Global Governance 14 (2008), 479–501

Compliance and Performance in International Water Agreements: The Case of the Naryn/Syr Darya Basin

Thomas Bernauer and Tobias Siegfried

Many case studies and some large-N research have shown that upstream-downstream cooperation in international river basins occurs quite frequently. The same holds for global water governance efforts more generally. Yet such findings are blind in one eye because they focus primarily on political commitments or compliance with international agreements. A policy performance metric (PER) allows for a more substantive assessment of success or failure in international water governance. To test its usefulness, this article applies this metric to the Naryn/Syr Darya basin, a major international river system in Central Asia. Management of the Toktogul reservoir, the main reservoir in the Naryn/Syr Darya basin, was internationalized in 1991 when the Soviet Union collapsed. *Compliance* with an international agreement, concluded in 1998, has been quite high. This agreement establishes an international trade-off between water releases for upstream hydropower production in winter and water releases for downstream irrigation in summer. However, *performance* of this agreement over time has been very low and highly variable. The management system in place is therefore in urgent need of reform. Studies of international and global water governance should pay more attention to the degree to which political commitments actually further de facto problem solving. KEYWORDS: international cooperation, governance, compliance, performance, water management, Naryn/Syr Darya, Toktogul dam.

The scientific literature on water governance issues has experienced a boom in recent years.[1] It has produced innovative concepts and theories that help make sense of hundreds if not thousands of collaborative efforts that are under way in water systems around the world at various levels, from the local to the global. Yet one weakness in the existing literature is its heavy focus on legal arrangements and institutional processes. Scant attention is given to the nexus between policy measures and changes in hydrological systems. Studying these connections is necessary to determine whether water governance efforts are effective not only in a discursive, legal, or institutional sense, but also in terms of solving concrete problems on the ground. Research on this issue creates exciting opportunities for collaboration across the social, natural, and engineering sciences, as is evident in this article, which has been coauthored by a political scientist and an environmental engineer.

The article speaks to the special issue theme of global water governance primarily by offering an analytical tool that helps to assess the performance of particular water governance efforts based on explicit and transparent standards. This performance assessment tool is useful for diagnostic purposes—that is, it identifies governance efforts that require improvement. The tool is also useful for comparing governance efforts between water systems, across political scales, and over time. It can thereby generate more generic insights that can inform efforts to establish water management principles at the global level.

Cooperation, when defined as a dependent variable in causal explanations of international water management, is usually measured in binary terms—that is, with a yes/no answer to the question whether an agreement, treaty, or international institution is in place. Examples of this approach can be found in the many qualitative case studies on international water management[2] and the few large-N quantitative studies that exist on the subject.[3]

Many case studies also assess the degree of substantive international cooperation. However, the criteria against which the depth of cooperation is measured differ across studies, and the assessments are usually qualitative.[4] Moreover, most assessments rely on noncausal criteria. The most common approach is to describe, over time, the development of a particular problem targeted by a cooperative effort (e.g., pollution) and to assess compliance with international obligations. This is usually done without a systematic analysis of whether international cooperation has, ceteris paribus, brought changes in environmental outcomes and in compliance levels. Coding of the contents of international agreements for purposes of measuring the depth of cooperation in large-N analysis is still in its infancy.[5]

Another approach has been to code cooperative and conflictual events among riparian countries, but this approach offers only indirect insights into the depth of cooperation.[6] International water management efforts are to some extent directly included in the codings of cooperative events. Moreover, deep cooperation may often be accompanied by conflict events. More cooperation than conflict may thus tell us little about whether international cooperation performs well in terms of problem solving.

Another line of research uses environmental parameters as proxies for cooperation. For example, two recent studies examine whether trade ties and other factors promote international efforts to clean up water pollution.[7] Since environmental outcomes are measured without causal reference to international cooperation (cleaner transboundary water is simply assumed to indicate more cooperation), this approach does not offer direct insights into the success or otherwise of cooperation.

Substantial progress has been made in recent years in measuring the performance (or depth) of international cooperation. Building on previous work,[8] the first part of this article outlines a methodology for estimating the performance of international cooperation. This policy performance metric (PER) is a time-dependent function of: (1) the outcome that should ideally be reached (optimum performance); (2) the outcome of a given policy at the time of measurement (actual performance); and (3) the outcome that would have occurred in the absence of this policy (counterfactual performance).

The PER measure has several advantages. First, it makes explicit reference to optimal performance and thus the target level for problem solving. Second, it focuses explicitly on the causal relationship between international policies and outcomes. Third, it can be used not only to assess international policy performance at specific points in time in contexts marked by rather little data, but also to assess performance dynamics over time in contexts where more data exist. Fourth, the measure also allows a disaggregation of cooperative efforts with reference to particular objectives.

The PER approach has broader relevance in that it addresses an ongoing debate in international relations about situation structures and their effects on international cooperation.[9] With respect to international water policy, this debate has concentrated on the difficulties of handling upstream-downstream settings where preferences of the countries involved are often antagonistic. Recent quantitative and qualitative research suggests that upstream-downstream cooperation is quite frequent.[10] However, the empirical evidence remains controversial. For example, M. Brochmann and N. P. Gleditsch find contradictory effects of upstream-downstream settings on international cooperation.[11] According to their analysis, international cooperation in water issues is more likely in upstream-downstream circumstances than in other settings in the time period 1820–2001, but the effect is insignificant in the sample period 1975–2001. When cooperation is measured by signed treaties (instead of ratified treaties), the effects are largely insignificant. Brochmann and Gleditsch also find that upstream-downstream settings produce both more cooperative and more conflictive events. The latter result indicates that such settings lead to more interaction, but it does not reveal whether such settings facilitate or hinder cooperation. Other, more process-oriented studies show that compensation or issue linkages to offset upstream-downstream asymmetries are often difficult to construct and that cooperation, if it emerges at all, remains shallow. J. Tir and J. T. Ackerman conclude that international water treaties are less likely in upstream-downstream settings.[12] T. Bernauer shows that it took countries of the Rhine River basin several decades to reduce upstream-downstream water pollution and that forces other than international cooperation have been key.[13] If upstream-downstream cooperation is very difficult among highly developed democratic countries, one should expect even greater difficulties in achieving similar levels of cooperation in less fortunate regions of the world.

The approach we take in this article is relevant also from a practical viewpoint. It focuses on problem solving and relates policy measures to specific notions of what should ideally be achieved (optimal performance) and what would have happened without cooperation (counterfactual performance). The PER tool thus produces a more accurate and policy-relevant diagnosis. This, in turn, provides a better foundation for finding ways to make cooperation more effective.

To demonstrate the empirical relevance of the PER concept, we examine international water management in the Naryn/Syr Darya basin, a major international river system in Central Asia. The analysis focuses on the Toktogul reservoir, the main reservoir in the Naryn/Syr Darya basin, and its downstream effects. The principal policy challenge in this case has been to design and implement an international trade-off between, on the one hand, water releases for upstream hydropower production in winter and, on the other hand, water releases for downstream irrigation in summer. That is, the main issue of conflict and cooperation in this case involves upstream-downstream water allocation. This situation facilitates measurement of the parameters in the PER metric and also allows for a systematic comparison between a compliance-based and a performance-based assessment of international water governance. The availability of new data for the Naryn/Syr Darya case has created an opportunity for the first systematic assessment of the performance of this water governance system. The Naryn/Syr Darya case is also interesting because it involves a transition from a top-down domestic water governance system in Soviet times to a more horizontal international scheme since 1998.

We start by examining institutional outcomes and the compliance-related behavior of riparian countries. A detailed agreement for the Naryn/Syr Darya basin was concluded in 1998, and compliance with this agreement is high. We then apply the PER measurement concept, noting that compliance could be high only because international obligations are weak and cooperation is therefore shallow. In other words, this second step assesses whether good news about compliance is also good news about cooperation.[14] This analysis shows that implementation of the 1998 agreement has in fact been characterized by low performance and high variability.

The principal policy implication of this finding is that, even though compliance is high, the management system in place is in urgent need of reform. We consider some options for improvement. The more general message is that many international upstream-downstream water agreements may rest on only shallow cooperation, and no quick fixes to such problems can be expected. Studies of international or global water governance should thus pay more attention to the relationship between de jure commitments and de facto problem solving.

Measuring Performance

The starting point for this analysis is a simple formula suggested by C. Helm and D. Sprinz,[15] in which

$$PER = \frac{AP - CP}{OP - CP} \qquad (1)$$

Here *AP* refers to actual performance, *CP* stands for counterfactual performance, and *OP* designates optimal performance. In international water management, a *PER* calculation might relate to hydropower production, irrigation water provision, water quality, or water provision for ecosystem functions.

PER can be estimated in relation to any public demand addressed by a public policy. In effect, this equation captures the extent to which a given problem has actually been solved ($AP - CP$) relative to the problem-solving potential ($OP - CP$). The first calculation alone would only indicate that the relevant policy has had some effect. Only by adding the second calculation (and *OP* in particular) do we gain information on the extent to which the problem has been solved. Moreover, adding the second calculation ($OP - CP$) facilitates comparisons across policies within and across policy domains and over time. Provided that one distinguishes between maximizing ($CP \geq AP \geq OP$) and minimizing ($CP \leq AP \leq OP$) cases, the *PER* measure sets a lower and an upper bound and (with some exceptions) standardizes values between 0 and 1.

We have developed a more complex version of the above formula.[16] It solves some conceptual problems in the simple formula (e.g., overcompliance scenarios and inefficiencies associated with them) and allows for the measurement of performance and its variation over time. Readers less interested in technical details can skip the remainder of this section and move to the empirical application to the Naryn/Syr Darya.

The more complex version of our performance assessment concept starts with the following definition

$$PER^{*}(t) = 1 - \left| \frac{AP(t) - OP(t)}{CP(t) - OP(t)} \right| \qquad (2)$$

where $PER^{*}(t)$ is a measure of policy performance at time t. $PER^{*}(t)$ measures performance relative to optimal performance *OP* at a specific observation time t. If we use the notation $\delta_{AP}(t) = |AP(t) - OP(t)|$ and $\delta_{CP}(t) = |CP(t) - OP(t)|$, then equation (2) becomes

$$PER^{*}(t) = 1 - \frac{\delta_{AP}(t)}{\delta_{CP}(t)} \qquad (3)$$

by the definition of the absolute value and its properties. If $CP(t)<AP(t)<OP(t)$ or $CP(t)>AP(t)>OP(t)$, the two performance measures as defined by equations (1) and (2) are equal, i.e., $PER^{*}(t)=PER$. Note that according to equation (3), $PER^{*}(t)$ is defined as long as $\delta_{CP}(t)\neq 0$.

Estimating performance over time means that one must look at *AP, CP,* and *OP* in terms of times-series data, that is, $AP(t)$, $CP(t)$ and $OP(t)$ (as well as the derived $\delta_{AP}(t)$ and $\delta_{CP}(t)$). In the subsequent analysis, the focus is restricted to stationary processes. The expected value as well as the variance of $PER^{*}(t)$ are used to characterize policy performance over time. The expected value of $PER^{*}(t)$ can be approximated by

$$\left\langle PER^{*}\right\rangle = 1 - \frac{\mu_{\delta_{AP}}}{\mu_{\delta_{CP}}} + \frac{1}{\mu^{2}_{\delta_{CP}}}\mathrm{Cov}(\delta_{AP},\delta_{CP}) \qquad (4)$$

where $\mathrm{Cov}(\delta_{AP,}\delta_{CP})$ denotes the covariance and $\mu_{\delta_{AP}}$ as well as $\mu_{\delta_{CP}}$ the mean of the time series $\delta_{AP}(t)$ and $\delta_{CP}(t)$.

The variance can be approximated by

$$\sigma^{2}_{PER^{*}} = \frac{4\sigma^{2}_{\delta_{AP}}}{\mu^{2}_{\delta_{CP}}} - \frac{\mu^{2}_{\delta_{AP}}\sigma^{2}_{\delta_{CP}}}{\mu^{4}_{\delta_{CP}}} - \frac{\mathrm{Cov}(\delta_{AP},\delta_{CP})^{2}}{\mu^{4}_{\delta_{CP}}} - \frac{2\,\mathrm{Cov}(\delta_{AP},\delta_{CP})\mu_{\delta_{AP}}}{\mu^{3}_{\delta_{CP}}} \qquad (5)$$

In equations (4) and (5), $\mu_{\delta_{AP}}$, $\mu_{\delta_{CP}}$, $\sigma^{2}{}_{\delta_{AP}}$, $\sigma^{2}{}_{\delta_{CP}}$ and $\mathrm{Cov}(\delta_{AP,}\delta_{CP})$ have to be estimated empirically from available data.

International Water Management in the Naryn/Syr Darya Basin

The Syr Darya River originates as the Naryn River in the mountains of Kyrgyzstan (see Figure 1). It then flows through Uzbekistan and Tajikistan and ends in the Aral Sea in Kazakhstan. Its total length is around 2,800 kilometers. About 20 million people inhabit this river catchment, which covers an area of some 250,000 square kilometers. The river is mainly fed by snowmelt and water from glaciers. The natural runoff pattern, with annual flow ranges of 23.5–51 cubic kilometers (around 40 cubic kilometers in the past few years), is characterized by a spring/summer flood that usually starts in April and peaks in June. About 90 percent of the Naryn/Syr Darya's mean annual

Figure 1 Naryn/Syr Darya Catchment

Source: T. Siegfried and T. Bernauer, "Estimating the Performance of International Regulatory Regimes: Methodology and Empirical Application to International Water Management in the Naryn/Syr Darya Basin," *Water Resources Research* 43 (2007) (W11406; doi: 10.1029/2006WR005738).

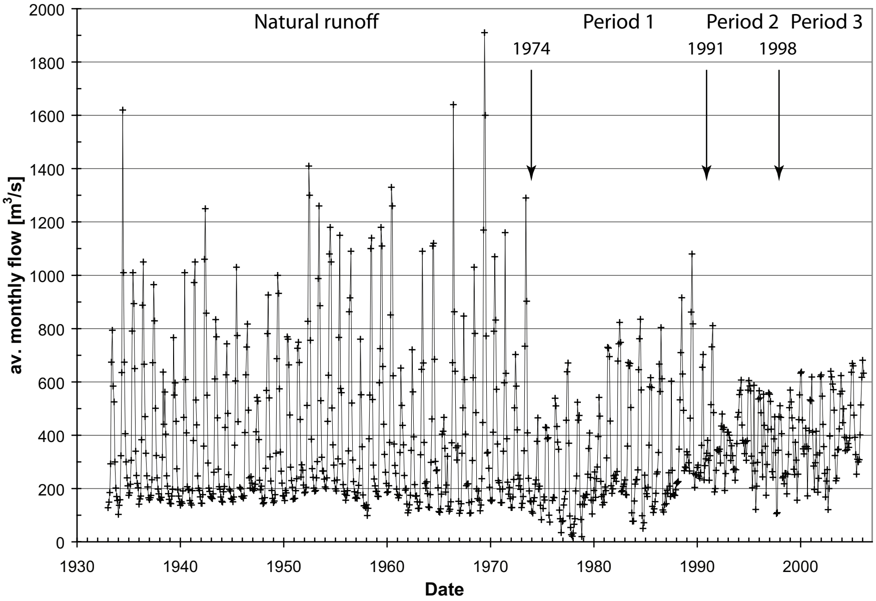
flow is now regulated by storage reservoirs. Approximately 75 percent of the runoff comes from Kyrgyzstan.[17] Consumptive water allocation from the Naryn/Syr Darya Basin is mainly for irrigated farming.

The runoff of the Naryn/Syr Darya, as measured at the Uch Kurgan gauge station, has varied strongly over time. As shown in Figure 2, the variation is characterized by four distinct periods. When the runoff was natural (1933–1974) and determined entirely by seasonal and climatic variability, the mean flow was around 390 cubic meters per second, with a high variability in summer. A substantial change in flow patterns occurred with the commissioning of the Toktogul dam in 1974. This event marks the beginning of the first river management period (1974–1990), which was characterized by centralized management by the USSR of the Toktogul reservoir and the river basin as a whole. The Toktogul dam is by far the largest storage facility in the Aral Sea basin, with a total storage volume of around 19.5 cubic kilometers, accounting for more than half of the total usable reservoir capacity in the whole Naryn/Syr Darya basin. The reservoir area is around 280 square kilometers, its length about 65 kilometers. The hydropower capacity of the Toktogul power plant is 1,200 megawatts, making it the second biggest in the Aral Sea basin.[18] After the dam was commissioned, a general attenuation of peak downstream flows was observed (see Figure 2). Moreover, an overall decline of monthly flow variability occurred, especially in the summer months.

During this first management period, the system was oriented primarily toward water provision for irrigated agriculture (particularly cotton production) in Uzbekistan and Kazakhstan. The timing of winter and summer flow releases did not change substantially compared with the natural runoff pattern. This is indicated by seasonal ratios *r* of inflow versus outflow that oscillate around $r=1$ (see inflow/outflow ratios for 1980–1990 in Figure 3).

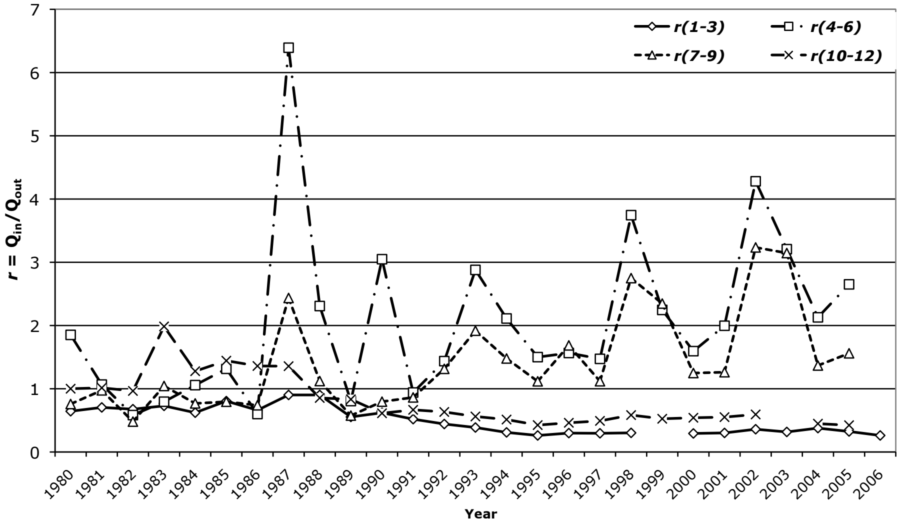
In the early 1980s, a water management organization for the Naryn/Syr Darya was set up in Tashkent, Uzbekistan. Its mandate was to operate and maintain all headwater structures with a discharge of more than 10 cubic meters per second. This management system and its infrastructure were fully funded from the federal budget of the USSR. In consultation with the governments of the riparian republics and using forecasts by the Central Asia Hydromet Service, the Ministry of Water Resources (Minvodgoz) in Moscow defined annually (based on a multiyear master plan) how much

Figure 2 Mean Monthly Flow of the Naryn/Syr Darya River at the Uch Kurgan Gauge



Sources: Global Runoff Data Center (GRDC); and Andrey Yakovlev, Uzbek Hydrometeorological Service, Tashkent, Uzbekistan.

Figure 3 Ratios *r* of Inflow to Outflow from the Toktogul Reservoir, Averaged over Three Months



Source: Andrey Yakovlev, Uzbek Hydrometeorological Service, Tashkent, Uzbekistan.

Notes: The switch from cooperative to noncooperative water resources management is characterized by the opening up of a gap from 1991 onward between inflow/outflow ratios in the winter (months 1–3, January–March; and months 10–12, October–December) and the summer (months 4–6, April–June; and months 7–9, July–September). The pronounced peaks in the summer month ratios characterize years of above average summer runoff.

water was to be released for irrigation during the growing season (April to September).

Minvodgoz was responsible for implementing the water allocations and maintaining the infrastructure. It also had the authority to increase or reduce allocations to each Soviet republic by up to 10 percent. The electricity produced at Toktogul during that period went into the Central Asian Energy Pool (CAEP) and was thus shared among the riparian republics. In exchange, the neighboring republics supplied coal, oil, and natural gas to Kyrgyzstan in winter to cover increased Kyrgyz energy demand during the colder months.[19]

The second river management period, as depicted in Figure 2, commenced with the collapse of the Soviet Union in 1991. This event brought an end to centralized management of water resources and water-energy tradeoff arrangements. Very quickly the newly independent states became involved in disputes over water allocation. Coal, oil, natural gas, and electricity supplies to Kyrgyzstan declined dramatically between 1991 and 1997.

Thermal and electric power output of Kyrgyz energy plants declined by more than half relative to the 1991 values. Since Kyrgyzstan has no fossil fuel sources of its own, it cannot rely on domestic fossil fuel for electricity production and thermal energy.

This circumstance, in turn, increased winter demand for hydropower by more than 100 percent. Purchases of energy from abroad were (and still are) difficult, because the government was (for political and administrative reasons) unable to increase and collect appropriate energy tariffs. Moreover, financial contributions from Moscow and the former republics in the basin for the maintenance of the reservoir ceased. In response to these developments, Kyrgyzstan switched the operation of the Toktogul reservoir from an irrigation to an electric power production mode. Since the winter of 1993, water flows have no longer peaked in summer but rather in winter. This change has opened a gap between the summer inflow/outflow ratios r and their winter counterparts, as seen in Figure 3.

The main political problem since 1991 concerned upstream-downstream antagonisms. Upstream interests derived from seasonal water demands are diametrically opposed to downstream water demands and interests. Kyrgyzstan is eager to store water between spring and autumn and to release this water between winter and spring for hydropower production. Conversely, downstream Uzbekistan and Kazakhstan, by far the largest consumers of irrigation water in the river basin, wish to obtain much more water during the growing season (April to September) than in the nongrowing season (October to March). They are also interested in electricity for operating irrigation pumps, as is produced upstream through water releases during the growing season. Moreover, downstream countries prefer low water releases in winter, because high flows in winter may cause floods due to ice in the river bed, which reduces water flow capacity.[20] Thus, the principal problem to be solved is to coordinate the management of the Naryn/Syr Darya cascade of reservoirs that are located entirely in Kyrgyzstan, and in particular the handling of trade-offs between consumptive water use for downstream irrigation purposes in summer and nonconsumptive use for upstream energy production in Kyrgyzstan in winter.

International negotiations focusing on the management of the Toktogul reservoir began shortly after the demise of the USSR. In February 1992, the five newly independent riparian states of the Naryn/Syr Darya basin set up the Interstate Commission for Water Coordination (ICWC). They agreed to keep the water allocation principles of the former USSR in place until a new system could be established, albeit without the funding for the infrastructure that previously came from Moscow. The most important hydraulic structures, in particular the biggest reservoirs in the basin (including the Toktogul), were *not* put under the control of the ICWC. That is, they were de facto nationalized by the newly independent countries.

This period of unilateralism continued until March 1998, when under the aegis of the Executive Committee of the Central Asian Economic Community, and assisted by the United States Agency for International Development (USAID), Kazakhstan, Kyrgyzstan, and Uzbekistan signed an agreement. This accord marks the beginning of Period 3, as defined in Figure 2. Tajikistan joined this agreement in 1999.[21] The release schedule for the Toktogul reservoir, the main element of the agreement, is shown in Table 1.

The 1998 accord includes a general framework agreement and a specific barter agreement on water-energy exchanges in 1998. The barter agreement holds that in the growing season (April 1–October 1), Kyrgyzstan will supply 2.2 million kilowatt hours (MkWh) of electricity to Kazakhstan and Uzbekistan (1.1 MkWh each). In exchange, Kazakhstan and Uzbekistan agree to deliver specific amounts of electricity, natural gas, fuel oil, and coal to Kyrgyzstan in specific months under conditions set forth in bilateral agreements concluded in 1997. Compensation can also be carried out in the form of "other products" (labor and services are mentioned) or money. Kyrgyzstan agreed to cut its energy consumption by 10 percent against 1997 levels. The framework agreement, also concluded in March 1998,[22] holds that these exchanges will subsequently be defined annually through negotiations.

In other words, the water management system put in place in 1998 holds that during the vegetation season, Kyrgyzstan releases more water than it needs for its own hydropower demand, and the resultant energy surplus is distributed to Kazakhstan and Uzbekistan. Outside the growing season (October 1–April 1), Uzbekistan and Kazakhstan supply Kyrgyzstan with energy resources in amounts that are approximately equivalent to the electricity they receive from Kyrgyzstan during the growing season. The exact amounts of water and energy are defined annually through negotiations among the governments. Typically, Kyrgyzstan has been scheduled to release around 6.5 cubic kilometers of water during the vegetation period and transfer about 2.2 MkWh of electricity to Uzbekistan and Kazakhstan.

Compliance

From the viewpoint of policy measures, the 1998 agreement has been good news. Its design follows the pattern of other international upstream-downstream

Table 1 Release Schedule of Toktogul Reservoir as Established in the 1998 Treaty

Month	1	2	3	4	5	6	7	8	9
q [m^3/s]	495	490	300	230	270	500	650	600	190

Source: Data available at http://ocid.nacse.org/tfdd/index.php.
Note: No values were defined for the months of October to December.

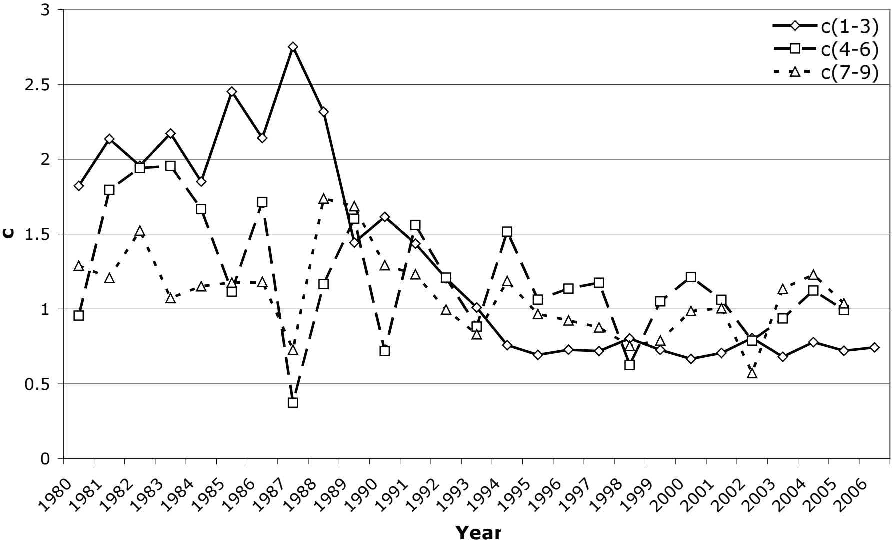
water agreements in that it addresses strong asymmetries of interests through economic exchanges. Moreover, the 1998 agreement for the Naryn/Syr Darya was reached relatively quickly, only seven years after the collapse of the Soviet Union. In other instances of strong upstream-downstream asymmetries, it has taken the riparian states several decades to arrive at an international agreement. Such was the case, for example, regarding chloride in the Rhine River and salinity levels in the Colorado River.[23]

To evaluate international cooperation with respect to the Naryn/Syr Darya regime of 1998, one can move beyond simple questions of "agreement, yes/no" and "how long it took to come to an agreement." *Compliance* is an obvious way to do so. Measuring compliance indicates to what extent the parties have fulfilled their obligations under the 1998 agreement. In the Naryn/Syr Darya case, this requires an assumption. The release schedule, as shown in Table 1, was set for the year 1998, with schedules for subsequent years to be negotiated annually. The riparian countries have not adopted and published any revised schedules since then. However, there has been a tacit bargain among the riparian states on maintaining the existing schedules. It may therefore be assumed that the 1998 schedule constitutes the benchmark for compliance in subsequent years as well. This assumption affects only the compliance assessment and has no relevance for the performance assessment.

To assess compliance, we computed ratios of actual water releases from the Toktogul reservoir (three-month averages) and the targets for the respective months as defined in the 1998 agreement. Figure 4 shows the results. As compared against a perfect compliance score of 1, average actual compliance levels were 1.6 in 1980–1990, 1.1 in 1991–1997, and 0.9 in 1998–2006. Hence the overall picture is one of high compliance, particularly in spring to autumn, and somewhat lower compliance (by 25 percent) in winter.

As noted by G. W. Downs, D. M. Rocke, and P, N. Barsoom, even high levels of compliance in international regulatory regimes do not necessarily imply good news about international cooperation.[24] The problem is that states often define treaty commitments so that meeting them requires little or no effort above and beyond what the states concerned would do in the absence of the respective international commitments. Consequently, a low level of compliance could still involve very substantial international cooperation if the states in question have engaged in very ambitious commitments. Conversely, high levels of compliance could involve very shallow cooperation if commitments merely register the status quo ante of state behavior. To deal with this problem, policy performance is determined in the following section using a measure of optimal performance *(OP)* and counterfactual performance *(CP)* rather than the 1998 treaty targets as benchmarks against which actual state behavior is compared.

Figure 4 Compliance with the 1998 Agreement



Notes: Compliance *c* in months 1–3 (January–March) is defined as the 1998 target for these months divided by the actual water release; compliance in months 4–6 (April–June) and 7–9 (July–September) is defined as the actual water release divided by the 1998 target. These definitions are based on the assumption that exceeding the target in the growing season is better for downstream countries than exceeding the target in winter. We also show the results for the years before 1998 to point out the general trend, though the 1998 agreement was, of course, not in force before 1998.

Performance

The measurement of performance is approached here in two steps. First, the simple performance formula (equation 1) is applied to long-term average runoff data. Second, the more complex performance metric (equations 4 and 5) is applied.

Whereas actual performance *AP(t)* is clearly defined in terms of the water releases in Period 3 (1998–2006), calculation of the counterfactual performance *CP* and the optimal performance *OP* require more attention. For this purpose, the period of breakdown of the centralized management system in 1991–1997, when there was no international agreement, is defined as counterfactual performance—that is, *CP(t).* The assumption is that, had the 1998 agreement not been reached, the riparian countries would have continued to behave as they had in 1991–1997. Another approach to measuring *CP* could be to assume unconstrained maximization by Kyrgyzstan of hydropower production to cover domestic energy needs and export excess energy to obtain foreign currency. However, experts on the region suggest that such a scenario would have been very unlikely.

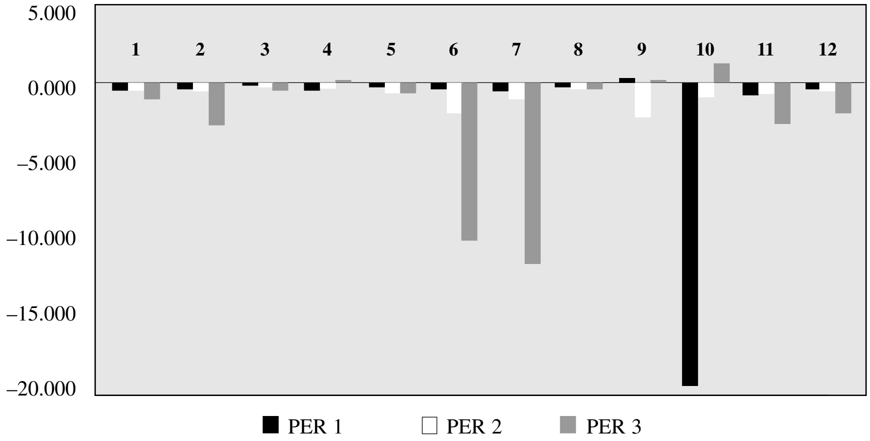
As to the definition of the optimal performance *OP,* there are three options: the natural runoff regime; runoff under Soviet rule; and estimates from optimization models. The natural runoff, $OP_N(t)$, is arguably the most problematic measure of the three, since it is quite difficult to see why no regulation of river flow should be Pareto-improving on a properly operated reservoir (in this case the Toktogul reservoir). However, one can use this measure of *OP* for purposes of comparison. The second measure for *OP* assumes that centralized management in Soviet times (Period 1, 1980–1990)[25] was optimal ($OP_S(t)$), because upstream and downstream interests were addressed through an integrated water-energy exchange system. Interviews with experts on the region confirm that the exchanges of water and energy under the Soviet management system worked relatively well, in terms of both providing water for irrigation downstream and facilitating energy production upstream. The disadvantage of the second measure of *OP* is that, from the perspective of the long-term Aral Sea problem and local economic and environmental interests there, Period 1 was certainly not optimal.[26] We thus also invoke a third notion of optimality, $OP_C(t)$, which emphasizes sustainability of natural resources management at the basin scale. This level, μ(optim), is not observed but is the result of a simulation-optimization approach that is denoted as $OP_C(t)$ (see Appendix). This simulation was undertaken by D. McKinney, X. M. Cai, and L. S Lasdon.[27] It considers risk minimization in water supply, environmental conservation of soil and water resources, spatial and temporal equity in water allocation, and economic efficiency in the development of future water infrastructure. The full optimization scenario operates under the assumption of long-term average precipitation in the basin and determines monthly reservoir releases, infrastructure development, and irrigated crop patterns and area, with the objective to maximize the resulting sum of irrigation and ecological benefits and hydropower profits.

We start with a simple estimation of *PER* based on equation (1). The results are shown in Figure 5. For all three measures of *OP* and for almost all months of the year, performance is very low.

We now move to the more complex approach and use the following notation to distinguish the scaling of $PER^*(t)$. $PER^*(t, OP_S)$ is calculated with respect to $OP_S(t)$ and $PER^*(t, OP_C)$ with respect to $OP_C(t)$. We restrict the estimation to $OP_S(t)$ and $OP_C(t)$, since $OP_N(t)$ is a rather problematic measure for optimum performance.

To compute the performance $PER^*(t)$ of the international management system installed in 1998, we use monthly averaged flow values for $OP_S(t)$ and $CP(t)$ (see Table 2). This is necessary for two reasons. First, comparing individual hydrological years with differing resource endowments (i.e.,

Figure 5 Simple Assessment of Performance



Notes: PER1 is based on $OP_N(t)$, PER2 on $OP_S(t)$, and PER3 on $OP_S(t)$. The calculation is based on long-term averages for each month; see the chapter appendix.

Table 2 Mean Monthly Flows for Period 1 (1980–1990) and Period 2 (1991–1997)

Month	1	2	3	4	5	6	7	8	9	10	11	12
μ ($OP_s(t)$) [m^3/s]	217	236	216	282	500	594	749	578	204	166	180	221
μ ($CP_s(t)$) [m^3/s]	479	464	429	350	348	450	481	354	199	234	344	480

Source: Andrey Yakovlev, Uzbek Hydrometeorological Service, Tashkent, Uzbekistan.

Note: As to $\mu(OP_S(t))$, we do not take into account the initial years of reservoir filling (1974–1979); $\mu(OP_C(t))$ is shown in the chapter appendix.

inflow as well as reservoir levels) and demand (for electricity as well as irrigation water) is problematic. Doing so would lead to an arbitrary comparison of reservoir outflows between years that are not necessarily comparable with respect to the key hydrological variables. Second, the individual periods have different lengths and so cannot be compared directly.

The calculation of $PER^*(t)$ based on $OP_S(t)$ may be problematic. The underlying assumption is that demand for irrigation water and hydroelectric power did not change during the period 1980–2006. Yet this assumption could be invalidated by the fact that, for example, the irrigated area in Uzbekistan grew by more than a quarter, from 3.5 million hectares in 1980 to 4.4 million hectares in 1998.[28] Demand in Kyrgyzstan for hydroelectric power has grown substantially as well.[29] This problem could have been addressed by scaling $\mu(OP_S(t))$ according to changes in demand for irrigation water and hydroelectric power, but such an approach would entail a high degree of arbitrariness. In particular, the very notion of optimality may lose sense after such scaling, since it does not take account of interseasonal shifts of optimal water allocation. In other words, optimal allocation is not a linear function of the quantity of water available. Note that such a problem does not apply to $PER^*(t, OP_C)$, because the relevant measure of OP reflects recent up- and downstream demand constraints.

The temporal development of $PER^*(t, OP_S)$ and $PER^*(t, OP_C)$ is shown in Figure 6. With respect to both notions of optimality, performance of the 1998 regime has been poor. The figure shows that extremely negative values of $PER^*(t, OP_S)$ start to occur from 2002 onward, usually in September. This can be explained by the fact that in this month, $|\mu(CP(t))-\mu(OP_S(t))|$, i.e., the denominator of $PER^*(t, OP_S)$ is small and the difference between actual performance and the monthly averaged performance of Period 1, i.e. $|AP(t)-\mu(OP_S(t))|$, is large.

Table 3 shows the overall results of our performance estimation. (We provide more details on the technicalities of the calculation in a previous paper.[30]) The calculations in Table 3 confirm the visual impression from Figure 6 that the performance of the international management regime for the Naryn/Syr Darya is very low.

Figure 6 ***PER*(t)*** **During Period 3 with Respect to the Two Notions of Optimality**

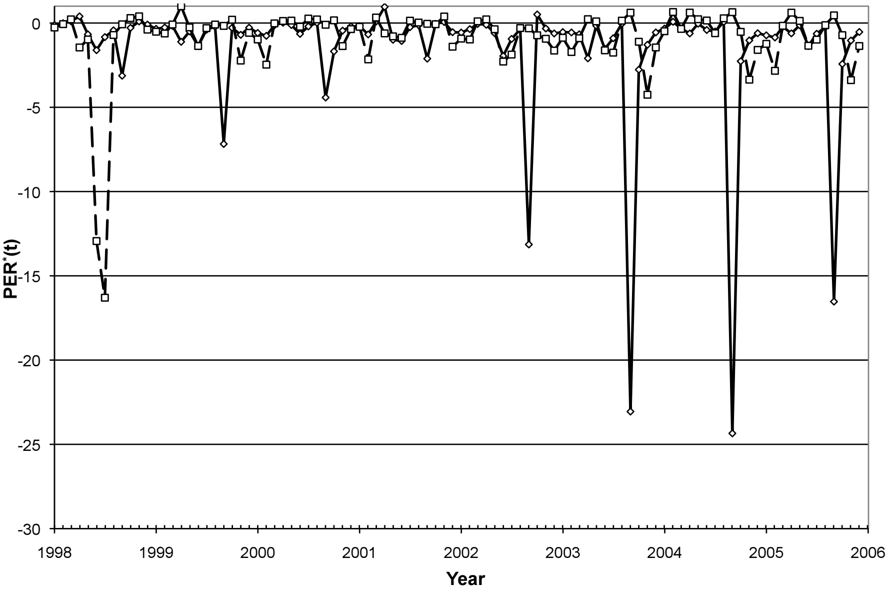


Table 3 Average Regime Performance and Variance with Reference to OP_S and OP_C

	Variance of *PER**	Average of *PER**
OP_S	–0.24	0.63
OP_C	–0.71	0.92

Note: The calculations are based on equations (4) and (5) and Table 3.

Conclusion

The theoretical literature stipulates that international environmental cooperation in upstream-downstream settings is very difficult.[31] However, the empirical evidence for this claim remains controversial. Many qualitative case studies and some quantitative research show that upstream-downstream water cooperation occurs quite frequently. For example, Brochmann and Gleditsch find no significant negative effects of upstream-downstream settings on the likelihood of cooperation.[32] Such empirical findings are indeed surprising, because they suggest that upstream-downstream asymmetries can be overcome through compensation payments and issue linkages offered by downstream countries in exchange for concessions by upstream countries at reasonably low transaction costs.[33] However, such findings may be overly optimistic, because they rely on definitions of the dependent variable (cooperation) that do not really capture the substance or depth of cooperation. As Downs, Roche, and Barsoon and others have argued, international cooperation, as measured by the existence of treaties and compliance with international commitments, may often be more shallow than it first appears.[34]

In the first part of this article, we presented a measurement concept that seeks to capture the depth or substance of cooperation. Assessing substantively the performance of international water governance efforts is important from an academic and practical viewpoint. Developing and testing generalizable explanations for success and failure in international water governance must rely on an accurate measurement of the dependent variable (i.e., success/failure). Moreover, helping policymakers understand whether or not a given water governance system performs well is usually the first step toward improving policies and institutions.

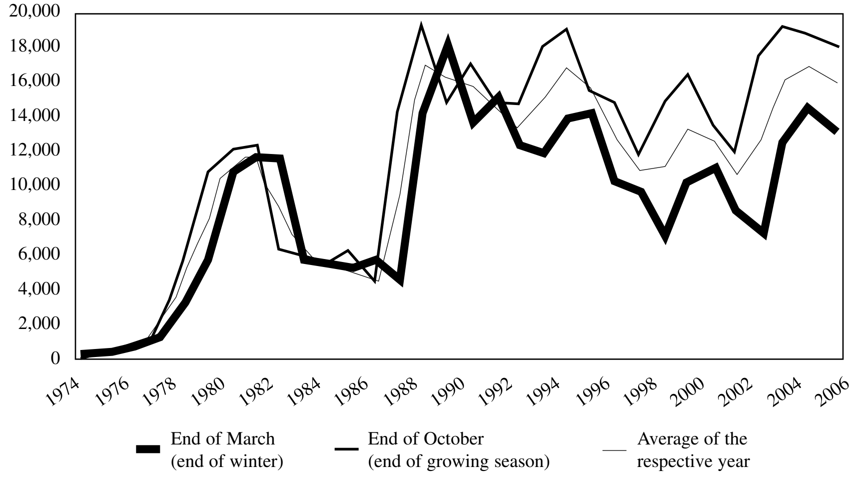
To demonstrate its empirical relevance, we have applied the performance metric to the Naryn/Syr Darya basin. We have seen that an international agreement was concluded in 1998, only seven years after the collapse of the USSR and thus comparatively very fast. Moreover, we have noted that compliance with this agreement has been quite impressive. Yet when the PER metric was applied, the initially positive picture changed entirely. Performance over time of the governance system established in 1998 has been very low and highly variable.

The principal policy implication of these findings is that the governance system in place for the Naryn/Syr Darya is in urgent need of reform. Conflicts over water allocation among the riparian countries have in the past few years been muted by high levels of precipitation upstream. The compliance analysis we have offered, together with Figure 7, show that because of high precipitation and thus high inflows into the Toktogul reservoir in recent years, excessive water releases from the reservoir in winter have not come at the cost of lower releases in spring to autumn. However, as soon as an extended period of low precipitation sets in, such as in the hydrological year 2008, seasonal trade-offs will become manifest again and conflict is likely to heat up very quickly.

Systematic analysis of the reasons for poor performance of the existing governance system for the Naryn/Syr Darya is beyond the scope of this article. However, the most apparent reasons relate to economic crisis and domestic political instability in the riparian countries. These problems have made it hard to establish credible long-term commitments. Most notably, as long as Kyrgyzstan does not receive credible commitments from the downstream countries that greater water releases from the Toktogul dam in spring to autumn (but not in winter) will be followed by more energy deliveries by downstream countries in winter, the incentive for Kyrgyzstan to release larger amounts of water for energy production in winter and lower amounts in summer will dominate.

Engineering solutions to the problem have been proposed and, to a minor degree, already undertaken. Uzbekistan has built several small reservoirs on

Figure 7 Water Volume of the Toktogul Reservoir, 1974–2006 (millions of cubic meters)



its territory to retain excessive water releases from Kyrgyzstan in winter for irrigation use in spring to autumn. However, there are topographical limits to this solution. Another solution would be to reactivate old plans to build a new reservoir upstream of the Toktogul reservoir. Releases from this new reservoir could serve to produce electricity for Kyrgyzstan in winter, and the released water could be retained further downstream in the Toktogul reservoir for release in spring to autumn for irrigation in the downstream countries. Such a solution might work if foreign investors could be attracted to this project, but this prospect is unlikely for the time being.

However, it is quite obvious that water-energy exchanges among the three riparian countries would be more cost-efficient than any unilateral allocation measure. International efforts should thus focus on establishing long-term hydrological forecasting systems for the Naryn/Syr Darya basin. A revised water-energy exchange mechanism that builds on such forecasting should include multiyear targets for the management of the Toktogul reservoir. To solve the time inconsistency problem in this upstream-downstream exchange, guarantees by advanced industrialized countries or international organizations could be established.

The more general message from this article is that many international upstream-downstream water governance systems may involve more shallow or unstable cooperation than is evident at first glance. Indeed, a closer look at a range of prominent cases—for example, the Rhine, Danube, and Colorado Rivers—suggests that solving upstream-downstream problems often takes decades. It usually goes hand in hand with growing income and intensifying political and economic ties among riparian countries. In the case of the Rhine, for instance, it took half a century to set up a system of international funding for pollution reduction to deal with salinization problems. Indeed, by the time this exchange was established, the problem had already been largely solved independently of international cooperation, as the main sources of salinity—coal and potash mines—were closed for economic reasons. Similarly, efforts to clean up the Danube and other rivers in Europe and North America have developed together with growing income, trade interdependence, and democratization.

A quite common pattern, at least with respect to water quality issues, seems to be that domestic public demand for stricter environmental policies grows with income and political and civil liberties. Stricter domestic standards (and thus lower pollution) tend to foster international cooperation in this area as well. For instance, when country A adopts higher water quality standards and commissions water treatment plants, these measures usually apply to areas near the national border, too. As a result, country A's water flowing into neighboring country B is bound to be of higher quality. To the extent that such domestic processes develop in parallel in two or more riparian countries, this will facilitate international cooperation on water quality issues. Further research should study interactions between policy processes at domestic and international levels in order to establish whether policy processes and policy outcomes in international water management are driven primarily by international cooperation or primarily by domestic processes that converge in more or less coordinated ways into higher international standards. 🌐

Appendix Means and Standard Deviations of Monthly Flows Under Different Management Systems

	Natural Runoff Regime		Period 1		Period 2		Period 3		Optim.
Month	μ	σ	μ	σ	μ	σ	μ	σ	μ
1	150.0	27.9	188.6	74.7	478.5	101.1	590.0	55.3	357.7
2	151.1	25.5	202.1	67.2	464.2	113.0	561.8	78.6	426.2
3	178.2	28.5	195.5	50.4	428.9	122.1	465.8	52.9	323.4
4	314.7	94.3	271.9	94.2	350.2	115.6	367.0	79.6	426.2
5	661.4	200.8	457.8	186.4	348.0	120.2	286.8	52.0	452.8
6	969.3	342.2	550.9	196.3	450.1	152.6	270.6	73.8	468.0
7	797.6	264.1	654.8	205.5	481.0	174.5	324.3	78.2	494.7
8	516.9	137.0	521.7	153.9	354.1	79.5	316.6	40.3	490.9
9	287.1	71.7	184.1	99.8	198.5	89.2	228.1	93.0	441.4
10	230.4	48.8	142.6	73.0	234.5	67.7	313.7	86.8	300.6
11	217.0	45.7	144.3	92.1	343.5	51.9	439.4	84.9	304.4
12	174.1	30.1	188.5	79.2	479.7	82.3	590.6	53.0	418.6
Overall	388	307	311	215	384	139	396	141	409

Sources: The last column shows data from X. M. Cai et al., “Integrated Hydrologic-Agronomic-Economic Model for River Basin Management,” *Journal of Water Resources Planning and Management* 129 (2003): 4–17. Data in other columns derived from Andrey Yakovlev, Uzbek Hydrometeorological Service, Tashkent, Uzbekistan.

Notes: The bottom row displays overall means and standard deviations for the duration of the management periods. Units are m^3/s for μ.

Notes

Thomas Bernauer is professor of political science at ETH Zurich. He and his research group are based at the Center for Comparative and International Studies (CIS) and the Institute for Environmental Decisions (IED). His research concentrates on the interaction of politics and markets, particularly in the environmental realm, and the conditions under which effective international environmental cooperation is possible (see www.bernauer.ethz.ch). Tobias Siegfried is adjunct assistant professor at the School of International and Public Affairs at Columbia University and a fellow at the Earth Institute. He investigates problems of freshwater depletion and degradation in the context of demographic and economic development. His work focuses on regions where sustainable resources management is difficult due to inadequate institutions and political conflict.

The authors are grateful to Andrey Yakovlev, Uzbek Hydrometeorological Service, and the Global Runoff Data Center, Federal Institute of Hydrology (BfG), Koblenz, Germany, for providing the hydrological data. Thanks also go to Christoph

Schär, Jean Fried, Detlef Sprinz, Gabi Ruoff, Vally Koubi, Oran Young, Gary Goertz, Aaron Wolf, Thomas Hinkebein, Nils Petter Gleditsch, Peter Bauer-Gottwein, Bianca Sarbu, and the reviewers of *Global Governance* for valuable comments on earlier versions of this article.

1. A. Dinar and S. Dinar, "Recent Developments in the Literature on Conflict and Cooperation in International Shared Water," *Natural Resources Journal* 43, no. 4 (2003): 1217–1287; T. Bernauer and A. Kalbhenn, "The Politics of International Freshwater Resources," manuscript, ETH Zurich, 2008, available at http://www.ib.ethz.ch1.

2. F. Marty, *The Management of International Rivers: Problems, Politics and Institutions* (Frankfurt: Peter Lang, 2001); T. Bernauer, "Protecting the Rhine River Against Chloride Pollution," in R. O. Keohane and M. A. Levy, eds., *Institutions for Environmental Aid: Pitfalls and Promise* (Cambridge: MIT Press, 1996), pp. 201–233.

3. See, for example, R. Durth, *GrenzüberschreitendeUmweltprobleme und regionale Integration: Zur Politischen Oekonomie von Oberlauf-Unterlauf-Problemen an internationalen Flüssen* (Baden-Baden: NomosVerlag, 1996); M. Brochmann and N. P. Gleditsch, "Conflict, Cooperation, and Good Governance in International River Basins," paper prepared for the workshop "Governance and the Global Water System: Institutions, Actors, Scales of Water Governance—Facing the Challenges of Global Change," Bonn, Germany, 20–23 June 2006; I. Dombrowsky, *Conflict, Cooperation, and Institutions in International Water Management* (PhD diss., Martin-Luther-Universität Halle-Wittenberg, Germany, 2005); J. Tir and J. T. Ackerman, "To Share or Not to Share: Politics of Cooperation Between Riparian States," paper presented at the annual meeting of the International Studies Association, Montreal, 17–20 March, 2004; A. T. Wolf, A. Kramer, A. Carius, and G. D. Dabelko, "Managing Water Conflict and Cooperation," in *State of the World 2005: Redefining Global Security* (Washington, DC: Worldwatch Institute, 2005), chap. 5; and S. B. Yoffe et al., "Geography of International Water Conflict and Cooperation: Data Sets and Applications," *Water Resources Research* 40, no. 5 (2004): 1–12.

4. See T. Bernauer, "Explaining Success and Failure in International River Management," *Aquatic Sciences* 64, no. 1 (2002): 1–19; and T. Bernauer, "The Effect of International Environmental Institutions: How We Might Learn More," *International Organization* 49, no. 2 (1995): 351–377.

5. R. B. Mitchell and St. B. Rothman, "Creating Large-N Datasets from Qualitative Information: Lessons from International Environmental Agreements," paper prepared for the annual meeting of the American Political Science Association, Philadelphia, 1–2 September 2006; Dombrowsky, *Conflict, Cooperation, and Institutions in International Water Management;* K. Conca, F. Wu, and C. Mei, "Global Regime Formation or Complex Institution Building? Principled Content of International River Agreements," *International Studies Quarterly* 50 (2006): 263–285.

6. See, for example, A. T. Wolf, A. Kramer, A. Carius, and G. D. Dabelko, "Managing Water Conflict and Cooperation," in *State of the World 2005: Redefining Global Security* (Washington, DC: Worldwatch Institute, 2005), chap. 5; and Brochmann and Gleditsch, "Conflict, Cooperation, and Good Governance in International River Basins."

7. H. Sigman, 2004, "Does Trade Promote Environmental Coordination? Pollution in International Rivers," *Journal of Economic Analysis and Policy (Contributions)* 3, no. 2 (2004), available at www.bepress.com/bejeap/contributions/vol3/iss2/art2 (accessed 20 April 2008); T. Bernauer and P. Kuhn, "Is There an Environmental Version of the Kantian Peace? Insights from Water Pollution in Europe," *European Journal of International Relations,* forthcoming 2008.

8. C. Helm and D. Sprinz, "Measuring the Effectiveness of International Environmental Regimes," *Journal of Conflict Resolution* 45, no. 5 (2000): 630–652; J. Hovi, D. Sprinz, and A. Underdal, "Regime Effectiveness and the Oslo-Potsdam Solution: A Rejoinder to Oran Young," *Global Environmental Politics* 3, no. 3 (2003): 105–107; J. Hovi, D. Sprinz, and A. Underdal, "The Oslo-Potsdam Solution to Measuring Regime Effectiveness: Critique, Response, and the Road Ahead," *Global Environmental Politics* 3, no. 3 (2003): 74–96; A. Underdal, "The Concept of Regime 'Effectiveness,'" *Cooperation and Conflict* 27, no. 3 (1992): 227–240.

9. R. B. Mitchell and P. M. Kielbach, "Situation Structure and Institutional Design: Reciprocity, Coercion, and Exchange," *International Organization* 55, no. 4 (2001): 891–917.

10. A. T. Wolf, A. Kramer, A. Carius, and G. D. Dabelko, "Managing Water Conflict and Cooperation," in *State of the World 2005: Redefining Global Security* (Washington, DC: Worldwatch Institute, 2005), chap. 5; Marty, *The Management of International Rivers.*

11. Brochmann and Gleditsch, *Conflict, Cooperation, and Good Governance in International River Basins.*

12. Tir and Ackerman, *To Share or Not to Share.*

13. Bernauer, "Protecting the Rhine River Against Chloride Pollution," pp. 201–233. See also T. Bernauer and P. Moser, "Reducing Pollution of the River Rhine: The Influence of International Cooperation," *Journal of Environment and Development* 5, no. 4 (1996): 391–417.

14. G. W. Downs, D. M. Rocke, and P. N. Barsoom, 1996, "Is the Good News About Compliance Good News About Cooperation?" *International Organization* 50, no. 3 (1996): 379–406.

15. Helm and Sprinz, "Measuring the Effectiveness of International Environmental Regimes," pp. 630–652. See also Underdal, "The Concept of Regime 'Effectiveness.'" pp. 227–240.

16. T. Siegfried and T. Bernauer, "Estimating the Performance of International Regulatory Regimes: Methodology and Empirical Application to International Water Management in the Naryn/Syr Darya Basin," *Water Resources Research* 43 (2007), W11406, doi: 10.1029/2006WR005738.

17. V. Dukhovny and V. Sokolov, "Integrated Water Resources Management, Experience, and Lessons Learned form Central Asia—Towards the Fourth World Water Forum," Inter-State Commission for Water Coordination in the Aral Sea Basin, Tashkent, 2005.

18. E. Antipova et al., "Optimization of Syr Darya Water and Energy Uses," *Water International* 27, no. 4 (2002): 504–516.

19. X. M. Cai, D. C. McKinney, and L. S. Lasdon, "A Framework for Sustainability Analysis in Water Resources Management and Application to the Syr Darya Basin," *Water Resources Research* 38, no. 6 (2002): 21/1–21/14.

20. O. Savoskul et al., "Water, Climate, Food, and Environment in the Syr Darya Basin," manuscript, 2003, available at www.weap21.org/downloads/ (assessed 18 April 2008).

21. See http://ocid.nacse.org/cgi-bin/qml/tfdd/treaties.qml (accessed 18 April 2008).

22. Ibid.

23. Bernauer, "Protecting the Rhine River Against Chloride Pollution," pp. 201–233; Bernauer and Moser, "Reducing Pollution of the River Rhine," pp. 391–417; Marty, *The Management of International Rivers.*

24. Downs, Rocke, and Barsoom, "Is the Good News About Compliance Good News About Cooperation?" pp. 379–406.

25. We start with 1980 because 1974–1979 were years of reservoir filling. We expect a strong trend effect in reservoir outflows during the latter years; this would introduce errors in our performance estimation.

26. O. Young, "Inferences and Indices: Evaluating the Effectiveness of International Environmental Regimes," *Global Environmental Politics* 1, no. 1 (2001): 99–121; O. Young, "Determining Regime Effectiveness: A Commentary on the Oslo-Potsdam Solution," *Global Environmental Politics* 3, no. 3 (2003): 97–104.

27. D. McKinney, X. M. Cai, and L. S. Lasdon, *Integrated Water Resources Management Model for the Syr Darya Basin,* Technical Report, US Agency for International Development, Environmental Policies and Institutions for Central Asia (EPIC) Program, Almaty, Kazakhstan, 1999.

28. See www.fao.org/nr/water/aquastat/countries/uzbekistan/index.stm (accessed 18 April 2008).

29. A similar argument applies to $\mu(CP(t))$.

30. Siegfried and Bernauer, "Estimating the Performance of International Regulatory Regimes."

31. Mitchell and Kielbach, "Situation Structure and Institutional Design"; Bernauer, "The Effect of International Environmental Institutions"; Bernauer, "Protecting the Rhine River Against Chloride Pollution."

32. Brochmann and Gleditsch, *Conflict, Cooperation, and Good Governance in International River Basins.*

33. S. Dinar, "Assessing Side-Payments and Cost-Sharing Patterns in International Water Agreements: The Geographic and Economic Connection," *Political Geography* 25 (2006): 412–437.

34. Downs, Rocke, and Barsoom, "Is the Good News About Compliance Good News About Cooperation?"

Electronic Sources

FAO Information System on Water and Agriculture: www.fao.org/nr/water/aquastat/main/index.stm (accessed 18 April 2008).

Global Runoff Data Center (GRDC): http://grdc.bafg.de/servlet/is/Entry.987.Display/ (accessed 18 April 2008).

Transboundary Freshwater Dispute Database, International Treaties: http://ocid.nacse.org/cgi-bin/qml/tfdd/treaties.qml (accessed 18 April 2008).

Uzbek Hydrometeorological Service (UZHYDROMET): www.worldweather.org/032/m032.htm (accessed 18 April 2008).