Unilateral Influence on International Bureaucrats:
An International Delegation Problem*

Johannes Urpelainen†
Assistant Professor
Department of Political Science
Columbia University

October 18, 2009

Abstract

The conventional wisdom emphasizes agency slack or bias as the central problem of international delegation. I show that the possibility of a unilateral influence contest is equally problematic. States can exert unilateral influence on autonomous international bureaucrats, either through rewards or punishments, to pursue their particular interests. A costly contest results, so some states could refuse to delegate because they expect others to be too influential. The analysis has four counter-intuitive empirical implications. First, international agreements often favor institutionally weak states that are disadvantaged in the unilateral influence contest. Second, states could limit the autonomy of an international organization even if this prompts bad policies. Third, a state can sometimes profitably exchange distributional concessions for autonomy. Finally, constraints on unilateral influence are possible only if a disadvantaged state can credibly commit to compensating an advantaged state for it. As illustrations, I discuss the Global Environment Facility and the World Trade Organization.

*I thank Terry Chapman, Leslie Johns, and Christina Schneider for comments.
†Contact information: +1-734-757-0161 (tel); ju2178@columbia.edu (e-mail); 420 W. 118th St, 712 IAB, New York, NY 10027, USA (postal).
1 Introduction

The research program on international delegation, defined here as “a grant of authority by two or more states to an international body to make decisions or take actions” (Bradley and Kelley 2007, 3), is largely based on the principal-agent model. States are principals and international bureaucrats are agents, so the central problem of international delegation is “agency slack” (Hawkins et al. 2006b; Johns 2007; Nielson and Tierney 2003; Pollack 1997). Accordingly, states must design international institutions that give international bureaucrats the proper incentives for effective policy implementation. In effect, states first bargain over the distribution of gains from international delegation, and then choose an international delegation contract that produces an outcome on the Pareto Frontier (Krasner 1991).

In this article, I examine an important but poorly understood strategic problem that the canonical principal-agent model cannot capture. While the theoretical literature does address the distributional conflict over the design of international institutions, it fails to consider the possibility that states informally exert unilateral influence on international bureaucrats so as to bias policy implementation (Hawkins et al. 2006a; Martin 2006; Pollack 1997; Nielson and Tierney 2003). This omission has important theoretical and empirical implications, for the possibility of unilateral influence implies that formal institutional rules could be highly incomplete and potentially misleading proxies of the de facto rules of the game (Stone 2004, 2008). The possibility of unilateral influence could reduce the value and change the optimal design of international delegation because states must prepare for a unilateral influence contest. If states can exert unilateral influence by rewarding or punishing international bureaucrats for biased policy implementation, when and how can states capitalize on effective policy implementation through international delegation?

The question of unilateral influence is not a theoretical curiosity. In the International Monetary Fund (IMF), the implementation of policies and enforcement of conditionality are selective. Allies of major powers can violate loan contracts without punitive consequences, despite formal rules that prescribe strict conditionality. As Stone (2004, 590) shows in his study of IMF conditionality in Africa, conditionality requires credible threats to enforce it, and “[t]he obstacle to enforcing these threats is interference by the major donor countries.” In the World Trade Organization (WTO), limited legal capacity prevents developing countries from using the Dispute Settlement Understanding (DSU) to hold major powers accountable, although the very idea of legalized dispute resolution was to go beyond power politics.
(Busch and Reinhardt 2003). In the Global Environment Facility (GEF), wealthy donors and the World Bank are in a dominant position despite seemingly egalitarian formal voting rules because developing countries lack the resources and information to influence practical policy implementation (Streck 2001, 92).

In my game-theoretic model, two states can draft an incomplete contract to delegate policy implementation to an international organization. The contract describes the sought distribution of gains and how autonomous policy implementation by an international bureaucrat will be. Upon delegation, they engage in a zero-sum unilateral influence contest. In equilibrium, the quality of policy implementation increases with autonomy because the international bureaucrat can expend effort at a lower cost. However, returns to unilateral influence increase because the international bureaucrat cannot be easily disciplined, so the costly unilateral influence contest intensifies. The dilemma that states face can therefore be characterized as follows. By giving autonomy to an international bureaucrat, states ensure effective and flexible policy implementation that capitalizes on informational advantages. But autonomy increases the level of unilateral influence, and this is costly to both states.

This dilemma has a number of notable empirical implications. First, the analysis shows that states sometimes deliberately formulate the delegation contract against the interests of influential states. The delegation contract forms a reference point for unilateral influence, so formal provisions that favor a disadvantaged state function as a commitment device that permits broader participation. Contravening the conventional wisdom that emphasizes the importance of aligning formal treaty provisions with the preferences of major powers (Hawkins et al. 2006a; Koremenos, Lipson, and Snidal 2001), my theoretical results reveal the possibility that written rules favor the underdog while the actual outcome favors the top dog.

Second, the analysis shows that states must sometimes limit the autonomy of international organizations even if they ascribe a high value to competent policy implementation. If the returns to unilateral influence increase rapidly with autonomy, high levels of autonomy induce a disastrously costly unilateral influence contest. Consequently, states limit the autonomy of the international organization even though this results in bad policies. Notably, this finding could explain why many international organizations, such as the United Nations, are often incapable of implementing badly needed policies even though major powers seem to value such outcomes as collective security.

Third, international delegation often enables a mutually profitable exchange of distributional concessions for autonomy. If one state cares about
the distribution of gains but does not value competent policy implementation, while another state is willing to concede as long as implementation is highly effective, such an exchange is possible. An example is a developing country that obtains increased foreign aid for environmental conservation and allows rigorous enforcement and intrusive monitoring by an international organization in exchange (Keohane and Levy 1996).

Finally, if states can condition international cooperation on limited unilateral influence, they achieve higher payoffs by avoiding the unilateral influence contest. However, this is only possible if the more influential state can be guaranteed a favorable distribution of gains in exchange for limiting its influence. Such a commitment seems to have been possible in the aftermath of the Second World War, as the United States reduced its unilateral influence in exchange for commitment to a liberal economic order by its allies (Ikenberry 2000).

I proceed as follows. In Section 2, I review and criticize the literature. In Section 3, I construct the model. In Section 4, I analyze the equilibrium. In Sections 5 through 7, I present the results. In Section 8, I conduct two illustrative case studies. In Section 9, I offer concluding remarks.

2 The Puzzle

In a principal-agent relationship, information is asymmetric between a principal and an agent, such as a politician and a civil servant (Bendor and Meirowitz 2004; Epstein and O’Halloran 1999; Gilardi 2002; McCubbins 1985; McCubbins and Schwartz 1984; Weingast and Marshall 1988). The principal needs the agent’s expertise for competent policy implementation, but the agent could “shirk” by expending limited effort or “drift” by implementing policies that are not favorable to the principal. Consequently, the principal must design a feasible incentive structure, usually referred to as a “contract.” In the international context, the conventional wisdom is that the act of delegation is particularly problematic because states prefer not to incur “sovereignty costs” in the form of lost autonomy (Abbott and Snidal 1998, 2000; Barnett and Finnemore 2004; Hawkins et al. 2006b; Moravcsik 2000).

The extant literature emphasizes the implications of preference heterogeneity (Hawkins et al. 2006a; Martin 2006; Nielson and Tierney 2003). If the delegating states are in substantial disagreement, it is difficult to design a contract that induces mutually acceptable policy implementation by the international organization (Martin 2006). Consequently, preference
heterogeneity should reduce the probability of successful delegation. This relationship is complicated by the presence of a “collective principal,” which exists if “an agent has a single contract with a principal, but the principal happens to be composed of more than one actor” (Nielson and Tierney 2003, 247). If decisions are made through voting, even the most powerful states must form coalitions to implement new policies. Additionally, warns Martin (2006, 144), “[w]hen state preferences diverge ... there is more likely to be a wide range of proposals that could gain majority approval. This gives the staff room to maneuver.”

Although the principal-agent model is a significant advance over exclusively state-centric theories, the literature is confused as to the nature of principal-agent problems in international politics. States are not a board of directors that first resolve their disputes, then write a legally binding contract to delegate, and finally discipline the bureaucrat. In addition to pursuing their interests by bargaining over the delegation contract, states can exert unilateral influence on policy implementation by the bureaucrat. Consequently, the international bureaucrat could fail to implement outcomes specified in the contract also because some states reward or coerce international bureaucrats to bias policy implementation. The international delegation contract is “incomplete” in that states can undermine it to pursue their interests at the expense of other states (Koremenos 2001; Tirole 1999). For example, empirical scholarship has documented how industrialized countries, and the United States in particular, frequently use power in the IMF, the World Bank, and the WTO far beyond their formal voting rights (Steinberg 2002; Stone 2004, 2008; Woods 2006).

Economic models of “common agency” capture some of these effects (Bernheim and Whinston 1986; Dixit, Grossman, and Helpman 1997; Grossman and Helpman 1994). In these models, multiple principals contribute money to an agent in exchange for favorable policies. However, existing models of common agency ignore the possibility that a principal refuses to delegate because it anticipates unilateral influence by other other principals in the implementation phase. This omission turns out to have important theoretical and empirical implications for international relations.

How do states actually exert unilateral influence in international delegation under distributional conflict? The permeability of international organizations implies that states can often exert unilateral influence on international bureaucrats (Hawkins and Jacoby 2008). The instruments of influence available to a state include career rewards and punishments. To begin with, Voeten (2008, 428) finds that judges in the European Court of Human Rights “are subject to increased pressure on controversial cases that directly deal
with the security of a country” and ascribes this finding to career incentives. Similarly, states can allocate tangible resources by conditioning funding on biased policy implementation, and they can collaborate directly with sympathetic agents (Addison, McGillivray, and Odedokun 2004; Pollack 1997; Selin 2007; Streek 2001). For instance, green Member States of the European Union can collude with the Directorate General Environment to ensure that their viewpoints are represented in policy implementation (Selin 2007). In the case of the UNEP or the GEF, wealthy donors have occasionally cut funding for these bodies out of fear that they either promote the interests of developing countries or push for overly ambitious environmental goals and policies (Ivanova 2007; Streek 2001).

States can also recruit staff that share their preferences or put ideological pressure on the international bureaucrats (Barnett and Finnemore 2004; Johns 2007; McKewown 2009). For example, in the context of the weapons inspections between the two Iraqi wars, Johns (2007, 267) argues that Russia, France, and China came to perceive the United Nations Special Commission created by the Security Council as “too aggressive and biased in favor of U.S. policy” and sought to exert influence “by forcing new personnel to the agency.” Although the mandate of the Commission was previously established in international negotiations, major powers continued to bias policy implementation to advance their interests.

This effect can also be indirect if the qualifications required for a career in an international organization induce a selection effect that favors certain countries. Perhaps the most important example are the Bretton Woods institutions that hire a large number of economists. As Woods (2006, 53) shows, officials trained in American or British universities have historically had a dominant position in senior management. Such overrepresentation creates a fertile ground for the diffusion of Anglo-American ideas about economic policy, such as the neoliberal Washington Consensus (Williamson 1990).

Often states can directly influence decisions made by international bureaucrats. Consider the following description of the IMF Executive Board by Stone (2008, 595):

“informal participation allows influential shareholders to control the substance of the management proposal, assuming the formal proposal-setting prerogatives of the chair for themselves. This allows the United States to exert effective control by participating much more actively than the other shareholders. The United States has a tremendous organizational advantage over
other countries because it has a more extensive diplomatic corps, particularly important private financial institutions, numerous advantages in gathering information, and all of the advantages of having the IMF located in the U.S. capital, in addition to issuing the international reserve currency and commanding the resources of a superpower."

If this statement is accurate, the United States has a major advantage in the unilateral influence contest.

Historical experience suggests that even flagrant forms of principal-agent collusion cannot be excluded. As Mathiason (2007, 28) writes, in the League of Nations, many supposedly neutral international bureaucrats were directly taking orders from their national government and even providing confidential information to the national capital. Especially problematic were the autocratic German and Italian regimes.

Less obviously, formal voting rules are also a potential vehicle of de facto unilateral influence. While the stated purpose of formal voting rules is to define the rules of the game, in practice they leave room for unilateral influence. If states form a collective principal, they can try to create “winning coalitions” by offering rewards to others for supporting certain proposals, or by imposing sanctions against states who fail to do so. For example, Kuziemko and Werker (2006) show that permanent members of the United Nations Security Council use foreign aid to secure favorable votes by the non-permanent members. Similarly, Miller and Dolsak (2007) document how Japan increases its bilateral foreign aid in exchange for votes in the International Whaling Commission. While these rewards are not directly accrued by the international bureaucrat, they indirectly bias policy implementation in favor of an entrepreneurial state willing to bribe others, against the spirit of the original contract that sought to aggregate preferences through sincere voting.

A related issue are the procedures to change the voting rules. Woods (2000, 830) describes the efforts by Japan to increase its voting quota in the Bretton Woods institutions during a period of phenomenal economic growth after the Second World War. While the contract to establish the Bretton Woods institutions prescribed that states be assigned votes according to their contributions, the United States and the major European donors fought against Japan’s attempts to increase its quota. Formally, the Bretton Woods institutions should have given Japan a larger vote share, but other major powers were able to slow down this undesirable change that would have shifted the policies implemented by the IMF and the World
Bank toward Japan’s ideal point. Notably, this success came at the cost of a bitter political struggle that hurt the relations between Japan and other industrialized countries.

Importantly, the covariates of unilateral influence could differ from the covariates of raw bargaining power. While powerful states usually exert substantial influence also within international organizations, seemingly weak states can sometimes thrive in institutionalized settings. For example, many United Nations bodies, such as the Development Program (UNDP), have a built-in bias towards developing countries simply because the staff have intrinsic developmental motivations. Similarly, an essential precondition for using the WTO dispute resolution mechanism is legal capacity, and many small industrialized countries in Europe have it in abundant supply (Busch and Reinhardt 2003). Throughout, I let “weak” and “powerful” refer to different levels of bargaining power under anarchy, while “advantaged” and “disadvantaged” refer to different opportunities for influence under the auspices of an international organization.

The possibility of unilateral influence reveals a tradeoff in how autonomous international organizations should be. If states are to capitalize on the expertise that international bureaucrats possess, they must give them substantial autonomy in policy implementation (Hawkins et al. 2006; Nielson and Tierney 2003). An agent that cannot do anything is simply not useful. While such autonomy could cause agency slack, I focus here on the problem of new opportunities for unilateral influence. If international bureaucrats are relatively autonomous, so that their behavior is not directly constrained by such rigid governance structures as unanimity voting on minute details of various policies, states can expect great benefits if they successfully collude with them. If the United States can dictate the economic doctrine used by the Bretton Woods institutions, economic statecraft is easier (Barnett and Finnemore 2004). If the European Union dominates international regulatory institutions, European hegemony in regulatory issues is strengthened (Mattli and Büthe 2003). If the Soviet Union and its satellites can hijack the United Nations Educational, Scientific, and Cultural Organization, they score a major symbolic victory in the Cold War of ideas (Osakwe 1972).

Notably, this conceptualization of autonomy differs from the notion of “independence” in the legalization literature (Alter 2008; Goldstein et al. 2000; Haftel and Thompson 2006). Autonomy aggravates the problem of unilateral influence because disadvantaged states cannot simply veto or otherwise prevent the implementation of biased policies. Independence could either mitigate the problem because international bureaucrats are not subject to threats by states, or amplify it because they can sell policy imple-
mentation for material or career rewards.

3 The Model

I model international delegation as follows. First, two states draft an incomplete contract that prescribes delegation of policy implementation to an international organization. Second, if they choose to delegate, each state can unilaterally influence an international bureaucrat so as to bias policy implementation toward its ideal point. Third, the international bureaucrat strategically implements a policy to achieve her preferred outcome.

First, states $i = A, B$ simultaneously announce a contract $(\overline{y}, \theta)$. If they announce different contracts, the game ends and both obtain a payoff $\lambda_i > 0$ that can be interpreted as the value of a unilateral outside option (Johns 2007; Voeten 2001). Throughout, I let $\lambda_i$ be low enough for both states $i$ so that international delegation is potentially profitable in the absence of the unilateral influence contest.

The first element $\overline{y} \in \mathbb{R}$ is a baseline distribution that the international bureaucrat is officially supposed to achieve through policy implementation. As $\overline{y}$ increases, state $A$ is supposed to obtain a larger share of the benefits at the expense of state $B$. Substantively, the baseline distribution comprises provisions that describe the distribution of costs and benefits, such as vote shares or general principles and norms of policy implementation.

The second element $\theta \in [0, 1]$ measures the autonomy of the international organization. As autonomy increases, the international bureaucrat becomes more sensitive to unilateral influence by states. However, its ability to competently implement policy is also enhanced. For example, suppose the European Commission is given a lot of autonomy in policy implementation. It can capitalize on expertise to implement optimal policies, but if France is in good terms with it, other states find it difficult to prevent bargains between the Commission and France because they cannot easily interfere with the operations of the Commission.

At the second stage of the game, both states $i$ simultaneously choose how much influence $z_i \in [0, \infty)$ to exert on the international bureaucrat at cost $\alpha_i \cdot c(z_i)$, where $c$ is increasing and strictly convex in $z_i$. Let $c(0) = c'(0) = 0$ so that neither state $i$ exerts zero influence unless influence has no effect whatsoever on the international bureaucrat. The parameter $\alpha_i > 0$ captures how costly it is for state $i$ to influence the international bureaucrat. As $\alpha_i$ increases, state $i$ becomes disadvantaged. Influence shapes the behavior of the international bureaucrat as detailed below.
At the third stage of the game, the international bureaucrat simultaneously selects a policy \( x \in \mathbb{R} \) and chooses an effort \( e \geq 0 \) to successfully implement that policy. The international bureaucrat obtains an informative implicit signal that is unobservable to states. This signal tells her how different policies \( x \) map into distributions \( y \). As a strategic actor, the international bureaucrat selects a policy \( x^* \) that implements her ideal point

\[
y^* = \mathcal{Y} + \theta \cdot (z_A - z_B)
\]

as the final distribution. The first term captures the idea that the baseline distribution \( \mathcal{Y} \) influences the ideal point, as states select mutually acceptable international bureaucrats and give them proper incentives to implement the baseline distribution. The second term captures the idea that unilateral influence by state \( A \) increases the ideal point of the international bureaucrat while unilateral influence by state \( B \) decreases it. This is logical because high distributions \( y \) will be favorable to state \( A \) at the expense of state \( B \). For parsimony, this formulation omits the possibility that the international bureaucrat is intrinsically biased (Johns 2007).

The international bureaucrat prefers to implement her final distribution as competently as possible, but effort is costly. Suppose the international bureaucrat chooses effort \( e \) to maximize a single-peaked policy payoff \( g(e, \theta) \). Assume the maximizer \( e^* = e^*(\theta) \) for \( g \) is strictly increasing in \( \theta \) to capture the idea that high levels of autonomy encourage effort. This premise captures the notion that autonomy is necessary for competent policy implementation, as the international bureaucrat can focus on the problem at hand without continuous interference by states. For example, if the IMF staff must have the approval of all major donors for any minutiae change in a conditional loan agreement, the cost of developing new innovative policies is prohibitive.

The final distribution \( y^* \) yields a policy implementation payoff

\[
v_i(y^*) \cdot u_i(e)
\]

to state \( i \). Let \( v_A \) be increasing and strictly concave with \( v_A'(-\infty) \to \infty \) and \( v_A(-\infty) = 0 \). Let \( v_B \) be decreasing and strictly convex with \( v_B'(\infty) \to \infty \) and \( v_B(\infty) = 0 \). These convenience assumptions ensure that a unique subgame equilibrium exists in the unilateral influence contest. Let \( u_i(e) \) be positive and strictly increasing. The idea is that the value of competent policy implementation \( u_i(e) \) is multiplied by the share of benefits that goes to state \( i \). Both increases in the competence and share exhibit decreasing returns to scale.
Combined with the cost of unilateral influence, the payoff from successful delegation for state $i$ is given by

$$v_i(y^*) \cdot u_i(e) - \alpha_i \cdot c(z_i^*).$$

(3)

Intuitively, the cost of the unilateral influence contest is subtracted from the payoff to policy implementation by the international organization.

## 4 Equilibrium

I use the subgame-perfect equilibrium. This is a contracting game, so I only examine equilibria on the Pareto-frontier. The central question is if mutually profitable international delegation is possible. To answer this question, I use backward induction.

The international bureaucrat implements a final distribution according to (1) and chooses effort $e$ optimally, so the question if there is a delegation contract $(\tilde{y}, \theta)$ that induces a payoff (3) higher than the value of the outside option $\lambda_i$ for both states $i$. Since equilibrium effort $e^*$ is strictly increasing in autonomy $\theta$, the payoff from policy implementation increases with it for both states $i$.

**Claim 1.** For any given final distribution $\tilde{y}$, the equilibrium payoff from policy implementation to state $i$ given by $v_i(\tilde{y}_i) \cdot u_i(e)$ is strictly increasing in autonomy $\theta$.

The cost of the unilateral influence contest notwithstanding, autonomy is unambiguously beneficial because it allows the international bureaucrat to expend effort towards successful policy implementation.

Consider next the stage at which each state $i$ chooses how much influence $z_i$ to exert. Autonomy $\theta$ and the baseline distribution $\bar{y}$ are given, but the final distribution $y^*$ is determined in the unilateral influence contest. The following first-order condition must hold for both states $i$:

$$\theta \cdot \frac{\partial v_i(y^*)}{\partial y} = \alpha_i \cdot \frac{\partial c(z_i^*)}{\partial z_i}.$$  

(4)

The benefits of additional influence must equal the costs, and autonomy $\theta$ increases the benefits of additional influence. **Figure 1** illustrates the best responses. As shown in the Appendix, the first-order condition holds with equality and a unique interior equilibrium exists.
The analysis focuses on unilateral influence by states, so it is useful to ex-
amine the choice of influence $z_i^*$ and the resulting cost $\alpha_i \cdot c(z_i^*)$ in some
detail. If $\alpha_i$ is low so that the marginal cost of influence is low and increases
only slowly for state $i$, it can easily influence the international bureaucrat,
so it will have an upper hand against the other state $j$. If $\alpha_i$ and $\alpha_j$ are
low so that the marginal cost is relatively high but increases only slowly for
both states, however, the unilateral influence contest will be intense. Both
states must choose high levels of influence just to prevent the other state
from gaining ground.

Autonomy $\theta$ can be harmful in the unilateral influence contest.

Claim 2. The equilibrium cost of unilateral influence $\alpha_i \cdot c^*$ is increas-
ing in autonomy $\theta$.

An autonomous agent can be easily influenced. As a result, autonomy in-
creases the returns to unilateral influence. By creating a constrained agent
instead, two states can avoid the zero-sum game in which they exert influ-
ence after they have agreed on delegation. If a state tries to influence a
constrained agent, the other state can directly intervene to undermine bi-
ased policy implementation. For instance, it could use a veto or withdraw
funding.

Consider finally the decision to delegate. For a final distribution $y^*$ and
agent autonomy $\theta$, state $i$ is willing to delegate if and only if
\begin{equation}
 v_i(y^*) \cdot u_i(e^*(\theta)) \geq \alpha_i \cdot c^*(z_i^*) + \lambda_i. \tag{5}
\end{equation}
Unless mutually profitable policy implementation is possible, one state $i$
obtains a strictly higher payoff $\lambda_i$ from the ex ante outside option, so dele-
gation is not feasible. Importantly, extremely asymmetric influence contests
do not permit international delegation.

Claim 3. If $y^*$ is high or low enough for some contract $(\pi, \theta)$, at least
one state $i$ cannot propose that contract in equilibrium.

If almost all benefits from international delegation go to state $j$ in equi-
librium, state $i$ cannot possibly benefit from international delegation. How-
ever, it could lose because the unilateral influence contest is costly. This
claim captures the commitment problem that unilateral influence prompts.
If state $j$ cannot guarantee some benefits to state $i$, international delegation is not feasible.

In conclusion, international delegation is possible if and only if there exists a contract $(\overline{y}, \theta)$ that does not yield an extreme final policy $y^*$. If such contract does not exist, each state $i$ proposes a contract that is unacceptable to the other state $j$. International delegation fails and unilateral action follows.

5 Contracts and Outcomes

To understand the unilateral influence problem, it is useful to investigate how the design of the international delegation contract $(\overline{y}, \theta)$ influences equilibrium behavior and payoffs. I first evaluate the effect of changes in the baseline distribution $\overline{y}$ and then in autonomy $\theta$.

The first choice that the states face is the baseline distribution $\overline{y}$. The possibility of unilateral influence precludes credible ex ante commitment to a final distribution $y^*$, but the baseline distribution turns out to have an unambiguous effect on the final distribution.

**Proposition 1.** The final distribution $y^*$ is increasing in the baseline distribution $\overline{y}$. The influence $z_A^*$ of the gaining state $A$ decreases while the influence $z_B^*$ of the losing state $B$ increases in the baseline distribution $\overline{y}$.

The first effect of increasing the baseline distribution $\overline{y}$ is that the reference point for unilateral influence shifts to the advantage of state $A$. Although changes in incentives to exert unilateral influence will partially offset this effect, the final distribution $y^*$ must be at least as good for state $A$. The second effect is that state $A$ reduces its influence while state $B$ increases its influence. The baseline distribution $\overline{y}$ is already more favorable to state $A$, so the value of influence decreases. State $B$ faces an increasingly costly outcome, so it must exert additional influence.

Both effects favor state $A$ at the expense of state $B$. The choice of baseline distribution $\overline{y}$ is central to the politics of international delegation because it has implications both for the distribution of gains from delegation and the cost of unilateral influence. As long as international delegation is feasible, the two states hold diametrically opposed preferences over the baseline distribution $\overline{y}$. Unfortunately, the baseline distribution $\overline{y}$ is an imperfect instrument. If unilateral influence cannot be prevented, both states exert some to obtain a better distribution. As a result, the final distribution
\(y^*\) drifts and both states incur a cost \(\alpha_i \cdot c(z^*_i)\). It would be better for both states if they could directly contract on the final distribution \(y^*\) and avoid the unilateral influence.

Notably, even the direction of equilibrium influence could change with the baseline distribution \(\overline{y}\). To see why, suppose first the baseline distribution \(\overline{y}\) is high so that state \(B\) has strong incentives to exert unilateral influence. State \(A\) obtains a favorable distribution in any case, so the most probable outcome is that the final distribution \(y^*\) will be better for state \(B\) after intense attempts to influence the international bureaucrat: \(y^* < \overline{y}\). But if the baseline distribution \(\overline{y}\) is low so that state \(A\) has strong incentives to exert influence, the final distribution \(y^*\) drifts towards state \(A\)'s ideal point: \(y^* > \overline{y}\). For example, if a superpower such as the United States accepts egalitarian decision rules for an international organization, it has great incentives to compel other countries to vote according to its preferences. But if the institutions rules reflect the preferences of the United States, the returns to aggressive use of power are lower.

The second choice that states face is autonomy \(\theta\). Both states gain from improved agent performance, but the effect of autonomy on the final distribution \(y^*\) or the cost of influence \(\alpha_i \cdot c(z^*_i)\) could induce distributional conflict.

**Proposition 2.** The cost of influence \(\alpha_i \cdot c(z^*_i)\) is increasing in autonomy \(\theta\). The final distribution \(y^*\) is increasing in autonomy \(\theta\) if and only if \(\alpha_B\) is low enough while \(\alpha_B\) is high enough.

As autonomy increases, the returns to successful unilateral influence increase. This has two effects on the level of unilateral influence that states exert. On the one hand, for any given influence exerted by the other state, a state has a greater incentive to exert influence because of the high benefits. On the other hand, competition will be more intense, as each state increases its influence as a response to an increase by the other state. How this affects the final distribution \(y^*\) depends on which state is better positioned to take advantage of the increased autonomy by exerting additional influence. The relative advantage of state \(i\) depends first and foremost on the cost parameter \(\alpha_i\). For example, if the United States is in a good position to influence decisions made by the IMF Executive Board, IMF autonomy could shift policy implementation toward the ideal point of the United States (Stone 2008).

The two states can thus reduce autonomy \(\theta\) in the delegation contract to change equilibrium behavior in two ways. First, by reducing autonomy, both states benefit from a less intense unilateral influence contest. Second,
a reduced autonomy mitigates the advantage that the more influential state holds. Unfortunately, these effects come at the expense of competent policy implementation, as the international bureaucrat expends less effort under low levels of autonomy.

6 Designing International Delegation

The propositions above permit an analysis of the design of international delegation contracts. I first find the conditions under which both states agree on the level of autonomy $\theta$ and then examine the choice of the baseline distribution $\mathcal{Y}$.

Consider first the choice of autonomy.

**Proposition 3.** If effort $e^*(\theta)$ increases rapidly (slowly) enough with autonomy $\theta$ while $\alpha_i, \alpha_j$ are high (low) enough, both states $i, j$ prefer maximal (minimal) autonomy $\theta = 1 (\theta = 0)$.

Both states prefer autonomy if it produces a high delegation payoff by inducing competent implementation without resulting in a costly unilateral influence contest. In contrast, if autonomy does not produce a high delegation payoff, but it causes an intense unilateral influence contest, both states prefer to give the international bureaucrat limited autonomy.

As an example of the virtues of high autonomy, competent policy implementation could be particularly valuable in international environmental affairs because the consequences of pollution and resource overconsumption are highly complex (Biermann 2001; Young 2002). Additionally, many non-governmental environmental organizations constantly monitor implementation, so an international organization cannot shirk or implement biased policies in any case (Dai 2005; Raustiala 1997). Consequently, states have incentives to allow high levels of autonomy.

The result could also shed light on those cases in which states choose limited autonomy although competent policy implementation is quite valuable. Even if international organizations hold valuable expertise, the high cost of unilateral influence could dissipate too much of the surplus from international delegation. Consequently, states choose to give international bureaucrats limited freedom to implement policies, not because states do not value policy implementation, but because they understand that the international bureaucrats are sensitive to unilateral influence. For example, although the Bretton Woods institutions deal with complex issues requiring
substantial expertise, the voting rules allow major industrialized countries to easily block policies (Woods 2006).

Equally notable is the possibility that states disagree on autonomy. According to the conventional principal-agent model, autonomy is something that usually aligns state interests (Alter 2008; Hawkins et al. 2006a; Nielson and Tierney 2003). For instance, in his seminal application of the principal-agent model to the European Union, Pollack (1997, 129) claims that “supranational agents may exploit differing preferences among the member states to avoid the imposition of sanctions against shirking and to ’push through’ legislative proposals via their formal agenda-setting powers.” If this is true, the member states have a strong collective incentive to agree on constraining the supranational agents ex ante, or before any particular contentious issue emerges, because they can thus at the very least avoid agency slippage.

The results presented here show that this premise could fail under rather undemanding conditions. If a state is very sensitive to poor performance and has a great advantage in the unilateral influence contest, it can but lose from reduced autonomy. If another state does not benefit from autonomy and is at a disadvantage in the unilateral influence contest, it unambiguously prefers less autonomy. Distributional conflict emerges on whether the principal should discipline the agent in the first place, so the European principal-agent problem has dimensions that Pollack (1997) fails to acknowledge. Indeed, an important theme in the history of European integration are pervasive disagreements between major member states, such as Germany and Great Britain, on how centralized the union should be (Konstantinidis 2008; Slapin 2009).

For another example, consider a wealthy democracy and a poor autocracy that agree on improving human rights in exchange for foreign aid (Hafner-Burton 2005). The poor autocracy is only willing to do this to receive foreign aid, so it does not really care about poor performance. If the international organization fails to monitor human rights, it might even benefit. However, the wealthy democracy is interested in avoiding bad performance because the only reason it is giving foreign aid is to improve the human rights conditions in the poor autocracy. This distributional conflict is further aggravated if the wealthy democracy is unable to exert unilateral influence because domestic institutional constraints shrink the room for maneuver, while the poor autocracy is willing to bribe officials in the international organizations.

What about the baseline distribution $\overline{y}$?

**Proposition 4.** Fix autonomy $\theta$. If $\alpha_i$ is high enough while $\alpha_j$ is low
enough, the baseline distribution $\overline{y}$ must be high (low) enough for $i = A$ ($i = B$).

Reduced autonomy mitigates the unilateral influence contest. However, for any level of autonomy $\theta$, it is also true that if state $i$ is disadvantaged enough in the unilateral influence contest, the baseline distribution $\overline{y}$ must reflect its preferences. A favorable baseline distribution reduces the extent of disadvantage, so such a baseline distribution $\overline{y}$ permits a credible commitment to a mutually acceptable final distribution $y^*$.

This effect should be most pronounced if autonomy per se is valuable because it enables competent policy implementation. Neither state is ready to sacrifice the quality of policy implementation, so they must find a way to credibly commit to giving some gains to the disadvantaged state. They can do this by drafting a contract that specifies large gains for the disadvantaged state, understanding that this status quo will not hold after the unilateral influence contest. For example, I argue below that the WTO DSU combines relatively high levels of autonomy with seemingly egalitarian procedures that wealthy industrialized countries then informally exploit by using their superior legal capacity (Busch and Reinhardt 2003; Guzman and Simmons 2005).

These findings are particularly important for two reasons. First, they contradict the conventional wisdom that delegation contracts should reflect the interests of powerful and influential states. Hawkins et al. (2006a, 22) argue that “[w]hen institutional rules fail to reflect accurately the distribution of power, powerful states will more readily choose to act alone outside the institution” while Koremenos, Lipson, and Snidal (2001, 792) write that powerful states control international institutions through international agreements. My findings corroborate the underlying logic of power, but they also indicate that the conventional empirical predictions are only correct if powerful states cannot operate within the institution by violating the rules. Instead, delegation outcomes should reflect their interests. If a state has a sufficiently large disadvantage in the unilateral influence contest, no matter how weak it is outside the international organization, the delegation contract will prompt an attractive baseline policy $\overline{y}$ that is then informally overturned. Conventional inferences on the advantage that major powers have in international organizations are therefore biased downward if one focuses on formal treaty provisions only.

This is essentially what Stone (2008) finds in his comprehensive empirical analysis of IMF conditionality. As discussed above, his results support the notion of “informal governance” by the United States. Where essential
American interests are threatened by stringent conditionalities, the latter are simply not being enforced, but to avoid dissipating the legitimacy of the Bretton Woods institutions, the United States exercises strategic restraint. In my model, informal governance is the special case in which one of the principals enjoys an overwhelmingly dominant position within and outside the international organization in focus.

Second, the findings provide a powerful rationalist argument for egalitarian institutional rules in international politics. Many scholars have highlighted the need to legitimize international institutions and organizations (Barnett and Finnemore 1999; Hall 1997; Hurd 1999; Najam 2005; Reus-Smit 2003; Ruggie 1982), but few have systematically examined the strategic implications of various principles and norms. My argument suggests that egalitarian principles, often seen as particularly legitimate and fair (Buchanan and Keohane 2006; Grant and Keohane 2005; Rubenstein 2007), have the added strategic benefit of committing powerful states to mutually acceptable delegation outcomes. On the other hand, the analysis suggests that such egalitarian principles could hide an exploitative relationship or an intense unilateral influence contest.

Consider now the relationship between autonomy and the baseline distribution. If preferences over autonomy $\theta$ and the baseline distribution $y$ vary, might two states profit from a distribution-autonomy exchange?

**Proposition 5.** Consider a delegation contract $(\theta, y)$ such that the autonomy parameter $\theta$ cannot be adjusted for mutual benefit. If the marginal value of autonomy $u'_i(e^*(\theta))$ is high enough for state $i$ while the marginal value of distributional gains $|\frac{\partial v_j(y^*)}{\partial y}|$ is high enough for state $j$, both states benefit from shifting the baseline distribution $y$ towards state $j$’s preference in exchange for increasing autonomy $\theta$.

Although changes in the baseline distribution $y$ are always detrimental for one of the states, their willingness to pay for distributional gains could differ. State $i$ could be willing to accept a less favorable baseline distribution $y$ in exchange for higher autonomy $\theta$ because the cost of poor implementation is so high. These gains from trade highlight an important relationship between the baseline distribution $y$ and autonomy $\theta$. As long as the two states value different things, simultaneous adjustments are something both of them value. Interestingly, Morrow (1991) proves a related result for asymmetric alliances in which a weak state concedes “autonomy” for “security” provided by a powerful state.
Consider an example. If a wealthy donor is concerned with the allocation of foreign aid in a poor recipient country, it could propose a contract that increases the sum of foreign aid at the expense of more stringent conditions (Stone 2004; Svensson 2000; Vreeland 2003). If the recipient is in dire need of funds but the donor is not, both countries could benefit from this exchange. Increasing foreign aid is a side payment that can be loosely thought of as a shift in the policy towards the ideal point of the recipient. If the cost of the side payment is lower to the donor than the cost of changing the outcome per se, the side payment permits international delegation that is otherwise infeasible.

7 Avoiding Unilateral Influence

A most important feature of the delegation game is the commitment problem that stems from unilateral influence. States can freely choose the baseline distribution \( \overline{y} \), so it would be in their common interest to replace the final distribution \( y^* \) with an identical baseline distribution \( \overline{y} = y^* \) and refrain from all unilateral influence. This would not change the expected policy implementation payoff, but it would save the costs of the influence contest.

Are there feasible solutions to this commitment problem? To examine this question, I consider the possibility that states can condition policy implementation on the distance between the baseline and final distribution \( |y^* - \overline{y}| \). Intuitively, states agree on suspending international cooperation if one of them exerts undue or illegal influence. To that end, either state can prevent policy implementation if the distance \( |y^* - \overline{y}| \) is strictly higher than \( d \). The parameter \( d \) measures the difficulty of observing and obtaining verifiable evidence on illegal influence, so it can be thought of as representing the “contractability” of unilateral influence (Bolton and Dewatripont 2005; Lipson 2003; Tirole 1999). As \( d \) increases, it becomes more difficult to hold a state accountable for unilateral influence. For example, \( d \) could increase if third parties, such as non-governmental organizations, lose interest in monitoring the international organization.

Does this additional commitment capacity matter? Consider delegation without conditionality. If the distance \( |y^* - \overline{y}| \) is smaller than \( d \) anyway, imposing conditionality has no effect on state behavior. But if \( |y^* - \overline{y}| \) exceeds \( d \), it is possible to reduce it to \( |y^{**} - \overline{y}| = d \) through conditionality. Figure 2 shows how conditionality can prevent excessive influence.

[FIGURE 2 ABOUT HERE]
Consider now some baseline distribution $\overline{y}$ and examine the effect of conditionality on the payoffs. On the one hand, the shift in the final distribution $y^*$ certainly harms one of the states. On the other hand, one might intuitively expect that the reduced cost of unilateral influence sometimes offsets this effect. This reasoning fails, however, and the cost savings from contractability cannot outweigh the loss of influence.

**Proposition 6.** Fix a contract $(\overline{y}, \theta)$. If conditionality has an effect on the final distribution $y^*$, it is unambiguously detrimental for state $A$ if $y^* > \overline{y}$ and for state $B$ if $y^* < \overline{y}$.

This proposition shows that for a given distribution of gains, as described in the delegation contract, an influential state prefers not to condition international cooperation. While it benefits somewhat from a less costly contest, the lost influence is simply too valuable.

Although the direct effect of conditionality is distributional conflict, it remains to consider possible changes to the delegation contract $(\overline{y}, \theta)$ to compensate the influential state that is losing:

**Proposition 7.** Given a baseline policy $\overline{y}$, if conditionality has an effect on the final policy $y^*$, the baseline policy $\overline{y}$ can always be adjusted so that the equilibrium payoff to both states strictly increases.

Conditionality sows the seeds of distributional conflict, but simultaneously enhancing conditionality and adjusting the baseline policy $\overline{y}$ to benefit the losing state allow a strict Pareto-improvement. To understand why, note that changes in the baseline policy $\overline{y}$ can be made so that the final policy $y^*$ moves only slightly. Since the distance $|y^* - \overline{y}|$ is strictly smaller, however, both states must be exerting less influence. Thus, the disadvantaged state benefiting from conditionality gains because the final policy $y^*$ helps it, while the advantaged state losing from contractability gains because the decrease in bargaining costs helps it.

These observations raise the possibility of an intriguing commitment problem in the second order. Suppose indeed that the disadvantaged state proposes conditionality and promises to adjust the baseline policy $\overline{y}$ in exchange. The advantaged state should certainly accept if it believes that the baseline policy $\overline{y}$ will be changed. But if it believes the baseline policy $\overline{y}$ cannot be changed, it can only expect to lose. Thus, mutually profitable conditionality requires that the influential state also be powerful in bargain-
ing under anarchy. This seems plausible because influence and powerful generally go hand-in-hand. For instance, Ikenberry (2000) shows that the United States exercised “strategic restraint” after the Second World War in exchange for creating a legitimate liberal economic order among the Western allies according to its own preferences. By agreeing on institutional constraints that reduced the vulnerability of its Western allies, the United States was able to widen and deepen international cooperation.

8 Case Studies

As illustrations, I investigate the GEF and the WTO DSU from a North-South perspective. Although the global North had greater influence within the international organization in both cases, the global South had a weaker outside option in international trade because the cost of exclusion from international economic exchange was so high. The design and practice of these two international organizations fit the predictions of the model and thus nicely illustrate the analysis.

8.1 Global Environment Facility

Primarily at the urging of France and Germany, the GEF was established to serve as the focal international organization responsible for North-South aid flows to implement environmental projects in developing countries (Streck 2001, 72). After an unsuccessful pilot phase from 1991 to 1994, the participating countries decided to restructure the GEF (Clémenton 2006, 52-53). The restructured GEF is an “institutional innovation” that strikes a delicate balance between demands made by wealthy donors and the concerns of the developing countries (Young and Boehmer-Christiansen 1997). First, the Council – the main governing body of the GEF – uses a double majority system. Any decisions must be supported both by a majority of the representatives and a group of countries responsible for the majority of contributions. Second, the composition of the Council has the developing countries in a narrow majority. Third, the GEF is a joint enterprise between the World Bank, the UNEP, and the UNDP. The World Bank was dominated by the wealthy industrialized countries, and in the international negotiations on the legal status of the restructured GEF, both the United States and Japan initially insisted that the GEF must operate within the World Bank (Sjöberg 1999, 34). Japan even threatened to suspend all contributions unless this condition was met, but it eventually chose to accept the tripartite compromise. As for the UNEP and the UNDP, Sjöberg (1999, 12) argues that “[t]heir support
for the UN system made the developing countries natural allies with the UN agencies on a number of issues. Both UNEP and UNDP valued their ties with developing countries and often shared and voiced their concerns."

The autonomy of the international bureaucrats at GEF is crucial for successful implementation of environmental projects. First, the GEF focuses on technically complex issue areas, such as global warming and biodiversity conservation (Streck 2001, 73). To implement and evaluate an environmental project, the GEF must first establish a baseline for biodiversity deterioration or the growth of greenhouse gas emissions and then estimate the effect of a proposed environmental project against this hypothetical baseline. The environmental projects must also be subjected to cost-benefit analyses and their economic and social impact must be addressed. As a 1998 evaluation of the performance of the GEF shows, both tasks have posed major difficulties (Porter et al. 1998).

Second, the task of the GEF is greatly complicated by the requirement that the environmental projects be only funded based on their “incremental cost” beyond what the developing country should do to maximize national benefits (Clémencçon 2006, 53). Wealthy donors were unwilling to help the developing countries promote their national interest, so they required that all contributions be used for globally beneficial environmental projects. Evaluating the increment of any given environmental project, however, is an enormously complicated technical task that gives the international bureaucrats a drastic informational advantage (Young and Boehmer-Christiansen 1997, 198).

In the model, the GEF has the following design features. First, the baseline distribution $y$ can be thought of as the governance structure and corresponds to a compromise between the wealthy donors and the poor recipients. This is somewhat counterintuitive, as the structural North-South asymmetry should overwhelmingly favor the wealthy donors (Krasner 1985; Thomas 1987). If North controls all the financial resources, why did they have to compromise? Second, the autonomy $\theta$ of the international bureaucrats is high. This is not surprising given the highly technical nature of the purpose of the GEF, but the analysis predicts that it creates scope for influence. However, the egalitarian governance structure also sets an upper bound for autonomy $\theta$ because the developing countries can vote down reform proposals made by the staff.

The compromise between North and South is congruent with the predictions of the model. While the wealthy donors can choose how much to contribute, the developing countries have virtually no intrinsic interest in the GEF, so they can threaten to simply not implement any environmental
projects (Barrett 1994). As Streck (2001, 83) writes, “[d]eveloping countries were almost exclusively interested in the size of the fund and in extending the coverage of the GEF to areas that donor countries did not regard as clearly global problems, such as desertification and drinking water quality.” The mixed governance structure reflects neither the ideal point of North, the “one dollar one vote” system, nor the ideal point of South, the “one country one vote” system (Streck 2001, 77).

The distribution of influence strongly favors North. As Streck (2001, 92) writes, “while on the organizational level the grievances of the developing world were taken into account, the major donor countries are still the most influential in the system. First, they occupy their own seat on the Council. Second, they have at their disposal the means for staff and coordination of their work on the Council. Some developing countries lack both the financial resources and knowledge to coordinate their interests.” In particular, the central role of the World Bank – an international organization governed by the “one dollar one vote” principle – in administering the financial flows gives North a unique opportunity to influence the practical implementation of environmental projects.

Further, according to an assessment of GEF projects by Porter et al. (1998, 70),

“the requirement for incremental costs is problematic for most recipient country officials. The country studies do refer to a few instances – mostly in the climate focal area – in which project managers in recipient countries have been actively involved in incremental cost calculation. However, the country studies also found that most officials of the relevant agencies had not been involved in initial increment cost calculations. These officials indicated that they not understand the concept of incremental cost and regard it as something done by the Implementing Agency. The agencies confirmed that in the vast majority of cases, international consultants are hired to undertake the incremental cost calculation.”

If the principle of incremental cost is used to determine how many funds a recipient is granted, but the recipient does not even participate in assessing the magnitude of the incremental costs, the wealthy donors have a tremendous advantage. A second evaluation, conducted four years later, found that the situation had not improved: the “team found that both [Implementing Agency] staff and other GEF stakeholders at the country level seemed unfamiliar with, and sometimes uncomfortable about their lack of understanding
of, the economic concepts and the GEF Operational Strategy relating to the incremental costs of delivering global environmental benefits” (Christoffersen et al. 2002).

Finally, as Clémençon (2006, 50) argues, the industrialized countries, and the United States in particular, have used the threat of temporarily suspending contributions. For example, “a controversial performance-based resource allocation framework (RAF) had been adopted by the GEF Council, largely as a precondition for securing U.S. participation in GEF4.” This framework impinges on the recipient country’s sovereignty by conditioning funding on “its performance as reflected in national policies and enabling environments that facilitate successful implementation of GEF projects” (Clémençon 2006, 59). The move is but a continuation of earlier bargaining tactics. Already in the first years after restructuring the GEF, major European donors and Japan had conditioned their contributions on participation by the United States (Clémençon 2006, 52).

These considerations shed light on the design of the GEF. Because the developing countries had a nontrivial outside option, they were able to extract concessions from the developed countries, as conventional bargaining theory predicts. However, the surprising extent to which the governance structure of the GEF reflects principles of universality and equality, as opposed to effectiveness and competence, points to the central empirical implication of my analysis: disadvantaged states with limited influence must obtain a favorable baseline distribution \( \overline{y} \) to delegate in the first place. Although the \textit{ex ante} outside option that South had was relatively strong, its ability to exert influence under the GEF was severely reduced by financial constraints and the dominance of the World Bank. As a result, the distance between the delegation contract and the practical reality, \(| y^* - \overline{y} |\), is unusually high.

This explanation is precise and powerful relative to those found in the extant literature. Principal-agent models predict that the distribution of power is codified in the international agreement (Hawkins et al. 2006a), so it cannot explain egalitarian voting rules that strongly favor the developing countries. Voting rules are irrelevant for realists (Hasenclever, Mayer, and Rittberger 1997; Waltz 1979), so they can explain away egalitarian voting rules, but they cannot explain why other international organizations, such as the IMF, have voting rules that strongly favor major powers such as the United States. Theories that focus on the difference between “deep” and “shallow” cooperation can also explain away egalitarian voting rules by claiming that cooperation is shallow (Downs, Rocke, and Barsoon 1996), but the distributional conflict surrounding these voting rules contradicts the idea that major powers were indifferent.

24
8.2 Dispute Settlement in the World Trade Organization

The WTO DSU is another innovative international institution, designed to resolve international trade disputes in a legalistic fashion. The officially recognized rationale for reforming the largely informal dispute settlement procedures of the GATT was to ensure that “right preserves over might” (Lacarte-Muro and Gappah 2000, 401). The WTO DSU is a particularly interesting application of my theory because wealthy industrialized countries can now formally exert unilateral influence by using their superior legal capacity to obtain favorable rulings. In the GEF, influence is largely informal and often contradicts formal rules, but in WTO dispute resolution, the use of legal capacity does not contradict the letter of international trade law, despite arguably being against the spirit (Alavi 2007; Busch and Reinhardt 2003; Guzman and Simmons 2005; Shaffer 2009).

To begin with, it is essential to understand that in the legalistic framework of the WTO DSU, the key to success in international trade dispute settlement is to win cases. However, as Shaffer (2009, 168-169) writes, “[j]ust to read through and understand the growing WTO case law is an immense task, including for specialized academics.” Indeed, he argues that the complexity of the international trade law has greatly increased over time. Overall, the WTO case law now comprises over 30,000 pages of decisions and other legal documents. The average length of a GATT decision was only twelve pages but the WTO panel rulings range anywhere from 100 to 500 pages. As a result, any state that does not possess substantial legal capacity or is unwilling to invest in legal expertise cannot meaningfully participate in legal dispute resolution.

The complexity of international trade law gives a dispute resolution panel substantial autonomy in policy implementation. Why? As my brief literature review below shows, the official reason for legalization, to level the playing ground, appears largely irrelevant in practice. Perhaps regrettably, legalization has only shifted the locus of influence from overt power politics to competition over legal influence (Busch and Reinhardt 2003; Guzman and Simmons 2005). However, the increasing scope and complexity of international trade law implies that legalized dispute resolution is simply not possible without a sophisticated legal apparatus and a body of precedents. Given that states have chosen to legalize the primary mechanism for settling international trade disputes, autonomy is plainly a necessary condition for successful implementation.

The distribution of power in the reform of the WTO DSU was, in terms of outside options, favorable to the wealthy industrialized countries. As the
success of export-led growth strategies and the failure of import-substitution strategies suggests, developing countries are fundamentally dependent on open access to the international market (Frieden and Martin 2002; Milner 1999). Their ability to shape the rules of international trade is much less pronounced than in the case of international environmental affairs because they cannot credibly threaten to “exit” the world economy (Gruber 2000). Thus, my model predicts that the WTO DSU should reflect the interests of the developed countries to a much greater extent than the GEF.

Second, the distribution of influence also disproportionately favors the wealthy industrialized countries. Given the highly legalistic nature of the WTO DSU, the primary way in which states can influence outcomes is through legal channels. Informal influence is less effective and more costly. The developing countries have not been successful in using the WTO DSU to promote their interests because they lack of legal capacity. Busch and Reinhardt (2003, 720) argue that the most important avenue for concessions under the WTO DSU is to settle the dispute in the shadow of international trade law, and most developing countries have consistently failed to do so: “our results indicate that the central problem for developing countries is that they are missing out on early settlement, not that they boast a worse record in winning pro-plaintiff rulings from Panels or the [Appellate Body] ... we find that this gap is due to a lack of legal capacity, not a lack of market power with which to threaten retaliation.” Guzman and Simmons (2005) concur: “Surprisingly, we find strong evidence that developing countries are constrained by their capacity to launch litigation and no evidence consistent with the power hypothesis.” Alavi (2007, 25) emphasizes Africa’s unfortunate position: “Although the importance of an adequate trade-policy infrastructure is difficult to underestimate, some of the more specific problems facing [sub-Saharan African] countries seem to be rooted in the nature of the [DSU] itself.”

My model offers an explanation for these design features. The baseline distribution $\overline{y}$ of de jure neutrality does not explicitly favor either group of countries, but asymmetries of legal capacity ensure that the wealthy industrialized countries dominate the WTO DSU in practice. Their superior legal capacity implies that the cost $\alpha_i \cdot c(z^*)$ of the unilateral influence contest is not an important factor, so the benefits of winning cases and setting precedents at the WTO DSU outweigh the cost of mustering the requisite legal capacity.

In the GEF, the industrialized countries had to commit to egalitarian voting rules because the developing countries had strong outside options. In the WTO, the developing countries did not have strong outside options. The
industrialized countries did not have to accept built-in provisions that help developing countries develop legal capacity. For instance, as Michalopoulos (1999) and Shaffer (2005) show, wealthy industrialized countries have done little to help the developing countries develop the requisite legal capacity. If the developing countries had stronger outside options, they might have demanded that the WTO DSU be accompanied by significant efforts to build legal capacity. Equally important, the industrialized countries did not simply force rules of international dispute resolution that clearly favor them. My model gives a parsimonious explanation: the combination of strongly favorable rules and superior legal capacity would have made the dispute resolution mechanism too costly for the developing countries.

9 Conclusion

In this article, I have focused on the problem of unilateral influence in international delegation. The extant literature has examined distributional conflict over the design of international agreements, but it has not investigated the time-inconsistency problems that arise if states cannot commit to restraint from unilateral influence in international organizations. My analysis reveals an important tradeoff in international delegation. By giving international organizations autonomy, states improve policy implementation by taking advantage of expertise. Conventional wisdom holds that the cost of this autonomy is agency slack, but the possibility of a unilateral influence contest points to two additional obstacles. First, if a weak state cannot trust that a powerful state refrains from biasing policy implementation, mutually profitable delegation is not possible. Second, the unilateral influence contest is costly to both states.

These results refine the nascent science of international delegation. While the standard principal-agent model helps us understand managerial problems in international delegation, this article focuses on the explicitly political problem of distributional conflict. It combines insights from international cooperation theory, and particularly the idea that international institutions can be “rationally designed” (Koremenos, Lipson, and Snidal 2001), with the realist notion that power politics are an essential element of international politics (Drezner 2007; Gruber 2000). The theoretical and empirical implications show that the design of international institutions is a complex strategic problem. In particular, states must understand that the rules of international institutions are imperfectly enforceable, and they must be chosen so that the expected effect of the unilateral influence contest is accounted
for. Most surprisingly, international institutions must often reflect the interests of weak states. This and the numerous other paradoxical implications of the analysis create exciting opportunities for progressive theoretical and empirical research in the future.
Appendix

Equilibria

Show that a unique subgame-perfect equilibrium exists given any contract \((\mathbf{y}, \theta)\). First, consider existence. The result is immediate for \(\theta = 0\). Fix \(\theta > 0\). A strictly increasing and continuously differentiable unique best response \(\tilde{z}_j(z_i)\) exists by the properties of \(v_j, c\). By strict convexity of \(c\), the derivative \(\tilde{z}_j'(z_i)\) approaches zero for high enough \(\tilde{z}_j\). This guarantees that the two best responses \(\tilde{z}_i, \tilde{z}_j\) meet at least once. Second, consider uniqueness. Towards a contradiction, let \((z^*_A, z^*_B)\) be a subgame equilibrium and suppose there exists another subgame equilibrium \((z^{**}_A, z^{**}_B)\). The best responses are strictly increasing. Set without loss of generality \((z^{**}_A, z^{**}_B) > (z^*_A, z^*_B)\). By the properties of \(v_A, v_B, c\), there is no final distribution \(y^*\) for which (4) is met for both states \(i = A, B\) simultaneously. Since \(c'(0) = 0\), only interior equilibria are possible, so the claim follows.

Consider now subgame-perfect equilibria at the contracting stage. Any commonly proposed contract \((\mathbf{y}, \theta)\) that induces a final distribution \(y^*\) such that \(v_i(y^*) \cdot u_i(e^*(\theta)) \geq \lambda_i\) for \(i = A, B\) is part of some subgame-perfect equilibrium of the game because no profitable deviation exists. If state \(A\) proposes \((\mathbf{y}_A, \theta)\) such that \(\mathbf{y}_A \to \infty\) while state \(B\) proposes \((\mathbf{y}_A, \theta)\) such that \(\mathbf{y}_B \to -\infty\), no profitable deviation exists either, so any such pair of proposals is a subgame-perfect equilibrium. A Pareto-frontier exists because the payoff to state \(A\) (\(B\)) given a contract \((y, \theta)\) is strictly increasing (decreasing) in \(y\) by Proposition 1 and autonomy \(\theta\) is chosen on the closed interval \([0, 1]\).

**Proof of Claim 1**

The payoff from policy implementation \(v_i(y_i) \cdot u_i(e^*)\) is strictly increasing in equilibrium effort \(e^*\). Equilibrium effort \(e^*\) maximizes \(g(c, \theta)\) and the maximizer is strictly increasing in autonomy \(\theta\).

**Proof of Claim 2**

The payoff \(v_i(y_i) \cdot u_i(e)\) is increasing and strictly concave in \(y_i\) for state \(i\). The effect of influence \(z_i\) on \(y_i\) is positive linear. The cost \(\alpha_i \cdot c(z_i)\) of influence \(z_i\) is increasing and strictly convex in \(z_i\). Thus, the unilateral influence contest has a unique solution. Examine (4) to see that for \(\theta' < \theta''\), this solution is found at \((z'_A, z'_B) \leq (z^*_A, z^*_B)\), where the inequality is strict for an interior equilibrium.
Proof of Claim 3

If \( y^* \to \infty \) or \( y^* \to -\infty \) given \((\overline{y}, \theta)\), at least one state \( i \) obtains an equilibrium payoff that is arbitrarily close to zero, so the \textit{ex ante} outside option \( \lambda_i \) will be a profitable deviation. ■

Proof of Proposition 1

Let \( \overline{y} \) increase by a small amount \( \Delta \overline{y} \) and examine (4) to establish that the first-order condition cannot be simultaneously met for both states \( i \) at any final policy \( y^* \) lower than previously. To see that \( z^*_A \) must decrease, first note that \( v_A \) is strictly concave in \( y \). By convexity of \( c \), the unique solution \( z^*_A \) to (4) given an increased \( y^* \) must decrease. To see that \( z^*_B \) must increase, proceed similarly to the case of state \( A \). ■

Proof of Proposition 2

Examine (4) to see that with \( v_i \) strictly concave and \( c \) strictly convex, equilibrium influence \( z^*_i \) must increase with autonomy \( \theta \) for both states \( i \). Now let \( \alpha_A \to 0 \) and \( \alpha_B \to \infty \). Now the increase in \( z^*_B \) is negligible while the increase in \( z^*_A \) is non-negligible. As a result, \( y^* \) increases. Exactly the opposite holds if \( \alpha_A \to \infty \) and \( \alpha_B \to 0 \). ■

Proof of Proposition 3

If \( e^*(\theta) \) increases rapidly enough in \( \theta \) while \( \alpha_i, \alpha_j \) are high enough, the equilibrium influence cost \( \alpha_i \cdot c(z^*_i) = \alpha_i \cdot \int_0^{z^*_i} c'(z_i) dz_i \) given by (4) is negligible for both states \( i \) while the policy implementation payoff (2) increases rapidly with \( \theta \). Exactly the opposite holds if \( e^*(\theta) \) increases slowly in \( \theta \) while \( \alpha_i, \alpha_j \) are low enough. ■

Proof of Proposition 4

Fix \( \theta \). Let \( \alpha_i \to \infty \) and \( \alpha_j \to 0 \). Then the distance \( |y^* - \overline{y}| \) will be large, so \( v_i(y^*) \to 0 \) unless \( \overline{y} \) is high (low) for \( i = A \) (\( i = B \)). But then (5) cannot hold for a given \( \lambda_i > 0 \). ■

Proof of Proposition 5

For sufficiently high \( u_i' \), small increases in autonomy \( \theta \) have a non-negligible effect on the payoff from delegation (3) that dominate the effect of increased cost \( \alpha_i \cdot c(z^*_i) \) for any given baseline policy \( \overline{y} \). By assumption, state \( j \) cannot
benefit from an increase in autonomy $\theta$, so a small increase in autonomy $\theta$ produces a large payoff increase for state $i$ and a small payoff decrease for state $j$.

By Claim 1, small shifts in $\gamma$ toward state $j$’s ideal point result in a shift of $y^*$ in that direction. For sufficiently high $\left| \frac{\partial v_j(y^*)}{\partial y} \right|$, small shifts in $\gamma$ toward state $j$’s ideal point have a non-negligible effect on the payoff from delegation (3) that dominate any effects on the unilateral influence contest. Thus, the shift produces a large payoff increase for state $j$ and a small payoff decrease for state $i$. □

Proof of Proposition 6

Without loss of generality, let $y^* > \gamma$. Conditionality has an effect on the final policy $y^*$ if and only if $y^* > \gamma + d$, in which case $y^*$ decreases. Thus, $v_A(z_A^*)$ decreases while $v_B(y^*)$ increases. Additionally, $z_A^*, z_B^*$ decrease. It remains to show that the decrease in $\alpha_A \cdot c(z_A^*)$ does not dominate the decrease in $v_A(z_A^*) \cdot u_A(e^*(\theta))$. Suppose towards a contradiction this was the case. With $z_B^*$ lower than previously and $c$ strictly convex, state $A$ could have previously profitably deviated from $y^*$ to $\gamma + d$, so the strategy profile prior to conditionality could not have been an equilibrium, a contradiction. □

Proof of Proposition 7

By Proposition 1, the final policy $y^*$ varies continuously and monotonically in the baseline policy $\gamma$. Impose conditionality $d > 0$ and set $\gamma$ such that the final policy $y^*$ remains unchanged. Now $\left| y^* - \gamma \right|$ decreases, so at least one state $i$ strictly benefits because it has to exert less influence $z_i^*$ to achieve $y^*$. For the other state $j$, the payoff remains unchanged. Now shift $\gamma$ slightly to the advantage of state $j$. □
Figure 1. The best-response correspondences $\tilde{z}_A, \tilde{z}_B$. 
Figure 2. Upon contractability, previous influences $z^*_A, z^*_B$ are replaced by lower influences $z^{***}_A, z^{***}_B$. If contractability limits the influence of state $A$, it chooses the influence $z^{***}_A$ such that $y^* = \overline{y} + d$ in equilibrium.
References


