International Theory (2010), 2:1, 32–49 © Cambridge University Press, 2010
doi:10.1017/S175297190999025X

Enforcement and capacity building in international cooperation

JOHANNES URPELAINEN*

Department of Political Science, Columbia University, 420 West 118th Street, 712 IAB, New York, NY 10027, USA

I examine enforcement and capacity building in international cooperation. In a game-theoretic model, a wealthy donor gives foreign aid in exchange for policy implementation by a poor recipient. The recipient has limited capacity to comply with international agreements, so the donor is not sure if cooperation failure is caused by willful disobedience or unintended error. I show that if perceived cooperation failure prompts reciprocal suspension of cooperation, the donor and recipient have a common preference for capacity building. But when the donor can request compensation for perceived cooperation failure, it only chooses to build capacity if cooperation is otherwise impossible. Consequently, the choice of enforcement mechanism shapes capacity building. This result lays a foundation for a genuine synthesis between the enforcement and managerialist schools of compliance. It generates falsifiable hypotheses and explains why reciprocal enforcement, which unfortunately inflicts collateral damage on the victim, is often considered legitimate.

Keywords: international cooperation; compliance; enforcement; capacity building; game theory

According to the enforcement school, a key problem in international cooperation is that states have incentives to violate international agreements (Keohane, 1984; Downs *et al.*, 1996; Downs, 1997, 2000; Gilligan, 2004; Carrubba, 2005). Mutually profitable international cooperation requires the defectors and free riders be punished through sanctions or suspension of cooperation. But not all international cooperation theorists find enforcement so important. For managerialists, the complexity and ambiguity of international cooperation problems dominate enforcement concerns (Mitchell, 1994; Chayes and Chayes, 1995; Young, 1999; Tallberg, 2002). In addition to transparency and rule interpretation, they emphasize capacity building as the key to successful international cooperation.

* E-mail: ju2178@columbia.edu

This study examines the relationship between enforcement and capacity building in international cooperation. In a game-theoretic model, a wealthy donor offers foreign aid to a poor recipient in exchange for policies that eradicate international externalities, such as environmental deterioration. I incorporate enforcement concerns by assuming that policy implementation is costly to recipient. Recipient has limited capacity to comply with international agreements, so compliance failure sometimes occurs despite genuine willingness to comply. A key problem is, therefore, that compliance failure could result either from inadequate capacity or willful disobedience, so the donor does not know if punishment is warranted. For example, Indonesia has consistently failed to reduce illegal logging despite promises made to donors of foreign aid, but it is unclear whether this failure should be mostly attributed to lacking political will or a genuine failure to address the rampant corruption that undermines environmental policies in the country.[1] This problem is so serious that deforestation makes Indonesia the third largest net emitter of greenhouse gas emissions, right after the United States and China.[2]

I compare two canonical enforcement mechanisms. Under *reversion*, mutual suspension of international cooperation follows perceived compliance failure (Axelrod, 1984; Axelrod and Keohane, 1985; Keohane, 1986; Conybeare, 1987; Downs and Rocke, 1995; McGillivray and Smith, 2000; Langlois and Langlois, 2001). If the recipient appears to defect, both sides lose mutually profitable international cooperation for a period of time. Under *compensation*, a defector must compensate the other state to restart international cooperation (Barrett, 1999; Rosendorff and Milner, 2001; Carrubba, 2005; Rosendorff, 2005; Svolik, 2006). If the recipient appears to defect, it must continue to cooperate whereas foreign aid is temporarily frozen. I focus on these stylized enforcement mechanisms because they capture an important feature of any international enforcement mechanism: does the enforcer prefer to implement the punishment or not? They are also variants of reciprocity, which according to Keohane (1984: 214) 'seems to be the most effective strategy for maintaining cooperation among egoists'.

The central theoretical result is that the choice of enforcement mechanism shapes donor incentives to fund capacity building. Under reversion, mutually profitable international cooperation is suspended every time the recipient appears not to comply. Thus, the donor has an incentive to build capacity so as to reduce the rate of spurious compliance failure.

[1] '" Corruption" in Indonesia Logging War'. *BBC News* 14 January 2003.
[2] 'Indonesia World's No. 3 Greenhouse Gas Emitter: Report'. *Reuters* 4 June 2007.

Under compensation, foreign aid is frozen if the recipient appears not to comply. Now the donor has an incentive *not* to build capacity because it can thus hold the recipient accountable for perceived compliance failure more frequently. *Reversion unambiguously aligns preferences for capacity building, while compensation reduces the donor incentive to fund capacity building unless international cooperation is otherwise impossible*. Strikingly, this result holds even if capacity building is *costless*.

The result is important for several reasons. Firstly, it contributes to the research program on international compliance (Chayes and Chayes, 1995; Downs *et al*., 1996; Tallberg, 2002). Although previous scholarship has posited enforcement and capacity building as competing accounts of compliance, this result shows that they interact in complex ways. Indeed, the result could explain why powerful states accept reversionary enforcement mechanisms despite potentially costly consequences. These mechanisms allow a credible commitment to capacity building, so the value of international cooperation to weak states increases, so that powerful states can build a legitimate and mutually profitable international order and avoid the high cost of coercion (Ikenberry, 2000; Stone, 2008). Secondly, it improves our understanding of international institutional design, especially as it depends on the relationship between distributional conflict and enforcement concerns (Abbott and Snidal, 1998; Fearon, 1998; Koremenos *et al*., 2001). If the choice of an enforcement mechanism influences willingness to build capacity, then these decisions are interdependent.

Thirdly, it can help us to understand the politics of legitimacy in international relations (Hurd, 1999; Clark, 2003; Reus-Smit, 2003). At first sight, reversionary enforcement appears normatively troubling because it hurts the innocent victims of perceived wrongdoings. However, it could ultimately align state preferences over capacity building and thus contribute to deeper cooperation in the long run. Finally, the result has notable implications for the study of foreign aid (Alesina and Dollar, 2000; Burnside and Dollar, 2000; Easterly, 2002; Bueno de Mesquita and Smith, 2009). Many scholars have emphasized that giving fungible resources to a government invites corruption and rent seeking. By contrast, capacity building could be particularly conducive to long term development. Thus, reversionary enforcement could reduce the benefits of giving foreign aid by shifting the focus from capacity building to cash transfers.

I first discuss the relationship between enforcement and capacity building in light of the extant literature. I then introduce the model and conduct the analysis. In the concluding section, I discuss the central theoretical and empirical implications of the results.

Enforcement and capacity building

An important debate in international cooperation theory is that between the enforcement and managerial schools of compliance.[3] The enforcement school builds on ideas articulated by Downs *et al.* (1996) who argue that states have incentives to violate international agreements even if they have previously ratified them. In the absence of a centralized enforcement mechanism, such as an omnipotent and benevolent world government, states monitor and sanction behavior by threatening defectors and free riders with punishment (Keohane, 1984; Axelrod and Keohane, 1985; Downs and Rocke, 1995; Carrubba, 2005). To facilitate these tasks, states design international institutions accordingly by establishing dispute resolution mechanisms, monitoring bodies, creating collective enforcement power through issue linkage, and so on (Snidal, 1996; Koremenos *et al.*, 2001).

For managerialists, enforcement is a secondary problem in international cooperation. International cooperation problems are complex and ambiguous, but states usually sign and ratify international agreements under a genuine intention to comply with them. Thus, Chayes and Chayes (1995: 22) write, '[i]f we are correct that the principal source of noncompliance is not willful disobedience but the lack of capability or clarity or priority, then coercive enforcement is as misguided as it is costly'. In addition to rule interpretation and transparency, Chayes and Chayes (1995) emphasize the importance of adequate capacity to implement policies. Tallberg (2002: 614) summarizes this concern as follows:

> Whereas some political and economic capacity problems are beyond the reach of international efforts, deficits in technical knowledge, bureaucratic capability, and financial resources maybe partially or entirely offset through capacity building. Often, capacity building is one of the main programmatic activities of international regimes; in other cases, technical and financial assistance is a targeted measure to alleviate a particular problem.

Capacity building is particularly important in the context of North-South international cooperation. Developing countries have limited capacity to implement complex policies necessary to improve trade policy, prevent environmental deterioration, eradicate or contain contagious diseases, reduce corruption, and so on (Barrett, 1994a, 2007; Keohane and Levy, 1996; Michalopoulos, 1999; Finger and Schuler, 2000; VanDeveer and Dabelko, 2001; Guzman and Simmons, 2005; Shaffer, 2005; Suwa-Eisenmann and Verdier, 2007). And they do not have the

[3] For a review, see Tallberg (2002: 611–614).

financial and organization resources to increase that capacity. Thus, capacity building funded by wealthy donors appears central to successful and mutually profitable international cooperation.

Although, the enforcement and managerialist schools are often seen as competing accounts, it is interesting to note that enforcement concerns and lack of capacity could be mutually *complementary*.[4] Downs (1997: 330) recognizes this arguing that '[t]he claim that the political economy model is suspect because most treaty violations are caused by factors such as the ambiguity of treaties and capacity limitations rather than the product of premeditated exploitation is problematic because the former does not necessarily preclude the latter'. Indeed, enforcement can be complicated by lack of capacity. To see why, recall that if a state with limited capacity fails to comply, others do not know if the reason is an intended defection or an unintended managerial failure. Thus, inadequate capacity provides a smokescreen for opportunism. In the case of illegal logging in Indonesia that I have introduced, failure to prevent it could either stem from lack of political will or a genuine failure to improve the quality of environmental regulation because of bureaucratic ineffectiveness. Should the international community respond by reducing foreign aid as a punishment, or by increasing foreign aid for capacity building?

Although this dilemma is now in common knowledge among international cooperation theorists, the extant literature has little to say about the effect of alternative enforcement mechanisms on a wealthy donor's incentives to contribute to capacity building. Intuitively, it seems plausible that the connection between enforcement and capacity building *encourages* donations because capacity building produces a double dividend. Not only does it solve managerial implementation problems, but it also improves the monitoring (Mitchell, 1994). If the United States helps Ecuador develop implementation capabilities in rainforest conservation, Ecuador can improve forestry practices – and blaming inadequate capacity for implementation failure is simply not credible. Both effects seem desirable to a wealthy donor.

In the analysis, I compare two alternative enforcement mechanisms. Under *reversion*, future cooperation is contingent on past compliance with an international agreement but compensation for damage is not necessary (Axelrod, 1984; Axelrod and Keohane, 1985; Keohane, 1986; Conybeare, 1987; Downs and Rocke, 1995; McGillivray and Smith, 2000; Langlois and Langlois, 2001). For example, if the recipient defects, both foreign aid and policy implementation are temporarily suspended. Under *compensation*,

[4] See Svolik (2006) for an application to democratic transparency in international cooperation.

a defector must compensate the other state (Barrett, 1999; Rosendorff and Milner, 2001; Carrubba, 2005; Rosendorff, 2005; Svolik, 2006). This can be done so that the defector cooperates while the other state free rides for some time. For example, if the recipient defects, policy implementation continues for sometime while foreign aid is temporarily frozen.

Both enforcement mechanisms are based on the idea of reciprocity (Axelrod, 1984). As Downs (2000: 323) writes, '[d]espite its name, the strategy of reciprocity is not designed to uphold some principle of equity-based justice'. Reversionary enforcement is based on reciprocity, but it does not allow compensation to the victims of non-compliance. In contrast, Keohane (1986: 6) argues that reciprocity has a normative dimension because it is based on the principle of 'mutual gain'. The compensation mechanism captures an important element of this normative dimension, because a victim of non-compliance will be compensated.

These enforcement mechanisms are highly stylized. In reality, states can choose from a broad inventory of enforcement mechanisms, ranging from simple reciprocity to issue linkage and institutionalized sanctioning mechanisms or binding dispute resolution (Keohane, 1986; Barrett, 1994b; Lohmann, 1997; Koremenos, 2007). The enforcement mechanisms that I study abstract away from the details and focus attention on the willingness of the enforcer to implement a punishment. Any enforcement mechanism that states could use also influences the payoff to the enforcer, so my main analytical results apply to a wide variety of possible enforcement mechanisms.

Intuitively, compensation seems superior for three reasons. Firstly, it does not inflict unnecessary damage on a victim of non-compliance (Chayes and Chayes, 1995; Rosendorff, 2005). Secondly, compensation is not 'renegotiable' because a victim of non-compliance actually prefers compensation to business as usual (Farrell and Maskin, 1989; Barrett, 1994b). In contrast, both states prefer *not* to implement the punishment under reversion. Finally, I show below that in an exchange relationship, compensation is easier to enforce than reversion. My key contribution is to show that states nevertheless often have good reasons to choose reversion over compensation, because the donor thus has stronger incentives to fund capacity building.

The Model

In the model, a wealthy *donor* and a poor *recipient* form an exchange relationship under which the donor funds policies that eradicate an international negative externality produced by the poor recipient. For example, the European Community and the United States have funded rainforest conservation in Brazil (Kolk, 1998). To improve the value of the exchange relationship, the donor can implement a capacity building

		Recipient	
		C	*D*
Donor	*c*	3 - *x*,*x* - 1	-*x*,*x*
	d	3,-1	0,0

Figure 1 Payoffs from international cooperation. The first (second) payoff is for the donor (recipient).

program. For example, the United States could train Brazilian officials to enhance measures against illegal logging.

The exchange relationship is modeled as an infinitely repeated game with a common discount factor δ. At every time $t \in \{0,\ldots,\infty\}$, the donor chooses between *contribution* (*c*) and *no contribution* (*d*) and the recipient chooses between *cooperation* (*C*) and *defection* (*D*). The payoffs to the two states are given in Figure 1. The idea is that policy implementation is costly to the recipient, but this cost is lower than the benefit to the donor. Thus, policy implementation for foreign aid is mutually profitable.

To model the dilemma of enforcement under limited capacity, I use a technique called imperfect public monitoring (Mailath and Samuelson, 2006). After every period, both states know if the donor contributed or not.[5] But recipient behavior is imperfectly observable, so it could accidentally send a misleading 'signal':

(1) If the recipient cooperates, both states see a 'defection' signal with probability p and a 'cooperation' signal with probability 1−p.
(2) If the recipient defects, both states see a 'defection' signal with probability q and a 'cooperation' signal with probability 1−q.

To ensure that cooperation increases the probability of a favorable 'cooperation' signal, let $p < q$. Intuitively, these signals correspond to the impression that the donor gets about the recipient's *intentions*. The 'cooperation' signal carries information that leads the donor to believe the recipient made a genuine effort whereas the 'defection' signal suggests otherwise. For example, the donor could consult local non-governmental organizations or audit implemented projects (Maxwell and Riddell, 1998; Dai, 2002).

To keep the analysis simple, I consider a single capacity building project that the donor can implement before the repeated game at a small cost $\varepsilon \to 0$. This project improves the accuracy of monitoring, so that the donor learns

[5] All results hold if donor behavior is misperceived with a small enough probability.

https://doi.org/10.1017/S175297190999025X Published online by Cambridge University Press

how the recipient really behaved more often. Formally, probability p decreases while probability q increases with capacity building. The results hold even if capacity building reduces the cost of cooperation to the recipient, but I omit the unnecessary complication.

The purpose of an international agreement is to ensure that the donor contributes and the recipient cooperates. A simplifying assumption is that if the donor ever fails to comply with an international agreement, international cooperation is permanently suspended. Although unrealistic, this convenient assumption is innocuous because donor behavior is transparent. As long as the donor has incentives to comply, international cooperation never breaks down. I assume this condition holds because the discount factor δ is high enough so that there is a large 'shadow of the future' (Axelrod and Keohane, 1985). I can now simply ignore donor incentives and only focus on the difficulties with disciplining the recipient.

I consider two alternative international agreements. The only difference between them is the enforcement mechanism. First, *reversion* is based on the idea that if the recipient appears to defect, both foreign aid and cooperation are suspended for one period. Given that the donor has no incentive to deviate, the equilibrium strategies are as follows:

(1) During 'compliance', play (c, C). During 'punishment', play (d, D).
(2) Begin with 'compliance' at time zero. Play 'punishment' at time $t + 1$ if 'compliance' should have been played but the signal was 'defection' at time t. Play 'compliance' at time $t + 1$ otherwise.

Second, *compensation* is based on the idea that if the recipient appears to defect, it must compensate the donor by cooperating for one period while foreign aid is frozen. Given that the donor has no incentive to deviate, the equilibrium strategies are as follows:

(1) During 'compliance', play (c, C). During 'punishment', play (d, C).
(2) Begin with 'compliance' at time zero. Play 'punishment' at time $t + 1$ if the signal was 'defection' at time t.

Under reversion, 'punishment' play need not be enforced because the states are playing mutual best responses. Under compensation, 'punishment' play must be enforced because the recipient prefers not to cooperate so as to compensate the donor.

Analysis

In equilibrium, the recipient never defects on purpose, but 'defection' signals must nevertheless result in punishment because otherwise the

https://doi.org/10.1017/S175297190999025X Published online by Cambridge University Press

recipient has no incentives to comply at all. Under reversion, it is clear that neither state ever deviates during 'punishment' because they are playing mutual best responses.

If the donor is supposed to contribute and the recipient is supposed to cooperate, the recipient has an incentive to comply if and only if

$$x + \delta \cdot 1 - q \cdot V \leq x - 1 + \delta \cdot 1 - p \cdot V, \tag{1}$$

where V is the average payoff from the game to the recipient. The only effect of replacing a 'cooperation' signal with a 'defection' signal is one unavoidable period of reversion, as if mutually profitable cooperation was postponed by exactly one period. Under imperfect monitoring, we have $0 < V < x - 1$, so the cost of the punishment is less than the value of successful cooperation.

For ease of interpretation, this expression can also be written as

$$1 \leq \delta \cdot (q - p) \cdot V. \tag{2}$$

The left side is the immediate reward for avoiding the cost of cooperation. The right side is the expected cost of reversion in the following period. Note in particular that loss of cooperation is never certain for the recipient, so the loss is weighted by the increase in probability $q - p < 1$. The discount factor δ is present because punishment follows a period later.

Under compensation, the recipient must have an incentive to comply at *any* time, be it during 'compliance' or 'punishment'. It is easy to verify that there is nevertheless only one constraint,

$$x + \delta \cdot 1 - q \cdot x \leq x - 1 + \delta \cdot 1 - p \cdot x, \tag{3}$$

as the punishment consists of exactly one period of compensation. This expression can be written as

$$1 \leq \delta \cdot (q - p) \cdot x. \tag{4}$$

The left side is again the immediate reward for defection and the right side is the cost of possibly losing foreign aid. With $x > x - 1 > V$, international cooperation is unambiguously easier to enforce under compensation than under reversion.

Is compliance possible? If the discount factor δ and the effect of defection on the probability of a 'defection' signal $q - p$ are high enough, there exists a donation x such that international cooperation is mutually profitable and neither state has an incentive to defect on purpose. In this case, international cooperation is *enforceable*. Otherwise it is *not enforceable*. If international cooperation is enforceable, states play according to the international agreement. If international cooperation is not enforceable, the donor never donates and the recipient always defects.

Results

Capacity building increases the value of q and decreases the value of p. Under reversion, this also increases the average payoff to the recipient V, so capacity building unambiguously improves enforceability. I assume that international cooperation is enforceable with capacity building, and possibly but not necessarily without capacity building, as the analysis is trivial if international cooperation is not enforceable even upon capacity building. Capacity building is funded by the donor, so only donor incentives are to be considered.

Consider first reversion. Every period preceded by a 'defection' signal produces a zero payoff to the donor, while the payoff after a 'cooperation' signal is always $3-x>0$. Capacity building increases the probability $1-p$ of a 'cooperation' signal, so it is profitable to the donor:

PROPOSITION 1 (*preference alignment*). Under reversion, the donor builds capacity.

As long as capacity building is not too costly, it is unambiguously profitable to the donor. The only effect of capacity building is to reduce the frequency of erroneous 'defection' signals that occur every now and then even though the recipient cooperates. Such 'defection' signals must trigger punishment that is unfortunately costly to both states under reciprocal enforcement. The silver lining for the recipient is that under reciprocal enforcement, the donor has strong incentives to build capacity.

This simple principle of preference alignment is very general for reversionary mechanisms. It does not depend on the assumption that capacity building increases the average payoff to the recipient, or the assumption that the length of the punishment period is one period. Under reciprocal enforcement, either both states or neither state profit, so the preference alignment is rather strong. Reversion implies that the donor suffers from the recipient's inadequate capacity to implement cooperative policies, so capacity building is something that both sides prefer.

Consider now compensation. Every period preceded by a 'defection' signal produces a payoff 3 while every period preceded by a 'cooperation' signal produces a payoff $3-x$. Clearly, $3>3-x$. Capacity building increases the probability $1-p$ of a 'cooperation' signal, so the frequency of the higher payoff decreases. Thus, capacity building is not profitable if international cooperation is enforceable without it:

PROPOSITION 2 (*preference non-alignment*). Under compensation, the donor builds capacity if and only if international cooperation is not enforceable otherwise.

https://doi.org/10.1017/S175297190999025X Published online by Cambridge University Press

If enforceability is guaranteed, donor–recipient preferences are diametrically opposed under compensation. Every time the recipient cooperates but the erroneous 'defection' signal results, the donor is given unwarranted compensation. As a result, the donor benefits from lack of capacity. But if international cooperation cannot be enforced without capacity building, the donor is willing to build capacity so as to achieve any international cooperation. Thus, the donor never really prefers to invest in capacity building, but it might have to do so for enforceability.

This result is almost diabolical because a seemingly benevolent enforcement mechanism removes the donor incentive to invest in capacity building even if this is *costless*. The two states have designed an international agreement that prescribes compensation, and the lure of this compensation prevents capacity building. It is essential to note that the result is qualified if capacity building increases the value of international cooperation to the donor (as opposed to the recipient). If this increase outweighs the reduced frequency of compensation, partial preference alignment is possible. But even here, full preference alignment is never possible because preferences over the quality of monitoring continue to be diametrically opposed.

The following example suggests that policymakers appreciate this logic. When the Global Environment Facility was first established in preparation for the 1992 Earth Summit, a largely unsuccessful pilot phase led donors and recipients to agree on rather egalitarian principles for the allocation of foreign aid for environmental protection (Streck, 2001; Clémençon, 2006). At the third replenishment of the institution in 2002, President Bush announced that the United States was no longer willing to give foreign aid to developing countries that do not meet stringent performance criteria related to good governance (Hicks *et al*., 2008: 54). This announcement provoked a furious response by the Group of 77 who argued that such conditionality would change the nature of the Global Environment Facility to the extent that it was completely unacceptable (al-Nasser, 2004). Although the United States ultimately prevailed, it is notable that other donor countries also resisted the idea (Hicks *et al*., 2008, 54).

The result on enforcement and capacity building can help us to understand these dynamics. It is not immediately clear why the developing countries should unanimously resist performance criteria, especially if the total amount of aid remains unchanged. It is equally surprising that other donor countries were unwilling to consider such performance criteria, given widespread public concerns regarding the efficacy of foreign aid. But if donors can legitimately reduce foreign aid upon perceived compliance failure, their incentives to support capacity

building in the long term are diminished. Performance criteria force developing countries to correct implementation failures and implement reforms before additional foreign aid is available, so they strongly resemble the compensation mechanism that I have analyzed. The resulting negative effect on capacity building could explain why the other donors agreed with the developing countries that this was not a good idea, even if they were actually committed to global environmental protection.

Concluding Remarks

I have examined the relationship between enforcement and capacity building in international cooperation. The key result concerns the possibility of preference alignment for capacity building. If states cooperate in the shadow of an international agreement based on a reversionary enforcement mechanism, the donor has a strong incentive to build capacity for the recipient because it thus reduces the risk of cooperation failure. But if enforcement is based on compensation for perceived failure to cooperate, the donor could deliberately refuse to build capacity so as to obtain compensation more frequently. Although the game-theoretic model that I have analyzed is simple, this alignment theorem holds in a wide range of circumstances.

International cooperation theorists have not previously recognized this relationship. Other rationalist explanations for limiting enforcement to reversion include the negotiation costs that dispute resolution mechanisms carry (Keohane, 1984; Koremenos *et al.*, 2001; Koremenos, 2005; Carrubba, 2009) and the possibility that enforcing compensation is sometimes too difficult (Downs and Rocke, 1995; Barrett, 1999). The argument advanced here is particularly notable, however, because it operates even if strong sanctions are enforceable and do not carry high negotiation costs. Additionally, it can explain reversion even if compensation is easier to enforce than reversion. But the argument also does not contradict the existing explanations in any way, and indeed the results that favor reversion are *stronger* if the compensation mechanism is costly or difficult to enforce.

These results can inform the debate on enforcement and capacity building as avenues to international compliance (Chayes and Chayes, 1995; Downs *et al.*, 1996; Tallberg, 2002). If the incentive to build capacity depends on the enforcement mechanism, it is not possible to understand one without understanding the other. This finding suggests that empirical and theoretical future research should integrate insights from both schools and serves as a warning against favoring one line of reasoning over the other in policy formation.

The argument is also relevant to the broader relationship between distributional conflict and enforcement. Fearon (1998) argues that while a large 'shadow of the future' facilitates enforcement because states assign higher cost to institutional breakdown, it creates incentives to engage in costly heavy-handed bargaining.[6] My argument shows that the choice of enforcement mechanism has direct implications for distributional conflict. Since the compensation mechanism results in 'regressive foreign aid' from the recipient to the donor, it is even possible that the recipient refuses to engage in costly international negotiations if the compensation mechanism is on the table. Thus, the compensation mechanism could *prevent* international cooperation. This is a stronger effect than an increase in bargaining cost, and it could explain why powerful states often prefer reversion over compensation.

The argument can inform the debate about the role of legitimacy in international cooperation (Young, 1989; Chayes and Chayes, 1995; Hurd, 1999; Wendt, 1999; Clark, 2003; Thompson, 2006; Chapman, 2007).[7] Although legitimacy appears to be an important determinant of compliance with international rules and standards, it is not obvious what the implications for the choice of enforcement mechanisms are. Should we consider reversion illegitimate because it imposes costs on an innocent state? According to my theory, this counterargument to reversion is often invalid because reversion prompts preference alignment *ex ante* and thus appears equitable. The compensation mechanism opens a window of opportunity for exploitation, so weak states and foreign publics concerned with the normative implications of international cooperation have good reasons to view reversionary enforcement as legitimate.

The relationship between enforcement and capacity building is also important for the political economy of development assistance. Recently, many scholars have argued that foreign aid could *harm* economic growth and democratization unless domestic institutions incentivize the recipient government to use resources for the provision of various public goods (Svensson, 2000; Easterly, 2002; Smith and Bueno de Mesquita, 2007; Bueno de Mesquita and Smith, 2009). If the reader accepts the notion that capacity building promotes good governance, it could be that for the population, as opposed to the government, capacity building – not foreign aid *per se* – is the real benefit of international cooperation. Then reversion is particularly desirable because it ensures that even a self-interested donor cooperating with an autocracy can produce valuable public goods.

[6] See Bearce (2002) for institutional breakdown and Krasner (1991) for the power politics of international cooperation.

[7] For legitimacy in social theory, see Suchman (1995).

The most important assumption underpinning the theoretical analysis is that the recipient cannot credibly threaten to terminate cooperation unless capacity building is forthcoming. If the recipient could condition cooperation on capacity building, an equilibrium would exist in which capacity building occurs under compensation enforcement. But given the asymmetric power relationship between the donor and recipient, the range of circumstances in which such bargaining tactics can be successful appears narrow. This boundary condition is also useful because it produces a falsifiable empirical hypothesis that can be operationalized against data on relative bargaining power.[8]

The analysis is robust to such variations as introducing multiple donors and recipients. As long as compliance is imperfectly observable, the fundamental problem of responding to perceived compliance failure remains intact. This problem is also present even if the donor holds 'altruistic' motivations and gives foreign aid to improve welfare in the recipient country. If the altruistic motive is strong enough, however, the commitment problem that I have identified could disappear because the donor holds an intrinsic preference for capacity building.

Theoretical analysis has two counterintuitive empirical implications. First, reversion should be most common in issue areas that require capacity building, such as regulatory and judicial reform or environmental conservation, and if the recipient is a least developed country. While an empirical test is beyond the scope of this study, this result is particularly important in light of the current emphasis on conditionality in the allocation of foreign aid (Vreeland, 2003; Stone, 2008). While some form of conditionality is necessary to incentivize the recipient, capacity building requires that the donor cannot exploit the conditions to demand compensation from the recipient.

Second, reversion should be less common if the two states operate in a highly institutionalized and legalized environment, such as the European Union, where credible commitments are possible.[9] If the donor can commit to capacity building, reversion can be augmented with compensation to the advantage of donor. And the donor could even increase the provision of foreign aid in exchange to ensure that *both* parties benefit from the compensation mechanism. But all this is only possible if mutual commitment capacity is readily available.

To conclude, the analysis creates opportunities for progressive theoretical research in international politics. The simple setup that I have used

[8] For complications of power analysis in international politics, see Baldwin (1979, 1986) and Barnett and Duvall (2005).

[9] See Abbott and Snidal (2000) and Goldstein *et al*. (2000) for legalization in international politics.

offers an accessible and flexible platform for game-theoretic analyzes of international cooperation between weak and powerful states. The mechanics of the baseline model are simple, so it is easy to add such new elements as domestic or transnational actors to it (Keck and Sikkink, 1998; Dai, 2002). The model can be applied to various international cooperation problems and the rationalist premises can be relaxed to incorporate sociological or constructivist theories of international politics (Ruggie, 1998; Hurd, 1999; Wendt, 1999; Checkel, 2001). Ideally, the theoretical analysis contributes to the research program on compliance with international agreements beyond the question that I have addressed.

Acknowledgements

I would like to thank Jennifer Kavanagh, the anonymous reviewers, and the editors of *International Theory* for their valuable comments and advice.

References

Abbott, K.W. and D. Snidal (1998), 'Why states act through formal international organizations', *Journal of Conflict Resolution* **42**(1): 3–32.

—— (2000), 'Hard and soft law in international overnance', *International Organization* **54**(3): 421–456.

al-Nasser, N.A. (2004), *Submission of the Group of 77*, New York, NY: Group of 77.

Alesina, A. and D. Dollar (2000), 'Who gives foreign aid to whom and why?', *Journal of Economic Growth* **5**(1): 33–63.

Axelrod, R. (1984), *The Evolution of Cooperation*, New York, NY: Basic Books.

Axelrod, R. and R.O. Keohane (1985), 'Achieving cooperation under anarchy: strategies and institutions', *World Politics* **38**(1): 226–254.

Baldwin, D.A. (1979), 'Power analysis and world politics: new trends versus old tendencies', *World Politics* **31**(2): 161–194.

—— (1986), *Economic Statecraft*, Princeton, NJ: Princeton University Press.

Barnett, M. and R. Duvall (2005), 'Power in international politics', *International Organization* **59**(1): 39–75.

Barrett, S. (1994a), 'The biodiversity supergame', *Environmental and Resource Economics* **4**(1): 111–122.

—— (1994b), 'Self-enforcing international environmental agreements', *Oxford Economic Papers* **46**(Suppl): 878–894.

—— (1999), 'A theory of full international cooperation', *Journal of Theoretical Politics* **11**(4): 519–541.

—— (2007), *Why Cooperate? The Incentive to Supply Global Public Goods*, Oxford: Oxford University Press.

Bearce, D.H. (2002), 'Institutional breakdown and international cooperation: the European agreement to recognize Croatia and Slovenia', *European Journal of International Relations* **8**(4): 471–497.

Bueno de Mesquita, B. and A. Smith (2009), 'A political economy of aid', *International Organization* **63**(2): 309–340.

Burnside, C. and D. Dollar (2000), 'Aid, policies, and growth', *American Economic Review* **90**(4): 847–868.
Carrubba, C.J. (2005), 'Courts and compliance in international regulatory regimes', *Journal of Politics* **67**(3): 669–689.
—— (2009), 'A model of the endogenous development of judicial institutions in federal and international systems', *Journal of Politics* **71**(1): 55–69.
Chapman, T.L. (2007), 'International security institutions, domestic politics, and institutional legitimacy', *Journal of Conflict Resolution* **51**(1): 134–166.
Chayes, A. and A.H. Chayes (1995), *The New Sovereignty: Compliance with International Regulatory Agreements*, Cambridge, MA: Harvard University Press.
Checkel, J.T. (2001), 'Why comply? Social learning and European identity change', *International Organization* **55**(3): 553–588.
Clark, I. (2003), 'Legitimacy in a global order', *Review of International Studies* **29**(1): 75–95.
Clémençon, R. (2006), 'What future for the Global Environment Facility?', *Journal of Environment and Development* **15**(1): 50–74.
Conybeare, J.A.C. (1987), *Trade Wars: The Theory and Practice of International Commercial Rivalry*, New York, NY: Columbia University Press.
Dai, X. (2002), 'Information systems in treaty regimes', *World Politics* **54**(4): 405–436.
Downs, G.W. (1997), 'Enforcement and the evolution of cooperation', *Michigan Journal of International Law* **19**: 319–344.
—— (2000), 'Constructing effective environmental regimes', *Annual Review of Political Science* **3**: 25–42.
Downs, G.W. and D.M. Rocke (1995), *Optimal Imperfection? Domestic Uncertainty and Institutions in International Relations*, Princeton: Princeton University Press.
Downs, G.W., D.M. Rocke and P.N. Barsoon (1996), 'Is the good news about compliance good news about cooperation?', *International Organization* **50**(3): 379–406.
Easterly, W. (2002), *The Elusive Quest for Growth*, Cambridge, MA: MIT Press.
Farrell, J. and E.S. Maskin (1989), 'Renegotiation in repeated games', *Games and Economic Behavior* **1**(4): 327–360.
Fearon, J.D. (1998), 'Bargaining, enforcement, and international cooperation', *International Organization* **52**(2): 269–305.
Finger, J.M. and P. Schuler (2000), 'Implementation of Uruguay round commitments: the development challenge', *World Economy* **23**(4): 511–525.
Gilligan, M.J. (2004), 'Is there a broader-deeper trade-off in international multilateral agreements', *International Organization* **58**(3): 459–484.
Goldstein, J., M. Kahler, R.O. Keohane and A.-M. Slaughter (2000), 'Introduction: legalization and world politics', *International Organization* **54**(3): 385–399.
Guzman, A.T. and B.A. Simmons (2005), 'Power plays and capacity constraints: the selection of defendants in world trade organization disputes', *Journal of Legal Studies* **34**(2): 557–598.
Hicks, R.L., B.C. Parks, J.T. Roberts and M.J. Tierney (2008), *Greening Aid? Understanding the Environmental Impact of Development Assistance*, New York, NY: Oxford University Press.
Hurd, I. (1999), 'Legitimacy and authority in international politics', *International Organization* **53**(2): 379–408.
Ikenberry, G.J. (2000), *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order after Major Wars*, Princeton, NJ: Princeton University Press.
Keck, M.E. and K. Sikkink (1998), *Activists Beyond Borders: Advocacy Networks in International Politics*, Ithaca, NY: Cornell University Press.

Keohane, R.O. (1984), *After Hegemony: Cooperation and Discord in the World Political Economy*, Princeton, NJ: Princeton University Press.
—— (1986), 'Reciprocity in international relations', *International Organization* **40**(1): 1–27.
Keohane, R.O. and M.A. Levy (1996), *Institutions for the Earth: Sources of Effective International Environmental Protection*, Cambridge, MA: MIT Press.
Kolk, A. (1998), 'From conflict to cooperation: international policies to protect the Brazilian Amazon', *World Development* **26**(8): 1481–1493.
Koremenos, B. (2005), 'Contracting around international uncertainty', *American Political Science Review* **99**(4): 549–565.
—— (2007), 'If only half of international agreements have dispute resolution provisions, which half needs explaining?', *Journal of Legal Studies* **36**(1): 189–212.
Koremenos, B., C. Lipson and D. Snidal (2001), 'The rational design of international institutions', *International Organization* **55**(4): 761–799.
Krasner, S.D. (1991), 'Global communications and national power: life on the Pareto frontier', *World Politics* **43**(3): 336–366.
Langlois, C.C. and J.-P. Langlois (2001), 'Engineering cooperation: a game theoretic analysis of phased international agreements', *American Journal of Political Science* **45**(3): 599–619.
Lohmann, S. (1997), 'Linkage politics', *Journal of Conflict Resolution* **41**(1): 38–67.
Mailath, G.J. and L. Samuelson (2006), *Repeated Games and Reputations: Long-run Relationships*, New York, NY: Oxford University Press.
Maxwell, S. and R. Riddell (1998), 'Conditionality or contract: perspectives on partnership for development', *Journal of International Development* **10**(2): 257–268.
McGillivray, F. and A. Smith (2000), 'Trust and cooperation through agent-specific punishments', *International Organization* **54**(4): 809–824.
Michalopoulos, C. (1999), 'The developing countries in the WTO', *World Economy* **22**(1): 117–143.
Mitchell, R.B. (1994), 'Regime design matters: intentional oil pollution and treaty compliance', *International Organization* **48**(3): 425–458.
Reus-Smit, C. (2003), 'Politics and international obligation', *European Journal of International Relations* **9**(4): 591–625.
Rosendorff, B.P. (2005), 'Stability and rigidity: politics and design of the WTO's dispute settlement procedure', *American Political Science Review* **99**(3): 389–400.
Rosendorff, B.P. and H.V. Milner (2001), 'The optimal design of international trade institutions: uncertainty and escape', *International Organization* **55**(4): 829–857.
Ruggie, J.G. (1998), 'What makes the world hang together? Neo-utilitarianism and the social constructivist challenge', *International Organization* **52**(4): 855–885.
Shaffer, G. (2005), 'Can WTO technical assistance and capacity-building serve developing countries?', *Wisconsin International Law Journal* **23**(4): 643–686.
Smith, A. and B. Bueno de Mesquita (2007), 'Foreign aid and policy concessions', *Journal of Conflict Resolution* **51**(2): 251–284.
Snidal, D. (1996), 'Political economy and international institutions', *International Review of Law and Economics* **16**(1): 121–137.
Stone, R.W. (2008), 'The scope of IMF conditionality', *International Organization* **62**(4): 589–620.
Streck, C. (2001), 'The Global Environment Facility: a role model for international governance?', *Global Environmental Politics* **1**(2): 71–94.
Suchman, M.C. (1995), 'Managing legitimacy: strategic and institutional approaches', *Academy of Management Review* **20**(3): 571–610.

Suwa-Eisenmann, A. and T. Verdier (2007), 'Aid and trade', *Oxford Review of Economic Policy* **23**(3): 481–507.
Svensson, J. (2000), 'Foreign aid and rent-seeking', *Journal of International Economics* **51**(2): 437–461.
Svolik, M. (2006), 'Lies, defection and the pattern of international cooperation', *American Journal of Political Science* **50**(4): 909–925.
Tallberg, J. (2002), 'Paths to compliance: enforcement, management, and the European Union', *International Organization* **56**(3): 609–643.
Thompson, A. (2006), 'Coercion through IOs: the Security Council and the logic of information transmission', *International Organization* **60**(1): 1–34.
VanDeveer, S.D. and G.D. Dabelko (2001), 'It's capacity, stupid: international assistance and national implementation', *Global Environmental Politics* **1**(2): 18–29.
Vreeland, J. (2003), *The IMF and Economic Development*, New York, NY: Cambridge University Press.
Wendt, A. (1999), *Social Theory of International Politics*, New York, NY: Cambridge University Press.
Young, O.R. (1989), 'The politics of international regime formation: managing natural resources and the environment', *International Organization* **43**(3): 349–375.
—— (1999), *Governance in World Affairs*, Ithaca, NY: Cornell University Press.