

# DISCUSSION PAPER SERIES

No. 5762

## DO CREDIBLE DOMESTIC INSTITUTIONS PROMOTE CREDIBLE INTERNATIONAL AGREEMENTS?

Paola Conconi and Carlo Perroni

*INTERNATIONAL TRADE*



**C**entre for **E**conomic **P**olicy **R**esearch

[www.cepr.org](http://www.cepr.org)

Available online at:

[www.cepr.org/pubs/dps/DP5762.asp](http://www.cepr.org/pubs/dps/DP5762.asp)

# DO CREDIBLE DOMESTIC INSTITUTIONS PROMOTE CREDIBLE INTERNATIONAL AGREEMENTS?

**Paola Conconi**, ECARES and University of Warwick and CEPR  
**Carlo Perroni**, University of Warwick

Discussion Paper No. 5762  
July 2006

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR, UK  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: [www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL TRADE**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Paola Conconi and Carlo Perroni

CEPR Discussion Paper No. 5762

July 2006

## **ABSTRACT**

### **Do Credible Domestic Institutions Promote Credible International Agreements?\***

We examine the relationship between international policy coordination and domestic policy reputation when both are self-sustaining. We show that domestic policy commitment does not necessarily facilitate international cooperation; rather, efficient policies may be most easily sustained when countries are unable to pre-commit to policy domestically. Moreover, lack of domestic commitment is more likely to facilitate international cooperation the larger are the international spillovers of domestic policies.

JEL Classification: F02 and F10

Keywords: domestic policy commitment and international agreements

Paola Conconi  
ECARES, ULB  
CP114  
Avenue F.D. Roosevelt 50  
1050 Bruxelles  
BELGIUM  
Email: pconconi@ulb.ac.be

Carlo Perroni  
Department of Economics  
University of Warwick  
COVENTRY  
CV4 7AL  
Email: c.perroni@warwick.ac.uk

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=159099](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=159099)

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=154747](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=154747)

\* We wish to thank participants to the Mid-West International Economics Group Spring 2003 Meeting, the first CEPR conference on Trade, Industrialization and Development, the CESifo Area Conference on the Global Economy and the second EURESCO Conference on the International Dimension of Environmental Policy; and seminar participants at the University of Illinois, Southern Illinois, Vanderbilt, Bristol, Royal Holloway, Cyprus, Leuven, Nottingham, Free University of Brussels (ECARES), and DELTA for helpful comments. We are also grateful to Mathias Dewatripont, Patrick Legros, Arik Levinson, Giovanni Maggi, David de Meza, Thierry Verdier, and John Whalley for their valuable suggestions. The usual disclaimers apply. We gratefully acknowledge support from CSGR at the University of Warwick and the ESRC.

Submitted 23 June 2006

## I Introduction

Do countries with more credible domestic institutions make for better international partners? Evidence on this question seems to be mixed. Indeed, members of the main international trade, environmental, and security agreements include countries with credible domestic institutions as well as countries whose credibility record is poor. Attempts by political scientists to link countries' international relations record to their domestic political institutions have also revealed mixed patterns.<sup>1</sup>

The literature on rules versus discretion<sup>2</sup> has mainly focused on the reverse question, i.e. how the participation in international agreements affects domestic policy credibility. In this respect, Staiger and Tabellini (1987) and Maggi and Rodriguez-Clare (1998), among others, have argued that *binding* international agreements, by making domestic policy changes more difficult to reverse, could enhance the credibility of policymakers when domestic policy commitment devices are not available.<sup>3</sup> The same argument can often be heard in the policy debate. For example, a desire to bolster the credibility of domestic reforms was central to Mexico's negotiations of the North American Free Trade Agreement (NAFTA).<sup>4</sup> Similarly, China's WTO accession has been viewed as a way to "lock-in the agenda for fundamental domestic reforms, which has been difficult to implement by domestic measures alone" (Bajona and Chu, 2004). This view would suggest that countries with more unreliable domestic institutions could actually be more willing partners in international agreements.

The above arguments, however, neglect enforcement considerations: absent a supranational authority with autonomous powers of enforcement, international agreements need

---

<sup>1</sup>Morgan and Campbell (1991), for example, show that democracies (which are generally characterized by comparatively more credible institutions) are not inherently more peaceful as one might expect.

<sup>2</sup>The idea that policy discretion might provide governments with an incentive to renege on earlier promises, undermining the sustainability of efficient policies, was introduced in a seminal paper by Kydland and Prescott (1977), and has since found applications in most areas of economic analysis, including the study of monetary and fiscal policy formation (see Persson and Tabellini 1994 for a review), trade policy (e.g. Staiger and Tabellini 1987; Matsuyama 1990; Tornell 1991), and environmental policy (e.g. Laffont and Tirole 1996).

<sup>3</sup>In a more recent paper, Staiger and Tabellini (1999) find that GATT rules did indeed help the US government to make domestic trade policy commitments to its private sector.

<sup>4</sup>Whalley (1998, pp. 71-72) argues that "Mexican negotiators were less concerned to secure an exchange of concessions. [...] The idea was clearly to help lock in domestic reforms through this process."

to be sustained by the threat of credible punishment between the parties involved.<sup>5</sup> Then, if international agreements are not automatically binding, one could actually conjecture a reverse linkage between domestic policy credibility and international agreements, namely that the inability to commit domestically might make it more difficult to undertake commitments vis-à-vis international partners. The nature and direction of the linkage between domestic credibility and international cooperation thus does not appear to be theoretically obvious a priori. This linkage is what we set out to investigate in this paper.

Despite being frequently alluded to in the informal debate, the two-way relationship between domestic and international policy credibility has never been formally examined before in the literature. Those who have come closest to addressing it are Rogoff (1985a) and Kehoe (1989); however, their analysis is only focused on how policy coordination between governments affects time-consistent policy choices in a single round of strategic interaction, thus abstracting from the problem of enforcement—an aspect that has become a central concern in the more recent debate on international agreements. Our analysis reveals that the manner in which time-consistent policymaking interacts with international policy cooperation under repeated interaction does not at all mirror what takes place within a single round.

If full commitment is unattainable, either domestically or internationally, vertical coordination—between each policymaker and its private sector—and horizontal coordination—between policymakers of different countries—must both be sustained by balancing each party’s temptation to deviate from a given policy against the threat of credible punishment triggered by unilateral defections. We show that, in this case, if the short-run costs associated with a lack of coordination between the government and its private sector during unanticipated policy deviations are large enough, the inability to commit domestically can make a policymaker less tempted to deviate unilaterally. In turn, this is more likely to be the case the larger the international policy spillovers, because the costs of vertical miscoordination during unanticipated deviations are increasing in the size of the spillovers. Thus, even when the international and domestic coordination problems do not directly offset one another, the need to sustain policy reputation vis-à-vis the private sector can help to achieve international cooperation; and, vice-versa, the need to sustain self-enforcing cooperation with other countries can boost domestic policy credibility. Consequently, partial coordination in the stage game—such as a binding international

---

<sup>5</sup>The problem of enforcement has been repeatedly stressed in recent literature on international cooperation (see, for example, Bagwell and Staiger 1997; Maggi 1999; Ederington 2001a,b).

agreement or a domestic policy commitment device—can actually make it more difficult to support efficient policies under repeated interaction. Self-enforcing international agreements and the need to sustain domestic policy reputation can then complement each other, working together to help support efficient policies.

The remainder of the paper is organized as follows. In Section II we model policy formation within a single round of interaction between policymakers and investors. We then consider repeated interaction over an infinite horizon in Section III. Section IV discusses implications for institution design. Section V concludes.

## II Vertical and Horizontal Miscoordination in Policy Formation

Our arguments are developed in a two-country model. In each country, the government faces a large number of individually small domestic players—its private sector. The domestic policy game is characterized by the government’s inability to precommit to a certain policy. The two governments are also engaged in a prisoners’ dilemma-type policy game, stemming from the presence of an international policy spillover. Policy formation thus suffers from both a *vertical* coordination problem between each government and its private sector, and a *horizontal* coordination problem between governments. This section describes policy formation in a single round of interaction; subsequent sections will deal with repeated interaction.

We develop our arguments by focusing on the case of environmental taxes. There is clear evidence that international environmental agreements (IEAs) suffer from an enforcement problem. Keohane (1995, p. 217) has argued that “every study that has looked hard to compliance [of all major IEAs] has concluded [...] that compliance is very spotty.” Also, Brown Weiss and Jacobson (1997) have found instances of violations to all major environmental agreements. For example, more than three hundred infractions of the CITES Convention on endangered species have been counted every year.

There is also ample evidence suggesting that environmental policy suffers from a domestic credibility problem. Politicians in both the United States and Europe have often pledged to introduce tough environmental policies but then adopted much softer policies. For example, according to the European Environmental Bureau, “there have been many words but little concrete change” in the EU taxation of energy products. A proposal to increase EU environmental taxes put forward in 1997 has not yet been adopted and

has been amended to introduce a number of exemptions for “sensitive” sectors.<sup>6</sup> There are also indications that this credibility problem stems from a tension between the goal of encouraging innovation and investment in environment-friendly technologies on the one hand and distributional concerns on the other: the prospect of future environmental taxes is instrumental to inducing firms to undertake abatement-related investment; environmental taxes, however, produce unwanted distributional effects,<sup>7</sup> which are difficult to offset through compensation.<sup>8</sup> Then, once innovation has taken place, policymakers are driven to reduce environmental taxes in order to minimize their distributional effects.<sup>9</sup> As private investors recognize the ex-post incentives of policymakers, the promise of high future emission taxes is not credible.<sup>10</sup>

Our analysis focuses on an infinitely-repeated policy game between governments, where two symmetric countries are linked by transboundary pollution externalities and where pollution abatement requires investment by the private sector. In order to focus exclusively on the environmental policy dimension, we assume that there is no trade.<sup>11</sup> The stage game consists of the following sequence: first the private sector selects levels of pollution-reducing investment on the basis of expected emission taxes; then governments choose emission taxes and pollution abatement decisions are finalized. Emission taxes are the only available policy instrument, and produce adverse distributional effects because different agents have different consumption requirements of the polluting good. Thus, when setting emission taxes governments trade off efficiency and distributional ob-

---

<sup>6</sup>See [www.eeb.org](http://www.eeb.org).

<sup>7</sup>Most energy and environmental taxes are well known to be regressive, since poorer people pay a disproportionate share of their income on these taxes relative to richer people (see, for example, Poterba 1991; OECD 1995).

<sup>8</sup>Compensation schemes typically run against incentive-compatibility problems. For example, grandfathering rules in the allocation of emission permits amongst firms can in principle neutralize distributional effects, but require verification of past emissions, which can generate ex-ante incentives for firms to increase emissions.

<sup>9</sup>In 2000, for example, truckers forced the UK government to roll back fuel taxes after successfully managing to disrupt automotive fuel distribution across the UK for almost a month.

<sup>10</sup>Much of the existing literature has focused on credibility problems arising from efficiency considerations only. An exception is Pearce and Stacchetti (1997), who analyze time-consistent taxation when a government cares about both efficiency and distribution.

<sup>11</sup>For an analysis of the interaction between trade and environmental policies in a game theoretical setting where international agreements are constrained to be self-enforcing, see Ederington (2001b).



jectives; however, since taxes are chosen after investment decisions are made, effects on investment will not be accounted for in this choice. This gives rise to a domestic commitment problem: in the absence of repeated interaction, the taxes chosen ex post by each government will be below their unilaterally optimal levels.

In this model, policy formation suffers from a *vertical* coordination problem between each government and its private sector, stemming from policymakers' inability to pre-commit to a certain level of emission taxation, as well as from a *horizontal* coordination problem between governments, stemming from the presence of transboundary emission spillovers.

This section describes the stage game, and examines policy formation within a single round of interaction. Subsequent sections will focus on repeated interaction over an infinite horizon.

## II.1 Pollution Abatement and Investment

In each country consumers consume a fixed amount  $X$  of a certain good, which can be produced by two alternative methods: a “dirty” technology, which produces one unit of the good at a constant marginal cost of unity while generating one unit of environmental emissions; and a “clean” technology, which generates no emissions but requires ex-ante investment.<sup>12</sup> If a total amount  $K$  of the good is produced using the clean technology, total domestic emissions are

$$Z = X - K. \tag{1}$$

The government levies a tax  $t$  per unit of emissions, which makes the gross-of-tax price of the polluting good and the net-of-tax price of its clean substitute both equal to  $p = 1 + t$ . Revenues from environmental taxation,  $R = tZ$ , are assumed to be returned to the consumers in equal shares in a lump-sum fashion.

---

<sup>12</sup>Investment in R&D and new equipment is the principal means by which pollution abatement takes place. Examples are the development of energy-efficient engine designs to reduce emissions by vehicles, the construction of refinery equipment to produce unleaded gasoline, or the installation of water cleaning equipment by chemical manufacturers. Estimates presented by the European Commission from studies carried out by several research institutions show that a European Carbon tax can only be effective in reducing CO<sub>2</sub> emissions if accompanied by substantial investment and innovation (DRI, 1992). Also, investment in emission-reducing projects plays a central role in the Kyoto Protocol through the Joint Implementation mechanism (Art. 6.1) and the Clean Development Mechanism (Art. 12).

There is a large number,  $n$ , of domestic firms having access to the clean technology. For each firm, the investment costs required to produce an amount  $k$  of the clean alternative is assumed to be quadratic in  $k$  and equal to

$$c(k) \equiv k + \frac{1}{2}\varphi k^2, \quad (2)$$

with  $\varphi > 0$ . In this formulation, the marginal cost of producing clean alternatives is always greater than that of producing the dirty alternative (unity) and is increasing in  $k$ .

If the private sector foresees a certain tax,  $t$ , the expected profits to a firm from producing an amount  $k$  of the clean good are

$$tk - \frac{1}{2}\varphi k^2, \quad (3)$$

and the first-order condition for a profit-maximizing abatement choice is

$$t - \varphi k = 0, \quad (4)$$

which implies that pollution abatement by a firm will take place up to the point where marginal abatement costs equal marginal abatement benefits. This identifies a function,  $k(t) = t/\varphi$ , linking the privately optimal level of investment in pollution abatement by each firm to the tax.

## II.2 Consumption, Damage, and Welfare

In each country, the population comprises  $h$  consumers of two types,  $A$  and  $B$ , each present in equal numbers ( $h/2$ ), and individually endowed with exogenous income levels equal to  $m^A$  and  $m^B$ , respectively.

Consumers all have equal stakes in production activities, implying that the total profits from abatement,  $S$ , are distributed uniformly in the population. Disposable income for each individual of type  $j$  is then  $y^j = m^j + (S + tZ)/h$ ,  $j = A, B$ . Consumption takes place in the second period, and it is assumed that individuals spend a fixed amount of their income on the pollution generating commodity— $x^j = \gamma^j X/h$ ,  $j = A, B$ , with  $\gamma^A + \gamma^B = 2$ —and spend the rest of their income on other non-polluting goods, in amounts equal to  $c^j = y^j - x^j$ .<sup>13</sup> In the rest of our discussion, we shall assume  $\gamma^A > \gamma^B$ .<sup>14</sup>

---

<sup>13</sup>Formally, such demand patterns are consistent with preferences that can be represented in terms of a utility function of the form  $u(x^j, c^j) = \lambda \min\{x^j - \gamma^j X/h, 0\} + c^j$ ,  $j = A, B$ , for  $\lambda$  sufficiently large.

<sup>14</sup>For example, group  $A$  could be identified with the rural population, who consume comparatively more automotive fuel.

Emissions are transboundary. The valuation of environmental damage by a representative domestic consumer is assumed to be linear in the global level of emissions:

$$\frac{\mu}{h}((1 - \alpha)Z + \alpha Z^*), \quad (5)$$

where  $Z^*$  denotes emissions by foreign firms, and where  $\alpha$  ( $0 \leq \alpha \leq 1$ ) represents the extent to which environmental damage is transboundary. Environmental damage is assumed to be additively separable in preferences. The (indirect) utility of consumers of type  $j$  in the home country can then be written as

$$u^j = m^j + \frac{-(1+t)\gamma^j X + S + tZ - \mu((1-\alpha)Z + \alpha Z^*)}{h}, \quad j = A, B. \quad (6)$$

We wish to model a situation where environmental taxes have undesirable distributional effects, i.e., where the distribution of welfare under  $t = 0$  is viewed by society as being desirable (so that emission taxes would not independently be used to pursue distributional objectives in the absence of environmental costs). This can be captured simply by specifying endowments as  $m^A = m^B + (\gamma^B - \gamma^A)X/h$ , so that  $t = 0$  implies  $u^A = u^B$ , and  $t > 0$  implies  $u^A < u^B$ . This means that any increase in  $t$  from zero will skew the distribution of welfare against group  $A$ .<sup>15</sup> If we then assume the government's objective to be a weighted linear combination of individual utilities,  $W(u^A, u^B) = h(w^A u^A + w^B u^B)$ , we can represent inequality aversion by attaching a premium  $\rho$  to the utility of the less favored group (type  $A$  if we assume  $\gamma^A > \gamma^B$ ) and specifying normalized weights as  $w^A = (1 + \rho)/(2 + \rho)$ ,  $w^B = 1/(2 + \rho)$ .<sup>16</sup> The government's payoff, as a function of the domestic tax and of the foreign level of abatement, can then be expressed as

$$\Pi(t, K(t), K^*) \equiv -\beta t - \frac{1}{2}\phi K(t)^2 + \mu((1 - \alpha)K(t) + \alpha K^*), \quad (7)$$

where  $\phi \equiv \varphi/n$ , and where  $K(t) = nk(t) = t/\phi$  is total domestic abatement when all domestic firms correctly forecast the tax,  $K^*$  is foreign abatement, and  $\beta = -\rho X(1 -$

---

<sup>15</sup>Much of the debate on the redistributive costs of environmental taxation revolves around the differential impacts these taxes can have across productive sectors, rather than consumers. A structure analogous to the one described here arises if ownership of the factors associated with the production of either the dirty or clean variety is concentrated within the economy. Then an increase in the tax would alter factor returns and would be distributionally nonneutral.

<sup>16</sup>This formulation can be derived from a hybrid Utilitarian/Rawlsian symmetric social welfare (or political support) function of the form  $W(u^1, \dots, u^h) = \min_i \{(1 + \rho)u^i + \sum_{l \neq h} u^l\}$ .

$\gamma^A)/(2 + \rho) > 0$  is a constant.<sup>17</sup> The term  $\beta t$  in (7) represents the distributional cost of emission taxes, an effect that we shall assume cannot be neutralized by any feasible compensation mechanism.<sup>18</sup>

In this framework environmental policy has both a direct effect on social welfare—because of its distributional costs—and an indirect effect—through the level of pollution abatement that it induces.

### II.3 Second-Best and Time-Consistent Emission Taxes

Consider first the case in which there is no coordination problem—between each government and its private sector and between the two governments. Vertical coordination would arise in a scenario where the government can rely on a commitment technology, i.e. where taxes are chosen, and credibly committed to, prior to investment taking place. Horizontal coordination is equivalent to a scenario in which  $\alpha = 0$ —i.e. there are no pollution spillovers between countries or the two governments manage to internalize the spillovers by choosing taxes in a coordinated manner. In this case, the globally efficient tax, denoted by  $t^E$ , will be selected:<sup>19</sup>

$$t^E = \mu - \beta\phi; \tag{8}$$

this is the tax for which the common payoff of the two countries is maximized.

Next, consider a scenario in which policymakers can credibly precommit to environmental taxation before investment decisions are made, but choose their policies unilaterally. Horizontal miscoordination between governments results in taxes that fail to inter-

---

<sup>17</sup>The full expression for social welfare also includes a constant term  $\Lambda = h(w^A m^A + w^B m^B) - (\gamma^A w^A + \gamma^B w^B)X + \mu((1 - \alpha)X + \alpha X^*)$ . For the sake of notational simplicity, we renormalize payoffs omitting the constant  $\Lambda$ ; this has no implications for our subsequent analysis, which only involves payoff differences.

<sup>18</sup>A consumption subsidy lowering the price of both the dirty good and its clean substitute could in principle neutralize the effects of the emission tax. Such a scheme, however, would not be feasible if the level of consumption of the clean substitute (reflecting how consumers substitute away from the polluting good) is either unobservable or nonverifiable; or it may not be viable if the social opportunity cost of the public funds required for such a subsidy is prohibitively high (e.g., because of high marginal efficiency costs from raising revenues through other taxes).

<sup>19</sup>In the rest of our analysis, we shall often refer to  $t^E$  as the optimal or efficient tax. However, the reader should keep in mind that this is a constrained (second-best) optimum; the unconstrained (first-best) optimum—a Pigouvian tax equal to  $\mu$ —could only emerge in the absence of distributional effects (i.e.  $\beta = 0$ ).

nalize the transboundary emission spillovers. Each government anticipates the effect of the policy on private choices and select a tax which maximizes (7). This yields

$$t^V = (1 - \alpha)\mu - \beta\phi, \quad (9)$$

where the subscript  $V$  refers to vertical coordination between the government and the investors. In this unilateral second-best solution, the concern for reducing the environmental damage associated with emissions is weighed against the distributional cost of higher emission taxes.<sup>20</sup>

Finally, consider a situation in which environmental policy suffers both from a horizontal coordination problem—stemming from the fact that the two governments act unilaterally—and a vertical coordination problem—resulting from a lack of coordination between each government and its private sector. In this scenario, the tension between efficiency and distributional goals combines with the dynamic dimension of pollution abatement to give rise to a time-inconsistency problem in the choice of emission taxes. The reason is that, once firms have installed a certain amount of investment, private abatement choices become unresponsive to changes in the tax; then, because of the adverse distributional costs of taxation, the government will be induced ex post to select a tax which is less than the one it would have committed to ex ante.

If the private sector's choice precedes the policy choice, the ex-post optimal policy level, for a given level of private investment will be simply

$$t^N = 0, \quad (10)$$

where the subscript  $N$  refers to the absence of any coordination (vertical or horizontal). Thus, if the government is unable to precommit to a given policy, it will be unable to adopt the efficient policy, and “promises” to do so will not be credible. Formally, suppose that investors foresee a tax  $\tilde{t}$ . Then each firm would invest  $k(\tilde{t}) = \tilde{t}/\psi$ . If the government selects a tax that maximizes welfare after a level of investment  $k(\tilde{t})$  is installed by each firm, it will be tempted to lower the tax so as to reduce its distributional effects. The government's ex-post optimal choice of tax, as a function of the tax that was expected by investors ex ante, is thus  $t(\tilde{t})$ ; in a perfect-foresight equilibrium, where investors correctly anticipate the ex-post optimal choice of the government, we must have  $\tilde{t} = t(\tilde{t})$ . It is easy to verify that the only value of  $\tilde{t}$  that satisfies the above fixed-point condition—the

---

<sup>20</sup>In the discussion that follows, we shall restrict our attention to scenarios featuring an interior unilateral solution where  $t^V = (1 - \alpha)\mu - \beta\phi > 0$ .

time-consistent level of taxation—is  $\tilde{t} = 0$ . Therefore, the unilateral second-best level of emission taxation cannot be sustained in the absence of a commitment mechanism.<sup>21</sup>

Notice that, consistently with our analysis' focus on the relationship between domestic policy credibility and international cooperation, the vertical relationship between government and investors is taken here as being exclusively domestic in nature: domestic policies have no effect on foreign investors, and, vice-versa, foreign policies have no effect on domestic investors.<sup>22</sup>

Comparing (8) with (9) and (10) it is straightforward to verify that, within a single round of interaction, any form of coordination (vertical or horizontal) leads to the adoption of higher taxes. Since payoffs are monotonically increasing in  $t$  for  $t < t^E$ , this implies that any form of coordination is beneficial, i.e. it results in a (weak) Pareto improvement:

**Proposition 1** *With a single round of interaction, partial coordination—either between governments or between governments and investors—results in a higher level of taxation and higher welfare.*

Notice that in our model the two forms of miscoordination operate in the same direction; both bias policies downwards, and cannot directly offset each other as they do, for example, in the problems studied by Rogoff (1985a), and Kehoe (1989). What this means is that coordination in either dimension—horizontal or vertical—can never be undesirable within a single round of interaction. In the next section, we will show that this result can be reversed when the interaction between governments and investors is repeated indefinitely.

### III Policy Formation under Repeated Interaction

Countries are typically unable to enter into binding coordination agreements—unless they enter into a political union—given that no enforcement power exists outside them. On

---

<sup>21</sup>Trivially, in a scenario where the second-best choice is  $t^V = 0$  (i.e., where  $(1 - \alpha)\mu - \beta\phi \leq 0$ ) no credibility problem arises.

<sup>22</sup>Other forms of policy intervention do not exhibit an exclusively domestic vertical structure. Consider, for example, import tariffs selected after investment in import-competing sectors (e.g. Matsuyama 1990; Tornell 1991), or capital tax rates selected after investment when capital is internationally mobile (e.g. Gordon 1986); in both cases, domestic policies also affect the choices of the foreign private sector. Nevertheless, as long as the domestic vertical policy linkages are stronger than the indirect vertical linkages—as is the case for import tariffs and capital taxes if there are mobility costs—the mechanisms we identify in our analysis will tend to dominate those indirect effects.

the other hand, institutional arrangements that can make domestic policy announcements binding are available in some cases (e.g. currency boards, guarantees of central bank independence, balanced budget rules). Nevertheless, even in the absence of any binding coordination arrangements, horizontal and vertical coordination can be achieved by repeated strategic interaction.

To rationalize the existence of international agreements between countries—as self-enforcing as opposed to binding arrangements—the international economics literature has appealed to the notion that, under repeated interaction, the threat of future punishments can be used in support of international cooperation. Intuitively, an agreement to maintain policy at the efficient level can be enforced if the one-time gain from cheating on the agreement is sufficiently small relative to the discounted future cost of the “policy war” that would be triggered as a consequence (Dixit 1987). The literature on policy credibility has independently appealed to the well-known idea that repeated interaction creates incentives to maintain reputation, and can therefore help overcome credibility problems, or at least mitigate them—an argument that was first formulated by Selten (1975).<sup>23</sup> As described in Stokey (1989), when the interaction between each government and its domestic sector is repeated indefinitely, time-inconsistency policy problems can be overcome by the credible threat of the private sector permanently reverting to the expectation of future inefficient policies.<sup>24</sup>

Both arguments are applications of folk theorems for repeated games. In such constructions, equilibrium strategies punish defections by switching to alternative continuation equilibria in which the defector experiences a lower average payoff than along the equilibrium path of play. Equilibrium strategies must therefore satisfy incentive constraints that involve a comparison between defection gains and punishment losses. Specifically, sustaining cooperation vertically in the domestic policy reputation game—in isolation from the international cooperation game—implies a comparison of the gains from vertical defections with the losses from vertical punishment; sustaining cooperation horizontally in the international policy game—in isolation from the domestic policy reputation game—

---

<sup>23</sup>Here the term “reputation” is used—somewhat loosely, but consistently with its use in some of the literature—to refer to policy credibility in the context of a game of complete information. For a discussion of reputation in games of incomplete information, see Fudenberg and Tirole (1996).

<sup>24</sup>Since investment projects are assumed to last only one period in the model, private agents are effectively finitely-lived players; nevertheless, their investment choices can be made to depend on past history.

implies a comparison of the gains from horizontal defections with the losses from horizontal punishment. In the problem we are analyzing, however, both a horizontal and a vertical reputation mechanism are at work. This effectively involves a pooling of sustainability constraints: the gains from deviations along both the horizontal and the vertical dimension must be weighed against the punishment that can be administered along both dimensions.

The theoretical implications of the pooling of incentive constraints under repeated interaction were first examined by Bernheim and Whinston (1990), for the case of oligopolistic firms sustaining collusion across multiple markets. What they show is that the effect of pooling is ambiguous, and that under some conditions pooling has no effect on cooperation. The form of pooling they analyze, however, is between different horizontal cooperation games, each involving a different strategic variable.<sup>25</sup> In our problem, pooling takes place instead between a horizontal and a vertical cooperation game that both involve the same strategic variable; to the best of our knowledge, this has not been studied before.

If fully eliminating both the vertical and the horizontal miscoordination problem were possible, then it would clearly always be desirable to do so—even under repeated interaction. But we are concerned here with situations where it is *not* possible to achieve full coordination within the stage game. Examining the relationship between the domestic policy credibility problem and the international cooperation problem under repeated interaction thus involves comparing a scenario where both coordination problems are simultaneously present in the stage game (the pooled case) with scenarios where only one type of miscoordination—horizontal or vertical—is at work (the unpooled cases). Given the specific structure of the incentive constraints that characterizes equilibria under repeated interaction, we should not expect pooling to have the same effect here as it has on the best responses in the stage game.

In the remainder of this section, we will look at purely hypothetical experiments in which coordination in the stage game—vertical or horizontal—is obtained by mechanically removing the corresponding coordination problem. Section IV will discuss how institutions that achieve coordination can be modeled explicitly.

---

<sup>25</sup>This is the same formal problem that has been studied by the literature on issue linkage in international economic relations (e.g. Ederington 2001b).



### III.1 Pooling and Defection Incentives

Folk theorems are fundamentally indeterminacy results; many equilibria are possible. In all cases, however, identifying equilibrium strategies involves a comparison of one-shot defection gains with the ensuing punishment. While the punishment depends on the particular punishment strategy invoked, the one-shot deviation incentives—and the effect that pooling has on them—are always the same irrespective of which punishment strategy one considers. Thus, any conclusions we obtain concerning the deviation phase have general applicability—unlike those concerning the punishment.

Consider first deviations from an agreed-upon policy,  $t^E$ , in the absence of horizontal coordination between governments and of vertical coordination between governments and their private sectors. If both governments keep to this choice, the policy level in both countries will be  $t^E$  in all periods. If, on the other hand, the government in one country contemplates a deviation from  $t^E$  in a certain period, such deviation would lie off the path of play anticipated by all other players; thus, unlike in the static analysis of Section II, an out-of-equilibrium deviation from  $t^E$  in any given period would not be anticipated by investors. This implies that, in the absence of domestic policy commitment, the ex-post optimal policy by a deviating government would be equal to  $t^N = 0$ , but total investment in pollution abatement in each country would still be equal to  $K(t^E)$ . Therefore, if a government were to deviate optimally from equilibrium play, it would experience a one-shot deviation gain equal to

$$\Gamma^N \equiv \Pi(t^N, K(t^E), K(t^E)) = \beta(\mu - \beta\phi). \quad (11)$$

Suppose instead that governments can commit to policy announcements in each round before private investment choices are made (vertical coordination) by relying on certain institutionally available commitment devices (we will discuss later, in Section IV, how such commitment devices can be explicitly modeled). In this case, if a government were to deviate from equilibrium play in a certain period, such a deviation would be accounted for by its investors. The optimal deviation policy would then be  $t^V = (1 - \alpha)\mu - \beta\phi$ , and the associated deviation gain would be equal to

$$\Gamma^V \equiv \Pi(t^V, K(t^V), K(t^E)) = \frac{\alpha^2\mu^2}{2\phi}. \quad (12)$$

Does domestic policy commitment increase or reduce the one-shot incentives to deviate from the optimal policy  $t^E$ ? To answer this question we can look at the effect of vertical coordination on deviation incentives, as captured by the difference between  $\Gamma^N$

and  $\Gamma^V$ —the deviation gains for the pooled and unpooled (vertically coordinated) cases respectively. If this difference is positive, then we can conclude that, other things equal (more specifically, abstracting from its effects on punishment) domestic policy commitment will make sustaining  $t^E$  easier.

The above difference can be written as

$$\begin{aligned}
\Gamma^N - \Gamma^V &= \beta(t^V - t^N) \\
&\quad - \frac{1}{2}\phi(K(t^E)^2 - K(t^V)^2) + \mu(1 - \alpha)(K(t^V) - K(t^E)) \\
&= \beta((1 - \alpha)\mu - \beta\phi) \\
&\quad - \frac{\alpha\mu}{2\phi}(\alpha\mu - 2\beta\phi).
\end{aligned} \tag{13}$$

The first term of expression (13) is always positive and identifies a “direct surprise effect” of noncommitment, consisting of the additional reduction in the direct distribution costs of the policy. The second line of (13) is an “investment surprise effect,” stemming from the fact that without commitment a defection surprises investors, resulting in a level of abatement  $K(t^E)$  in the deviation phase which is higher than the level  $K(t^N)$  that would have occurred if a deviation had been anticipated; in turn, this implies both higher abatement costs for firms and higher environmental damage. When negative—which happens whenever  $\alpha > 2\beta\phi/\mu$ —this latter effect can be thought of as the fraction of the net cost of cooperation which remains sunk in the deviation phase when a government cannot coordinate its deviation with its own private sector. This net sunk cost can outweigh the direct surprise gain, so that the ability to commit domestically makes a policymaker more tempted to deviate. For this to be possible, the investment surprise effect must be negative, i.e. cooperation must take each country to a point where each country would individually gain from unilaterally lowering its level of private investment, even when neglecting the direct effect of the policy. In turn, this means that the size of the international spillover must be sufficiently large.

**Proposition 2** *When efficient policies must be sustained by repeated strategic interaction, for large enough international spillovers, domestic policy commitment implies a greater temptation to deviate from optimal policies.*

PROOF: We have

$$\frac{\partial(\Gamma^N - \Gamma^V)}{\partial\alpha} = -\frac{\alpha\mu^2}{\phi} < 0, \tag{14}$$

which implies that  $\Gamma^N - \Gamma^V$  is monotonically decreasing in  $\alpha$ . Moreover, we can show that this difference is positive for  $\alpha = 0$  (where  $t^V = t^E$ ):

$$(\Gamma^N - \Gamma^V)_{\alpha=0} = \beta(\mu - \beta\phi) > 0; \quad (15)$$

the reason for this result is that, in the absence of international spillovers, the lack of commitment generates only the surprise gain associated with the reduction in the distribution costs of environmental taxation. The difference  $\Gamma^N - \Gamma^V$  is instead negative for  $\alpha = 1$  (where  $t^V = t^N$ ):<sup>26</sup>

$$(\Gamma^N - \Gamma^V)_{\alpha=1} = \beta\mu - \frac{\mu^2}{2\phi} - \beta^2\phi < 0. \quad (16)$$

Hence, there exists an interval  $(\underline{\alpha}, 1]$ ,  $\underline{\alpha} > 0$ , over which it is negative.  $\square$

Thus, in the absence of full coordination, vertical coordination alone can be undesirable, because it can make a country more effective at defecting from an international agreement. The reason why vertical miscoordination affects the deviation phase (Proposition 2) and the stage game equilibrium (Proposition 1) in different ways is simply that, in the former, the deviation from the efficient policy level is not anticipated by investors. This implies both a benefit and a cost for the deviating government: the distributional cost of the policy can be reduced without consideration for the associated effects on investment; at the same time, investors are unable to reduce investment to its domestically optimal level. If the second effect dominates the first, a reduction in the extent of the horizontal coordination problem is also undesirable, i.e. vertical miscoordination can help to sustain efficient policies.<sup>27</sup>

This is more likely to be the case the larger is the international policy spillover, because a larger international spillover entails a comparatively larger cost of miscoordination between a government and its private sector during unanticipated deviations. If so, then a reduction in the extent of the horizontal coordination problem is also undesirable, i.e. the need to sustain policy reputation vis-à-vis the private sector and the need to sustain self-enforcing cooperation with other countries are complementary in helping to sustain efficient policies:

---

<sup>26</sup>If we look at the case of “global” emissions ( $\alpha = 1/2$ ), the difference  $\Gamma^N - \Gamma^V$  is negative whenever the investment surprise effect associated with a deviation from agreed upon policy in the absence of vertical coordination is negative. This requires the distribution costs to be low enough, i.e.  $\beta < \mu/(4\phi)$ .

<sup>27</sup>In the Appendix, we show that this result carries over to a framework in which the degree of vertical miscoordination is allowed to vary continuously.

**Proposition 3** *When efficient policies must be sustained by repeated strategic interaction, the temptation to deviate under vertical coordination becomes comparatively larger, relative to that under no coordination, the larger is the size of the international spillover.*

PROOF: In the proof of Proposition 3, it has been shown that  $\partial(\Gamma^N - \Gamma^V)/\partial\alpha < 0$ , establishing the result.  $\square$

Whether or not a lack of domestic policy commitment results in smaller deviation gains will also depend on the size of the direct domestic effect of the policy, which depends on the size of the distributional costs of environmental policy: the smaller these costs are, the less likely it is for the direct surprise effect of noncommitment to dominate the investment surprise effect.

**Proposition 4** *The minimum level of international spillover above which lack of domestic policy commitment lowers the temptation to deviate from efficient policies is smaller the smaller are the direct domestic effects of the policy.*

PROOF: The spillover level  $\underline{\alpha}$  for which defection gains are just as high with policy commitment than without it is identified by  $\Gamma^N - \Gamma^V = 0$ . Totally differentiating this condition with respect to  $\alpha$  and  $\beta$ , we obtain  $d\underline{\alpha}/d\beta = -(\partial(\Gamma^N - \Gamma^V)/\partial\beta) / (\partial(\Gamma^N - \Gamma^V)/\partial\alpha) = \phi(\mu - 2\beta\phi)/(\alpha\mu^2) > 0$ . Moreover, with  $\beta = 0$ , policy deviations from  $t = 0$  produce no direct effect—i.e. there are no distributional costs—and the direct effect of the policy is increasing in  $\beta$ . As  $\beta$  approaches zero,  $\Gamma^N - \Gamma^V$  approaches  $-\alpha^2\mu^2/(2\phi)$ , which is negative for any value of  $\alpha$ . This implies that for any given level of spillover, no matter how small, there will exist scenarios in which noncommitment results in lower deviation gains.  $\square$

This result implies that, other things equal, noncommitment will be more likely to discourage defections precisely under the conditions for which the efficiency cost of noncommitment in a one-shot, perfect-foresight equilibrium are largest: when the direct effects of the policy are large, the commitment policy,  $t^V$ , will be close to  $t^N$ , and so vertical miscoordination will have little effect on investment; if instead the direct effects of the policy are small, vertical miscoordination will produce a larger gap between  $t^V$  and  $t^N$  and thus a larger effect on investment.

The main focus—and novelty—of our analysis concerns the effects of vertical coordination in the stage game on the sustainability of efficient policies under repeated interaction. There are good reasons for this emphasis: unlike for vertical coordination, no institutional means of achieving fully binding horizontal coordination are available, short of creating

a political union; thus, in the context of the problem we are studying, the parallel between vertical and horizontal coordination within the stage game is merely formal. If we nevertheless consider the effects of horizontal coordination in the stage game (equivalent to the case  $\alpha = 0$ ), we can note that equation (11) is independent of  $\alpha$ ; this implies that, in the absence of domestic policy commitment, horizontal coordination generates no effect on the deviation gains. In this model, the effect of horizontal coordination on the sustainability of  $t^E$  depends exclusively on its effect on the punishment, as discussed below.

### III.2 Pooling and Punishment

Domestic policy commitment will also influence the sustainability of efficient policies via the effect it produces on the punishment that follows defections. However, the effect of incentive pooling on punishment will depend on the particular type of punishment strategy adopted.

Consider, for example, how vertical coordination affects Nash-reversion punishment strategies—whereby a deviation from  $t^E$  triggers indefinite reversion to noncooperation. Lack of domestic policy commitment implies not only that the other government would cease to select policies cooperatively following a defection, but also that investors in each country would mistrust their governments forever. This course of action—with policy levels  $t^N$  in the reversion phase being fully anticipated by investors—would result in the lowest possible equilibrium level of policy in all periods of the continuation game; and, since payoffs are monotonically increasing in  $t$ , this is the worst punishment that can be administered through a stationary punishment strategy in the continuation game. In this case, if the international policy spillover is larger than the critical level that has been identified in the proof of Proposition 3, lack of domestic commitment unambiguously makes it easier to sustain optimal policies, both because it decreases gains in the deviation phase and because it increases the punishment following a deviation. In the presence of a smaller spillover, however, vertical coordination in the stage game raises both the temptation and the punishment, so its combined effect would be ambiguous.

Yet, despite the fact that multiple types of punishment strategies are in principle possible in equilibrium, there are compelling theoretical reasons to view some of them as being more plausible. In particular, punishment strategies relying on continuation equilibria that are Pareto dominated by other feasible alternative equilibria seem implausible, because they might be vulnerable to renegotiation. For example, although Nash rever-

sion represents a credible threat—since playing  $t^N$  indefinitely is always an equilibrium strategy in the continuation game—once a defection has occurred, players will have incentives to forgo punishment and re-coordinate to an equilibrium that gives them all a higher continuation payoff. In the remaining of this section we shall show that, if we restrict our attention to renegotiation-proof equilibrium strategies—strategies that leave no joint incentive to renegotiate to an alternative continuation equilibrium upon entering the punishment phase (Farrell and Maskin 1989)—we obtain clear-cut predictions about the effects of removing vertical or horizontal miscoordination on the sustainability of efficient policies.

It can be shown that the following strategy profile is renegotiation-proof: each country plays the efficient policy  $t^E$  as long as the other country does the same; if country  $i$  defects in a given period (and country  $j$  does not), then country  $j$  (the punisher) will play  $t^V$  until the defector country  $i$  (the defector) reverts to  $t^E$ ; as soon as country  $i$  has “repented” by playing  $t^E$ , country  $j$  “forgives” the initial defection and returns to playing  $t^E$ .<sup>28</sup> These punishment strategies identify a symmetric renegotiation-proof equilibrium *independently* of whether or not the vertical coordination problem is present, since a policy choice of  $t^N$  by the punisher in the punishment phase would be Pareto dominated by  $t^V$ . In this case, the only effect of vertical coordination would be through the deviation gains, and so Proposition 2 translates directly into predictions about the effect of vertical coordination on the sustainability of efficient policies.

**Proposition 5** *When efficient policies must be sustained by renegotiation-proof punishment strategies under repeated strategic interaction, for large enough international spillovers, domestic policy commitment raises the minimum discount factor for which efficient policies can be sustained.*

PROOF: For the above strategy to be a subgame perfect, renegotiation-proof equilibrium, the following condition must be satisfied: (i) the punisher must be better off during punishment than under cooperation (this condition is what especially distinguishes a renegotiation-proof equilibrium), which is trivially satisfied by the strategies described; (ii) the reduction in the continuation payoff experienced by defectors (the punishment) must be large enough, in discounted terms, to discourage defections; for the strategies described, the drop in payoff experienced by a defector in the punishment phase equals  $\Omega = \alpha^2 \mu^2 / \phi$ ; denoting with  $\delta$  the discount factor, the

---

<sup>28</sup>These punishment strategies have a structure analogous to that first described by Van Damme (1989) for a discrete prisoners’ dilemma game.

no-defection condition is thus  $\Gamma \leq \delta\Omega$ —where  $\Gamma = \Gamma^V$  with policy commitment and  $\Gamma = \Gamma^N$  without policy commitment; (iii) the punishment must be credible, i.e. the continuation equilibrium must be subgame perfect, which means that the players must have no incentive to deviate from the stated punishment; if the defector optimally deviates from  $t^E$  during the punishment phase, it experiences a one-shot gain equal to either  $\Gamma^V$  or  $\Gamma^N$ , but suffers a loss equal to  $\delta\Omega$  (in discounted terms) from the postponement of reversion to cooperation. Conditions (ii) and (iii) are therefore formally identical in this game, and the minimum discount factor for which  $t^E$  can be sustained in this way is identified as  $\underline{\delta} = \Gamma/\Omega$ .<sup>29</sup> Whether this value is smaller or larger in the presence of policy commitment depends directly on the comparison between the deviations gains  $\Gamma^V$  and  $\Gamma^N$ .  $\square$

Results analogous to Propositions 3 and 4 also immediately follow.

If we next consider the effect of horizontal coordination on punishment—with the caveats discussed earlier concerning the real-world relevance of such an exercise—for the case of Nash-reversion punishment strategies, we conclude that horizontal coordination in the stage game has no effect on either the deviation gains or the punishment: the Nash-reversion policy is  $t^N$ , independently of  $\alpha$ .<sup>30</sup> With renegotiation-proof punishment strategies, however, horizontal coordination ( $\alpha = 0$ ) implies that no degree of patience is high enough to solve the domestic policy credibility problem, since there are no credible punishment strategies giving investors a higher continuation payoff, once punishment is triggered, than the payoff they would obtain by renegotiating a reversion to cooperation jointly with government. Therefore, in the absence of any strategic interaction with a foreign country, any policy level above  $t^N$  would be unsustainable.

---

<sup>29</sup>Conditions (ii) and (iii) identify the minimum discount factor  $\underline{\delta}$  for which the efficient policies are sustainable. For this critical discount factor—the focus of Proposition 5—the punishment strategies described above are *strongly* renegotiation-proof in the sense of Farrell and Maskin (1989). In scenarios where the discount factor exceeds  $\underline{\delta}$ , there might be alternative continuation equilibria that Pareto dominate the one described (and so the punishment strategies we have described would only be *weakly* renegotiation-proof, i.e. undominated by alternative strategies only within a set suitably defined). It should be stressed, however, that the result obtained here for  $\underline{\delta}$  holds independently of which notion of renegotiation proofness we consider for scenarios where  $\delta > \underline{\delta}$ .

<sup>30</sup>We can view this result as being somewhat analogous to the “irrelevance result” described by Bernheim and Whinston (1990) for the horizontal (multimarket) pooling of incentive constraints under constant-returns-to-scale technologies.

## IV Institution Design

There are a number of institutional commitment devices that governments can rely upon to achieve some degree of binding policy commitment—such as policy delegation or budgeting rules—many of which have been extensively studied in the literature.<sup>31</sup> The analysis carried out in the preceding sections allows us to draw implications for the structure of institutions—not necessarily in a normative sense, but rather thinking of which institutions we would expect to arise through the choice of independent governments.

Institutions that achieve binding coordination in policy choices—vertically or horizontally—can be formally represented as follows. Suppose we distinguish, in each round of interaction, the choice of a central decision-making authority—which selects common policies with the objective of maximizing the joint payoff of the two countries—and the individual policy choices of the two countries—which may or may not coincide with the centralized choice, depending on the strength of the horizontally coordinating institutions. Similarly, we can distinguish, for both centralized and decentralized decisionmakers, ex-ante announcements and ex-post policy choices—which may or may not coincide with ex-ante announcements depending on the strength of the policy commitment institutions. We can thus imagine the following sequence of moves in each round: (i) the central authority announces a common policy choice,  $t_C^A$ ; (ii) the individual policymakers choose individual policy announcements,  $t_i^A$ ,  $i = 1, 2$ ; (iii) investors make their investment choices; (iv) the central authority selects an ex-post common policy choice,  $t_C$ ; (v) the individual policymakers choose individual policies,  $t_i$ ,  $i = 1, 2$ .

The strength of the horizontal coordination institutions can be represented in terms of an additional penalty,  $L^H$ , that countries incur for deviating from central policy choices either at the announcement stage or at the ex-post stage; while the strength of the vertical coordination institutions can be represented in terms of a penalty,  $L^V$ , for deviating from announcements, either at the central level or at the country level.<sup>32</sup> Then, depending on the size of the penalties  $L^H$  and  $L^V$ , outcomes will correspond to those of scenarios without any form of coordination in the stage game, with vertical coordination, with horizontal coordination, or with full coordination.<sup>33</sup>

---

<sup>31</sup>The role of delegation of authority in monetary policy was first studied by Rogoff (1985b). Lucas and Stokey (1983) were the first to study debt structure as a commitment device.

<sup>32</sup>This specification could be generalized by allowing for the cost of deviations to depend on the size of the deviation.

<sup>33</sup>Horizontal coordination here only means that countries coordinate their choices in each round without



In practice, the possibility of achieving binding horizontal coordination (as we characterize it above) between sovereign countries is extremely limited. On the other hand, as already noted, countries *can* rely on domestic policy commitment devices, and the choice of whether or not to do so can be thought of as being endogenous.<sup>34</sup>

This endogenous choice can be modeled as follows. Suppose that the penalties for vertical deviations can be country-specific. Also, suppose that these penalty levels are selected by each country at an initial stage which precedes all other moves in each round (stage zero), and that they are fully observable immediately after they are selected. Then, if a symmetric policy choice of  $t^E$  requires a certain common choice of  $L^V$  in all periods to be sustainable, this institutional choice could itself be sustained by the threat of punishment as part of an efficient equilibrium. Moreover, the efficient choice of  $L^V$  along the path of equilibrium play will be unrelated to the corresponding efficient choice in the punishment phase, which implies that the choice of policy commitment institutions in equilibrium will be dictated by the same conditions described in Proposition 2 (i.e. independently of the punishment strategies adopted):

**Proposition 6** *When efficient policies must be sustained by repeated strategic interaction and the choice of whether or not to make binding domestic policy announcements is endogenous, for large enough international spillovers, supporting efficient policies may require countries to choose—in equilibrium—not to make domestically binding policy commitments.*

PROOF: The ability of selecting  $L^V$  endogenously will dictate a particular institutional choice along the path of equilibrium play, so as to minimize the gains from unilateral deviations. Let the country-specific penalty levels be denoted by  $L_i^V, i = 1, 2$ , and suppose that  $L_i^V$  can take on one

---

being directly tied to efficient policies—as would be the case, for example, under a political union. One may think of international institutions as being able to achieve more than just period-by-period horizontal policy coordination if other forms of punishment, along different policy dimensions, can also be used. Note, however, that the horizontal pooling of policy games does not necessarily facilitate enforcement, since both defection incentives and punishments become pooled—just as in Bernheim and Whinston’s (1990) analysis. The conditions for cross-issue policy linkage to facilitate sustainability of efficient outcomes under repeated strategic interaction have been studied by Spagnolo (2001).

<sup>34</sup>Domestic institutions for policy commitment can also be used, to some extent, in direct support of international agreements. This is the case, for example, when agreements are ratified by domestic legislators—which, to the extent that domestic law is binding on policymakers, directly raises the cost of noncompliance in the short run (until the ratification is “undone”).

of two values, either zero or a value,  $\underline{L}^V$ , large enough that deviations from announced policies could never generate a gain. Then, if  $\alpha \leq \underline{\alpha}$  and noncommitment results in larger deviation gains (Proposition 2), a common choice of  $L_i^V \geq \underline{L}^V$  would make it easier to sustain  $t^E$ ; whereas if  $\alpha > \underline{\alpha}$ , sustaining  $t^E$  will be easier for  $L_i^V = 0$ . An efficient choice of  $L_i^V = 0$  along the equilibrium path can always be sustained: any individual deviation from a prescribed common choice of  $L_i^V = 0$  in stage zero of any given round (which would immediately be observed by all players) could trigger immediate punishment in the subsequent stages of the same round—with the government in the other country selecting a policy level  $t^N$  in that round, and its private sector correctly anticipating this choice by selecting  $K(t^N)$ ; then the deviating country could never gain from such a deviation. The selection of  $L^V$  in the punishment phase that follows policy deviations will be dictated by the need to support continuation equilibria that can in turn best support efficient policies in the cooperation phase. Which institutional choice in the punishment phase is most effective, and whether or not it can be sustained, is fully independent of which institutional choice is most effective in the cooperation phase.  $\square$

In conjunction with Propositions 3 and 4, this result also implies that countries should be comparatively less prone to rely on domestic institutional commitment mechanisms for policies where the level of spillover is comparatively larger and/or where the direct effect of the policy is comparatively small. In support of this prediction one could argue, for example, that international trade and climate change—where typically only weak forms of domestic policy commitment are used—are in fact areas where policies produce comparatively larger transboundary spillovers (relative to their corresponding domestic effects) than is the case for, say, monetary and fiscal policies—where we do often observe commitment devices being employed.

## V Conclusion

It is well understood that in noncooperative games partial coordination amongst a subset of players can give rise to an even less efficient noncooperative outcome.<sup>35</sup> Our analysis can be thought of as providing a counterpart of that general principle in a setting with repeated interaction, where cooperative choices must be supported by noncooperative equilibrium strategies: under repeated interaction, partial binding coordination can be an

---

<sup>35</sup>Krugman (1991), for example, applies this idea to preferential international trade policy coordination, and shows that the simultaneous formation of customs unions can lower welfare.

obstacle to supporting efficient policies, independently of whether or not coordination is beneficial when interaction is not repeated.

In the policy game we have analyzed, two different forms of miscoordination are simultaneously present—vertical and horizontal—both operating in the same direction within the stage game; consequently, partial coordination always results in more efficient policies in a single round of interaction. Nevertheless, when the game is repeated indefinitely, partial coordination in the stage game can become counterproductive. Thus, the effects of partial binding coordination under repeated interaction do not parallel its effects within a single round of interaction.

Under repeated interaction, lack of domestic policy commitment can facilitate international cooperation because it impedes vertically-coordinated horizontal defections. International policy spillovers can help because they generate larger vertical miscoordination costs during deviations. It follows that, when policies must be self-enforcing, countries may voluntarily forgo the use of institutional mechanisms for achieving partial binding coordination even if these are available. The need to maintain policy reputation through repeated interaction with the private sector can indeed help support international cooperation, implying that not “tying one’s hands” by means of binding commitment mechanisms can actually boost prospects for international cooperation. Thus, absent a system of institutions that can eliminate all forms of miscoordination, international policy coordination may be better served by weak domestic institutions. While the potentially adverse effects of partial horizontal coordination have long been understood in the literature—for example, the threat that regional agreement can pose to multilateral cooperation—the threat that domestic policy commitment can pose to international cooperation had not before been identified.<sup>36</sup>

Some writers have argued that lack of policy commitment can make a country a more willing partner in an agreement because such a country has comparatively more to gain from sustaining cooperative policy levels. However, our analysis shows that what matters for the sustainability of international cooperation is not only how much a country has to gain from solving the commitment problem, but also how much the miscoordination between policymakers and investors helps or impedes gainful defections—and under a

---

<sup>36</sup>The idea that reputation in one vertical relationship may positively spill over to other vertical relationships is implicit in the analysis of Cole and Kehoe (1998)—via a very different mechanism from the one we describe in this paper—but has never before been considered for the case of interlinked vertical and horizontal relationships.

renegotiation-proofness restriction this is indeed all that matters.

Is policy credibility imported or exported? The results of our theoretical discussion suggest that the answer to this question hinges on how we interpret domestic policy credibility and international cooperation—as binding or as self-supporting. Under the latter interpretation, our analysis has shown that credibility can both be imported from a self-enforcing international agreement and exported from a government’s ability to sustain policy reputation at home. Moreover, importing and exporting credibility are not reciprocally exclusive possibilities. International agreements are therefore not simply a direct source of outside commitment as earlier literature has characterized them; rather, self-enforcing international agreements and the need to sustain policy reputation vis-à-vis the private sector can complement each other in helping to support overall coordination to efficient policies. The possibility of such co-causation represents a distinctive prediction of a model where reputation and international cooperation must be attained by repeated strategic interaction, and would be at odds with a model where international agreements are binding.

Evidence on the relationship between domestic policy commitment and international policy cooperation is mixed, as our results would indeed suggest (Proposition 2). For example, if we compare the members of the Kyoto Protocol against non-members according to the World Bank indicator of institutional credibility (Brunetti *et al.* 1998), we find that, although members score higher on average, credibility scores vary widely within each group.<sup>37</sup> A similar pattern emerges when looking at the members of the WTO and at the countries that have ratified the Non-Proliferation Treaty.<sup>38</sup>

Our analysis also predicts that the relationship between domestic policy commitment and international cooperation should vary with the size of the international policy spillovers: countries with more credible institutions should be comparatively more likely to make better international partners the smaller are the international spillovers of domes-

---

<sup>37</sup>This index—ranging from a minimum of 1 to a maximum of 6—is meant to measure the credibility of governments’ policy announcements. It was constructed by the World Bank and the International Finance Corporation on the basis of a private sector survey conducted during 1996-1998 in seventy-four countries. The average credibility index of the countries that have ratified the Kyoto Protocol is 3.9, with a standard deviation of 0.6; for non-ratifiers the average index is 3.6, with a standard deviation of 0.6.

<sup>38</sup>The average credibility index is 3.6 for WTO members and 3.0 for non-members, with a standard deviation of 0.7 and 0.6, respectively. In the case of the Non-Proliferation Treaty, the average credibility index for ratifying countries is 3.8, with a standard deviation of 0.6; the average index for non-ratifiers is 3.7, with a standard deviation of 0.9.

tic policy (Proposition 3). Looking again at evidence on institutional credibility and membership in international treaties, we do find patterns that appear to be aligned with this prediction: if we rank countries according to their degree of openness, and compare countries in the second and third quartile of the distribution so obtained, we find that the credibility index differs significantly between WTO members and non-members only for those countries that are characterized by a comparatively lower degree of openness.<sup>39</sup> However, without a more systematic empirical investigation—accounting for the effect of other variables and addressing measurement and endogeneity problems—we cannot exclude that there may be alternative explanations for this pattern.

A richer analytical framework would be required to address a number of aspects that we have abstracted from in our analysis. For example, our analysis has not addressed the implications of asymmetries between countries with respect to policy spillovers or with respect to the severity of the policy credibility problem they face.<sup>40</sup> Also, if international cooperation is multilateral, rather than bilateral, then how partial binding coordination affects cooperation also depends on how it affects countries' ability to coordinate joint punishments against defectors—an aspect that is not captured by our two-country model.

## Appendix: Continuous Degree of Commitment

The results described in the text carry over if we allow the degree of vertical miscoordination to vary continuously. Suppose that a fraction  $1 - \theta$  of firms invest after the policy is selected, and a fraction  $\theta$  of firms invest beforehand. Then,  $\theta$  will capture the degree of vertical miscoordination. In this case, if the government in one country contemplates a deviation from  $t^E$  to  $t^D$  in a certain period, such deviation would be anticipated by a fraction  $1 - \theta$  of firms, which will each choose a level of pollution abatement equal to  $t^D/\varphi$ , but would lie off the path of play anticipated by

---

<sup>39</sup>The openness indicator used is the sum of merchandise exports and imports divided by the value of GDP, all in current U.S. dollars (constructed using WTO and World Bank GDP estimates). The mean credibility index of member countries in the second quartile (3.5) is statistically higher than that the mean of non-members (2.7) in the same quartile at the ten percent level. However, if we compare the credibility index of the member and non-members of the WTO in the third quartile, we find no statistically significant difference. The lowest and highest quartiles contain too few non-WTO member countries for any comparison to be statistically significant.

<sup>40</sup>The role of these asymmetries in a trade policy application is considered by Conconi and Perroni (2004). Self-enforcing trade agreements between countries asymmetric size—in the absence of domestic commitment problems—are also the focus of Bond and Park (2004).

the remaining fraction  $\theta$  of firms, which would keep pollution abatement at  $t^E/\phi$ . In this case, it can easily be shown that the optimal deviation tax is given by

$$t^D = \max \left\{ 0, \mu(1 - \alpha) - \frac{\beta\phi}{1 - \theta} \right\} < t^E. \quad (17)$$

If the deviation tax is positive, the deviation gain can be written as

$$\Gamma(\theta, \alpha) = \frac{(\beta\phi\theta + \alpha(\mu - \mu\theta))^2}{2\phi(1 - \theta)}. \quad (18)$$

The sign of the total derivative of (18) with respect to  $\theta$  is generally ambiguous. However, if we look at the cross partial derivative of the deviation gain with respect to  $\theta$  and  $\alpha$ , we find

$$\frac{\partial(\partial\Gamma(\theta, \alpha)/\partial\theta)}{\partial\alpha} = \frac{\partial^2\Gamma(\theta, \alpha)}{\partial\theta\partial\alpha} = \mu \left( \beta - \frac{\alpha\mu}{\phi} \right), \quad (19)$$

which is negative in any scenario in which there exists a commitment problem—since this implies  $t^V > 0$  and  $\beta < \alpha\mu/\phi$ . This leads to results analogous to Propositions 2 and 3; i.e. a marginal increase in the degree of vertical miscoordination discourages defections as long as the international spillover is sufficiently large, and more so the larger the spillover is. A result analogous to Proposition 4 can also be proven.

## References

- Bagwell, K., and R. Staiger (1997). “Multilateral Tariff Cooperation during the Formation of Regional Free Trade Areas,” *International Economic Review* 38: 291-319.
- Bajona, C. and T. Chu (2004) “China’s WTO Accession and Its Effect on State-Owned Enterprises,” mimeo.
- Bernheim, D., and M. Whinston (1990). “Multimarket Contact and Collusive Behavior,” *Rand Journal of Economics* 21: 1-26.
- Bond, E., and J-H. Park (2004). “Gradualism in Trade Agreements with Asymmetric Countries,” *Review of Economic Studies* 69, 379-406.
- Brown Weiss, E. and H.K. Jacobson (1997). “Compliance with International Environmental Accords,” in M. Role, H. Sjöberg and U. Svedin (eds), *International governance on environmental issues*, Kluwer, Dordrecht, 78-110.
- Brunetti, A., G. Kisunko, and B. Weder (1998). “Credibility of Rules and Economic Growth: Evidence from a Worldwide Survey of the Private Sector,” *World Bank Economic Review* 12: 353-84.

- Cole, H., and P. Kehoe (1998). "Models of Sovereign Debt: Partial versus General Reputations," *International Economic Review* 39: 55-70.
- Conconi, P., and C. Perroni (2004). "The Economics of S&D Trade Regimes," CEPR Discussion Paper No. 4508.
- Dixit, A. (1987). "Strategic Aspects of Trade Policy," in Bewley, T. (ed.) *Advances in economic theory: Fifth World Congress*, Cambridge University Press, Cambridge, MA: 329-62.
- DRI (1992). *Impact of a package of EC measures to control CO<sub>2</sub> emissions on European industry*, European Commission (DGXI).
- Ederington, J. (2001a). "International Coordination of Trade and Domestic Policies," *American Economic Review* 91: 1580-93.
- Ederington, J. (2001b). "Global Environmental Agreements and Trade Sanctions," University of Miami, mimeo.
- Farrell, J., and E. Maskin (1989). "Renegotiation in Repeated Games," *Games and Economic Behavior* 1: 327-60.
- Fudenberg, D., and J. Tirole (1996). *Game theory*, MIT Press, Cambridge, MA., and London, UK.
- Gordon, R. (1986). "Taxation of Investment and Savings in the World Economy," *American Economic Review* 76: 1086-102.
- Kehoe, P. (1989). "Policy Cooperation Among Benevolent Governments May Be Undesirable," *Review of Economic Studies* 56: 289-96.
- Keohane, R.O. (1995). "Compliance with Environmental Standards: Environmental Case Studies," in J.L. Hargrove (ed.), *Proceedings of the eighty-ninth annual meeting of the American Society of International Law*, Hein & Co., Buffalo, NY, 206-224.
- Krugman, P. (1991). "Is Bilateralism Bad?" in Helpman, E., and A. Razin (eds.) *International trade and trade policy*, MIT Press, Cambridge, MA.
- Kydland, F., and E. Prescott (1977). "Rules rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85: 473-91.
- Laffont, J.-J., and J. Tirole (1996). "Pollution Permits and Environmental Innovation," *Journal of Public Economics* 62: 127-40.
- Lucas, R., and N. Stokey (1983). "Optimal Fiscal and Monetary Policy is an Economy without Capital," *Journal of Monetary Economics* 12: 53-93.
- Maggi, G., and A. Rodriguez-Clare (1998). "The Value of Trade Agreements in the Presence of Political Pressures," *Journal of Political Economy* 106, 574-601.
- Maggi, G. (1999). "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review* 89: 190-214.

- Matsuyama, K. (1990). "Perfect Equilibria in a Trade Liberalization Game," *American Economic Review* 80: 480-92.
- Morgan, T., and S. Campbell (1991). "Domestic Structure, Decisional Constraints, and War: So Why Kant Democracies Fight?" *Journal of Conflict Resolution* 35: 187-211.
- Pearce, D., and E. Stacchetti (1997). "Time Consistent Taxation by a Government with Redistributive Goals," *Journal of Economic Theory* 72: 282-305.
- Poterba, J. (1991). "Tax Policy to Combat Global Warming: On Designing a Carbon Tax," in Dornbusch, R., and J.M. Poterba, (eds.), *Global warming: Economic policy responses*, MIT Press, Cambridge, MA.
- Persson, T., and G. Tabellini (1994). *Monetary and fiscal policy, Volume I: Credibility*, MIT Press, Cambridge, MA.
- Organisation for Economic Co-operation and Development (1995). *Climate change, economic instruments and income distribution*. OECD, Paris.
- Rogoff, K. (1985a). "Can International Monetary Policy Cooperation Be Counterproductive?" *Journal of International Economics* 14: 199-217.
- Rogoff, K. (1985b). "The Optimal Degree of Commitment to an Intermediate Monetary Target," *Quarterly Journal of Economics* 100: 1169-90.
- Selten, R. (1975). "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory* 4: 22-55.
- Spagnolo, G. (2001). "Issue Linkage, Credible Delegation, and Policy Cooperation" CEPR Discussion Paper No. 2778.
- Staiger, R., and G. Tabellini (1987). "Discretionary Trade Policy and Excessive Protection," *American Economic Review* 77: 823-37.
- Staiger, R., and G. Tabellini (1999). "Do GATT Rules Help Governments Make Domestic Commitments?" *Economics and Politics* 11: 109-44.
- Stokey, N. (1989). "Reputation and Time Consistency," *American Economic Review, Papers and Proceedings* 79: 134-9.
- Tornell, A. (1991). "On the Effectiveness of Made-to-Measure Protectionist Programs," in E. Helpman and A. Razin (eds.), *International trade and trade policy*, MIT Press, Cambridge, MA.
- Van Damme, E. (1989). "Renegotiation-Proof Equilibria in Repeated Prisoners' Dilemma," *Journal of Economic Theory* 47: 206-17.
- Whalley, J. (1998). "Why Do Countries Seek Regional Trade Agreements?" in J. Frankel (ed.), *The regionalization of the world economy*, Ch. 3, University of Chicago Press, Chicago.