictness of MAI of $B(\chi) = (u-\chi)/(2u-\chi)$.

Proof of part (ii): From part (i) follows that there exists an incentive of some subgroup to deviate from any MAI that is negotiated by a group of countries C heterogeneous in χ. Consider the following algorithm: of any possible group C the countries c' with $\chi_{c'} = \chi_l$, form an own MAI. Applying this step to all possible groups C yields a system of investment treaties for each different value of χ. There is no incentive to deviate from such a system, since $B(\chi)$ is the rent extraction rate that maximizes the government objective function of all countries c with lobbying term $\chi_c = \chi$. The assumption on marginal negotiation costs rules out that a country is member of several agreements given that only the strictes agreement is economically relevant. □

A.10 Appendix 10. Proof of Proposition 5.

Call $U(i\,|A)$ country i welfare at the MAI regime under rule B^* and $U(i\,|B)$ the country welfare under the BIT regime, respectively. By Corollary 2, all countries i join MAI under rule (23). Then, I have

$$U(i\,|A) = w\left(E\left[\beta_{iI}\,|I\right], \beta_{iI}, 1\right) + w\left(\beta_{iI}, \beta_{iI}, 1\right) \tag{79}$$

and

$$U(i\,|B) = 2w\left(\frac{u-\chi_i}{2u-\chi_i}, \frac{u-\chi_i}{2u-\chi_i}, 1\right), \tag{80}$$

respectively. (i) I recall that $dw(x,y,\chi)/dx < 0$.(ii) The function $w(h,h,1)$ has the property of single-peakedness with peak $h = (u-1)/(2u-1) = B^*$. From rule (23) and optimal β choice of a country within MAI follows

$$\beta_{iI} = \max\left[\frac{u-\chi_i}{2u-\chi_i} + \frac{u\left(1-E[\beta_{iI}\,|I]\right)}{[2u-\chi_i]}, B^*\right] \equiv \beta(\chi, I). \tag{81}$$

Hence, I have the ranking

$$\frac{u-\chi_i}{2u-\chi_i} \leq \beta_{iI} \leq B^* \tag{82}$$

and

$$\frac{u-\chi_i}{2u-\chi_i} \leq E\left[\beta_{iI}\,|I\right] \leq B^* \tag{83}$$

with at least one strict inequality for some χ_i. Next, I find

$$w\left(\frac{u-\chi_i}{2u-\chi_i}, \frac{u-\chi_i}{2u-\chi_i}, 1\right) < w\left(E\left[\beta_{iI}\,|I\right], E\left[\beta_{iI}\,|I\right], 1\right) \tag{84}$$

and

$$w\left(\frac{u-\chi_i}{2u-\chi_i}, \frac{u-\chi_i}{2u-\chi_i}, 1\right) < w\left(\beta_{iI}, \beta_{iI}, 1\right) \tag{85}$$

by applying property (ii) of the function $w(.)$ and inequalities (82) and (83). World welfare in a world with MAI rule B^*, i.e. $U(A)$, and world welfare in a regime of BITs $U(B)$ are the aggregates of all countries' welfare conditional on the regimes given in (79) and (80). Hence, I can conclude

$$\begin{aligned}
U(A) &\equiv \int U\left(i\,|A\right) di \\
&= \int_1^{\bar{\chi}} w\left(E\left[\beta_{iI}\,|I\right], \beta\left(\chi, I\right), 1\right) + w\left(\beta\left(\chi, I\right), \beta\left(\chi, I\right), 1\right) d\chi \\
&= \left(\bar{\chi}-1\right) w\left(E\left[\beta_{iI}\,|I\right], E\left[\beta_{iI}\,|I\right], 1\right) + \int_1^{\bar{\chi}} w\left(\beta\left(\chi, I\right), \beta\left(\chi, I\right), 1\right) d\chi \\
&= \int_1^{\bar{\chi}} w\left(E\left[\beta_{iI}\,|I\right], E\left[\beta_{iI}\,|I\right], 1\right) + w\left(\beta\left(\chi, I\right), \beta\left(\chi, I\right), 1\right) d\chi \\
&> \int_1^{\bar{\chi}} 2w\left(\frac{u-\chi_i}{2u-\chi_i}, \frac{u-\chi_i}{2u-\chi_i}, 1\right) d\chi \\
&= \int U\left(i\,|B\right) di \\
&\equiv U(B),
\end{aligned}$$

where the first equality uses Corollary 2, the second equality follows, since y enters linear in the function $w(x, y, \chi)$ and the law of large numbers is applied, and the inequality follows from (84) and (85). □

References

[1] Bagwell, K., Staiger, R. W., 1999. An economic theory of GATT. American Economic Review 89 (1), 215-48.

[2] Barry, F., Bradley, J., 1997. FDI and trade: the Irish host-country experience. The Economic Journal 107, 1798-1811.

[3] Braconier, H., Norbäck, P.-J., Urban, D.M. , 2005. Reconciling the evidence on the knowledge-capital model. Review of International Economics 13, 770-786.

[4] Brecher, R.A., Díaz-Alejandro, C.F., 1977. Tariffs, foreign capital, and immiserizing growth. Journal of International Economics 7, 317-322.

[5] Burgess, N., Stern, N., 1993. Taxation and development. Journal of Economic Literature 31, 762-830.

[6] Che, J., Willmann, G., 2006. The economics of a multilateral investment agreement. Mimeograph.

[7] Doyle, C., van Wijnbergen, S., 1994. Taxation of foreign multinationals: a sequential bargaining approach to tax holidays. International Tax and Public Finance 1, 211-225.

[8] Eaton, J., Gersovitz, M., 1984. A theory of expropriation and deviations from perfect capital mobility. The Economic Journal 94, 16-40.

[9] Ethier, W., 1998. Regionalism in a multilateral world. Journal of Political Economy 106 (6), 1214-1245.

[10] Facchini, G., Willmann, G., 2005. The political economy of international factor mobility. Journal of International Economics 67, 201-219.

[11] Feenstra, R.C., Hanson, G.H., 1997. Foreign direct investment and relative wages: evidence from Mexico's maquiladoras. Journal of International Economics 42, 371-393.

[12] Fernandez-Arias, E., Spiegel, M.M., 1998. North-South customs unions and international capital mobility. Journal of International Economics 46, 229-251.

[13] Ferrarini, B., 2003, A multilateral framework for investment? In: State Secretariat of Economic Affairs and S.J. Evenett (eds.): the Singapore Issues and the World Trading System: The Road to Cancun and Beyond, forthcoming.

[14] Ganesan, A.V., 1998. Developing countries and a possible multilateral framework on investment: strategic options. Transnational Corporations 7 (2), 1-38.

[15] Grossman, G.M., Helpman, E., 1994. Protection for sale. American Economic Review 84, 833-50.

[16] Grossman, G.M., Helpman, E., 1996. Foreign investment with endogenous protection, in: Feenstra, R.C. , Grossman, G.M., Irwin, D.A. (Eds.), The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati. MIT Press, Cambridge MA, pp. 199-223.

[17] Helpman, E., 1984. A simple theory of international trade with multinational corporations. Journal of Political Economy 92 (3), 451-71.

[18] Hoekman, B., Saggi, K., 2000. Multilateral disciplines for investment-related policies, in: Guerrieri, P., Eckart-Scharrer, H. (Eds.), Global Governance, Regionalism, and the International Economy. Nomos Verlagsgesellschaft, Baden-Baden, pp. 167-191.

[19] Janeba, E., 2002. Attracting FDI in a politically risky world. International Economic Review 43 (4), 1127-1155.

[20] Krugman, P.R., 2000. Technology, trade and factor prices. Journal of International Economics 50, 51-70.

[21] Lucas, R., 1990. Why doesn't capital flow from rich to poor countries? American Economic Review 80 (2), 92-96.

[22] Markusen, J.R., 2001. Commitment to rules on investment: the developing countries' stake. Review of International Economics 9, 287-302.

[23] Markusen, J.R., 2002. Multinational Firms and the Theory of International Trade, MIT Press, Cambridge MA and London.

[24] Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory, Oxford University Press, New York.

[25] Murphy, K., Shleifer, A., Vishny, R., 1989. The big push. Journal of Political Economy 97, 1003-1026.

[26] OECD, 1998. The multilateral agreement on investment. Draft consolidated text, DAFFE/MAI(98)7/REV1, OECD, Paris, available on www1.oecd.org/daf/mai/pdf/ng987r1e.pdf.

[27] Raff, H., 1992. A model of expropriation with asymmetric information. Journal of International Economics 33, 245-265.

[28] Reinhart, C.M., Rogoff, K.S., 2004. Serial default and the "paradox" of rich-to-poor capital flows. American Economic Review 94 (2), 53-58.

[29] Schnitzer, M., 2002. Debt v. foreign direct investment: the impact of sovereign risk on the structure of international capital flows. Economica 69, 41-67.

[30] Thomas, J., Worall, T., 1994. Foreign direct investment and the risk of expropriation. Review of Economic Studies 61, 81-108.

[31] Turrini, A., Urban, D., 2001. For Whom is MAI? A Theoretical Perspective on Multilateral Agreements on Investment. C.E.P.R. Discussion Paper No. 2774.

[32] UNCTAD, 1998, 1999. World Investment Report. United Nations, New York and Geneva.

[33] Weichenrieder, A., 2006, What determines the use of holding companies and ownership chains? Johann-Wolfgang-Goethe-University Frankfurt, mimeograph.