Article

Journal of Conflict Resolution
2019, Vol. 63(8) 1860-1888

DOI: 10.1177/0022002718822127
journals.sagepub.com/home/jcr
SAGE

# Do Binding Beat Nonbinding Agreements? Regulating International Water Quality

**Martin Köppel[1], and Detlef F. Sprinz[1]**

**Abstract**

Does the form of a multilateral agreement (MEA) by itself improve environmental performance? In particular, do legally nonbinding MEAs pose a rival to the effect of more traditional legally binding international agreements? Our theory builds on the legal and international regimes literatures and postulates that legally binding agreements (LBAs) have more benign effects on water quality than legally nonbinding agreements (LNBAs). We probe two operationalizations of the form of MEAs. First, we purely focus on the form: of legally binding versus legally nonbindings. Second, we combine the form of an agreement each with an index of precision and an index of delegation. The empirical focus is on upstream–downstream water quality in Europe during 1990 to 2007. Our regression analyses, regardless of specification, find that LBAs beat LNBAs and that LBAs with high degrees of precision and delegation beat the effect of any other configuration of agreements with respect to enhancing water quality.

**Keywords**

international regimes, international institutions, international law, international treaties

[1]Research Domain Transdisciplinary Concepts and Methods, Potsdam Institute for Climate Impact Research, Potsdam, Germany

**Corresponding Author:**

Detlef F. Sprinz, Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany.

Email: dsp@pik-potsdam.de

## Introduction

Over the past decades, research on international regimes had a marked impact on the study of world politics. In particular, the reasons for their creation (Roberts, Parks, and Vásquez 2004; Stinnett and Tir 2009; Sprinz and Vaahtoranta 1994), the degree of behavioral change (Goldstein, Rivers, and Tomz 2007; Von Stein 2005), and the degree of their effects (Köppel 2009a; Breitmeier, Young, and Zürn 2006; Miles et al. 2002) have been debated. On the impact or effect of international regimes, the conclusion emerged that they do matter—yet are far from perfect in solving the problems that led to their creation (Miles et al. 2002; Young 2011). Comparatively little attention has so far been placed on the question whether the *form* of (environmental) agreement matters and, in particular, whether legally binding agreements (LBAs) or legally nonbinding agreements (LNBAs) have greater effect—and which their directional impact is.[1] This article provides empirical evidence whether LBAs as compared to LNBAs lead to improvements in the quality of our environment.

The existing literature on the effectiveness of LBAs and LNBAs offers contradictory empirical evidence. Breitmeier, Young, and Zürn (2006) show that legalization of international institutions raises their effectiveness. Yet case studies by Gurtner-Zimmermann (1998), Holtrup (1999), and Tschanz (2001) indicate that soft law can be at least as effective as binding treaties. In addition, Victor, Raustiala, and Skolnikoff (1998) argued that soft law has proven more effective in changing the behavior of relevant actors than legally binding treaties. Breitmeier, Underdal, and Young (2011, 12) showed that when rules are legally binding, compliance is, at best, only marginally higher than is the case for LNBAs, yet LNBAs outperform LBAs in terms of problem-solving. Moreover, Skjærseth, Stokke, and Wettestad (2006) assessed the effects of hard and soft law on international nutrient agreements in Norwegian agriculture. Skjærseth found that adding legally binding to LNBAs in domestic law does not necessarily lead to improved implementation of international agreements. Thus, the effect of LBAs as compared to LNBAs has so far remained unresolved.

In order to shed light on whether LBAs versus LNBAs have the greater effect on environmental problem-solving, we chose the field of transboundary river quality as a test case as it embraces a long time span, offers substantial variation in the dependent variable of river water quality across time and space, allows us to keep the problem structure constant, is politically relevant to broader sections of the population (human consumption, biodiversity, fishing), and affords a reasonable balance of cases that have not been regulated with those with different combinations of LBAs and LNBAs.

This article makes three major contributions. First, we employ multiple operationalizations of LBAs versus LNBAs.[2] The initial operationalization purely focuses on the form of legally binding versus legally nonbinding multilateral environmental agreements (MEAs). Subsequently, we refine this approach by combining the form of an agreement, each with an index of precision (combining depth and clarity of obligations) and an index of delegation (combining a scale on monitoring provisions with a scale on dispute resolution mechanisms).

Second, we contribute to the literature on the effectiveness of transboundary river pollution. In general, fresh water is widely seen as humanity's most precious but also most vulnerable natural resource that is directly threatened by anthropogenic activities (Bernauer and Kalbhenn 2010; Vörösmarty et al. 2010).

Third, the public policy literature proposes a distinction between three potential indicators: *outputs* refer to the norms, principles, and rules that states adopt when implementing a regime. *Outcomes* can be thought of as the regime-induced changes in human behavior. And *impacts* in this context can be interpreted as changes in environmental quality—the biophysical environment itself (Underdal 2002, 5-6; Easton 1965; Mitchell 2007, 896-897; Young 2004, 12-13). While studies on the effectiveness of international institutions have focused mainly on the output and outcome dimensions (e.g., Bernauer et al. 2013; Spilker and Koubi 2016; Bernauer, Böhmelt, and Koubi 2013), we focus on the *impact* dimension. Our analysis uses the border-measurement-station-location year (short: "river dyad year") as the unit of analysis. This allows us to look at the water quality of rivers along river basins rather than at the river basin level at large. Interventions undertaken along a river basin will be attributed to specific locations.

## Theory and Hypotheses

States that decide to solve transboundary environmental problems through cooperation have several options at their disposal (Lipson 1991; Guzman 2005). International agreements have proven to be one important mechanism for solving transboundary problems.

In general, the literature distinguishes between two different types of international agreements: LBAs and LNBAs. On the one hand, legally binding treaties often appear to be the preferred approach for solving international environmental problems. At least some LBAs have proven ineffective if not counterproductive for solving environmental problems. For instance, the Kyoto Protocol has received fundamental critique for being the "wrong tool for the nature of the job" (Prins and Rayner 2007, 973). Furthermore, states increasingly use soft law (Skjærseth, Stokke, and Wettestad 2006, 104; Thürer 2000, 453).

Scholars offer a variety of reasons why agreements can lead to changes in norms, principles, and rules or in state behavior—that is, the output and outcome dimensions of regime effectiveness. However, there largely remains a gap in explaining the impact dimension. International agreements do not exercise direct control over the environmental conditions they are expected to improve. Only by way of national implementation of international law is international environmental law able to tackle the environmental challenge (Bodansky 2010, 131). A state's behavioral change constitutes a necessary condition for changes in environmental quality.

We restrict the analysis to upstream–downstream settings because this allows us to control for problem structure. In addition, upstream–downstream problems are widely seen as the most difficult challenge to solve in international environmental politics (Mitchell and Keilbach 2001). If LBAs beat LNBAs under such unfavorable circumstances, their effect may be more pronounced in more benign settings.

Does the form of an agreement matter? States employ international law with care (Raustiala 2005, 587; Von Stein 2008). We propose several hypotheses for the effects of international agreements on environmental quality by distinguishing between the different effects of LBAs versus legally nonbinding MEAs. We proceed in five steps. First, we build the argument for LBAs being effective. Second, we focus purely on the legal form of an international agreement: legally binding versus legally nonbinding (LBA vs. LNBA). Third, we use a more nuanced approach by combining the form of agreement with indices of precision and delegation (LBA+ vs. LBA−). Finally, we contrast the effects of LBA+ with LNBA− and of LBA+ with LNBA+.

## *Why LBAs Are Effective*

LBAs may influence environmental quality through changing the behavior of actors via several causal mechanisms.[3] First, LBAs can modify the cost and benefit calculus of actors (Young and Levy 1999). Actors change their behavior because agreements reward compliance and punish noncompliance. For example, countries install new wastewater treatment plants for achieving the goals of reducing water pollution. This then leads to water quality improvements that would not have materialized otherwise. Information and transparency can play an important role in this regard.

Second, LBAs can enhance cooperation among actors (Young and Levy 1999, 23). Since the problem structure of many international environmental issues hampers joint and effective solutions, agreements can help alleviate these problems; for example, the upstream country in an upstream–downstream configuration has low incentives compared to the downstream country, to avoid polluting a river. LBAs can enhance transparency of the actor's behavior and mitigate fears of cheating. As Mitchell (2010, 43) noted, transparency makes the impacts of polluting behavior visible and clearly links them to actors. Jointly collecting information on the sources and the magnitude of water pollution enables trust among riparian states (see Salamé and Van der Zaag 2010, 172). Highly detailed and specialized data, comprising the magnitude as well as the sources of pollution, are necessary for evaluating whether riparian countries change their polluting behavior after they have committed to an agreement (Tir and Stinnett 2012, 216). Transparency may lead municipalities and industries to alter their behavior because they expect other actors to reciprocate.

Third, LBAs can operate as learning facilitators (Young and Levy 1999, 24-25), such as individual and social learning (Social Learning Group 2001a, 2001b). Learning can relate to the identification of new measures for effectively solving problems. Implementing these new measures may lead to water quality improvements that would not have happened otherwise. Learning can lead municipalities and industries to implement new technologies capable of reducing or eliminating water pollution or install new wastewater treatment plants. This change in behavior subsequently leads to an improvement of water quality that would not have occurred otherwise.

Fourth, LBAs can function as agents of internal realignment (Young and Levy 1999, 26-28). The establishment of an international agreement affects the behavior of actors by generating new constituencies or by changing the balance between national actors. LBAs can increase governmental concern or can be used by governmental as well as societal actors as commitment devices. They legitimize certain behaviors and delegitimize others (Simmons 2010, 192). The existence of an international rule can be used by domestic actors either to justify their own actions or to call the legitimacy of other actors into question (Conca 2006, 11). Governmental officials who pursue environmental protection, for instance, can make use of a recently signed agreement to justify new expenditures to improve water quality.

Finally, LBAs can act as bestowers of authority. Rather than calculating what the costs and benefits of available options are, actors respect authority. Treaties affect domestic policy through the integration of obligations into domestic law. Enforceable domestic laws make breaches of the law costly—which many polluters wish to avoid. At its simplest, an international agreement is transformed into domestic law. Polluters who do not want to bear the consequences of breaching the new national law change their behavior. For instance, they may switch to cleaner production technologies (Bernauer and Moser 1996, 397-398). Consequently, this leads to water quality improvements due to the international agreement.

In view of the aforementioned points, we hypothesize:

> **Hypothesis 1:** Signing a LBA leads to higher water quality in transboundary rivers than not signing a LBA.

## *Why LBAs Are More Effective Than LNBAs*

Are LBAs more effective than LNBAs? Theoretically, there are some aspects pointing in this direction. However, as Köppel (2009b, 2012) has shown elsewhere, there are also arguments in favor for LNBAs being more effective than LBAs (see note 3).

First, LBAs strengthen a commitment's credibility by constraining "self-serving autointerpretation" (Abbott and Snidal 2000, 427). In addition, it increases the costs of reneging. For many countries, signing an LBA at the international level requires domestic, legislative approval. This approval makes an LBA more "resistant to backsliding" compared to an LNBA (Bodansky 2010, 179), or put succinctly by Lipson (1991, 508-509), LBAs "visibly stake the parties' reputations to their pledges." In a nutshell, states implement the obligations of an agreement due to its legally binding form and enforceability in the courts.

Second, LBAs increase the probability of compliance (Köppel 2009b). Stronger legalization raises the political costs of noncompliance. As Bodansky (2010, 179) argued, "[t]here is a difference between saying that one will *try* to do something and *committing* to doing it, particularly when the commitment is made at the highest levels of government." Pacta sunt servanda displays the respect of a government for the law itself (Simmons 2010, 277). Hence, when a state violates an international commitment, it may face several consequences ranging from loss of reputation as a reliable partner to specific and costly retaliation. In some cases, noncompliance may also be combined with some form of direct sanction (Lipson 1991, 508-12; Guzman 2005, 582).

While LBAs can affect environmental quality directly along all of the aforementioned causal mechanisms, LNBAs cannot. In general, an LNBA at the international level leads to weaker forms of domestic oversight as compared to LBAs. As Brown Weiss (2007, 550) argued, "[t]he most significant difference [between LBAs and LNBAs] is that some of the pathways for enforcing compliance at the national level, namely through courts, may not be available for nonbinding legal instruments." Consequently, internal coordination, parliamentary approval, and the obligation to publish are weaker for LNBAs than is the case of LBAs. We therefore hypothesize:

> **Hypothesis 2:** Signing a LBA leads to better water quality in transboundary rivers than signing a LNBA.

## *Why LBA+s Are More Effective Than LBA−s*

In particular, Abbott et al. (2000) have argued that besides the form of an agreement, two other dimensions are important: (i) *precision*, that is, the depth of obligations as well as the precision of the agreement and (ii) *delegation*, that is, the existence of monitoring and enforcement provisions as well as the inclusion of a dispute resolution mechanism.

Are precise LBAs with delegation mechanisms (LBA+) more effective than imprecise LBAs with low delegation (LBA−)? There is reason to argue that precision and delegation are crucial for at least most of the above discussed causal mechanisms for LBAs.

On *precision*, Abbott et al. (2000, 412) argued, precise rules narrow the "scope for reasonable interpretation." Imprecise obligations give the signatory states the flexibility to interpret the obligations as they like. Precise rules, in contrast, specify unambiguously and clearly what and how states or other actors are expected to behave (Jacobson and Brown Weiss 1998, 524-525; Victor, Raustiala, and Skolnikoff 1998; Wettestad 1995, 43-44). This may, for instance, be important for the causal mechanism where agreements function as agents of internal realignment (see above).

To illustrate precision, a LBA may oblige signatory states to reduce their amount of water pollution by 40 percent until a specified date. Governmental and societal actors can use this obligation for legitimizing certain behaviors and delegitimizing others. New financial burdens can be justified because the signatory state precisely has committed to do so under the LBA.

In addition, precisely formulated agreements may lessen the probability of unintended violations due to ambiguities in agreements (Chayes and Chayes 1993). Ambiguity in treaty design may have detrimental effects as Fischhendler (2008) argued. Hence, precise rules provide states with clarity about expected behavior.

With respect to *delegation*, the existence of monitoring and enforcement provisions as well as the inclusion of a dispute resolution mechanism implies that authority is granted to a third party to implement, interpret, and to apply the rules as well as to resolve disputes (Abbott et al. 2000; Bernauer et al. 2013). Interpreting an agreement in a "self-serving or biased manner" (Von Stein 2008, 248) as well as to "accept or reject proposals without legal justification" (Abbott et al. 2000, 415) poses challenges for signatory states. When authority is granted to third parties, this means that states relinquish some of their sovereignty and authority. Commissions or secretariats take over numerous tasks such as the authority to monitor compliance, collect and publish information, or the power to enforce rules (Böhmelt and Pilster 2010, 248-249).

Monitoring, in particular, appears to take on a crucial role.[4] River commissions, established by an agreement, can help collect information regarding flow rates, pollution levels, and the sources of pollution. Monitoring increases transparency. As this information is collected under the mandate of a neutral international organization, it is perceived as more trustworthy if contrasted with data collected for purely national purposes and otherwise unaudited (Stinnett and Tir 2009, 232-233). This may, for instance, be crucial for causal mechanisms emphasizing the costs and benefit. Information, such as water quality data, is produced by the river commission due to the agreement.

Finally, the inclusion of a dispute mechanism can help stipulate peaceful settlements or prevent disputes escalating to armed conflict. Dispute resolution procedures can help to resolve interstate disagreements such as divergent perspectives on compliance. Also, it may help national governments to "neutralize domestic political pressure to violate an agreement by moving the matter to an international forum" (Stinnett and Tir 2009, 233-234).

We hypothesize:

> **Hypothesis 3:** Signing a LBA with high precision and delegation (LBA+) leads to better water quality in transboundary rivers than signing an LBA with low precision and delegation (LBA−).

## *Why LBA+s Are More Effective Than LNBA−s*

Besides the mechanisms mentioned above, there are further reasons why LBAs with high precision and delegation should be more effective than LNBAs with low precision and delegation.

LBAs strengthen a commitment's credibility by constraining "self-serving auto-interpretation" (Abbott and Snidal 2000, 427) and increase the costs of reneging. This seems even stronger when the agreement is precise and when it includes delegation provisions. An agreement's credibility seems to be higher when the rules unambiguously and clearly specify what and how states or other actors are expected to behave. Furthermore, the costs of reneging ought to be higher when authority is granted to third parties. In addition, LBAs increase the probability of compliance (Köppel 2009b). An LBA+ that contains precise obligations as well as major monitoring provisions is expected to be more effective than LNBAs with imprecise obligations and minor monitoring provisions.

We hypothesize:

> **Hypothesis 4:** Signing a LBA with high precision and delegation (LBA+) leads to better water quality in transboundary rivers than signing a LNBA with low precision and delegation (LNBA−).

### *Why LBA+s Are More Effective Than LNBA+s*

Further above, we have already argued that there are good arguments for LBAs being more effective than LNBA−. But are LBAs with high precision and delegation also more effective than LNBAs with high precision and delegation?

Theoretically, precision and delegation seem to be important, yet the configuration of high precision and delegation is kept constant in this comparison. Other things held constant, the legally binding status, as argued more extensively under Hypothesis 2, is expected to deliver additional impetus for reducing pollution.

We hypothesize:

> **Hypothesis 5:** Signing a LBA with high precision and delegation (LBA+) is more effective than signing a LNBA with high precision and delegation (LNBA+).

## Data Sources and Control Variables

Freshwater pollution and the pollution of transboundary rivers[5] are widely regarded as one of the most significant environmental problems worldwide (Bernauer and Kuhn 2010, 78). The statistical analysis is conducted at the river dyad year level.

First, we did not include other transboundary water issues such as navigation or water quantity because of the desire to keep the analysis amenable to systematic comparison. Second, we chose transboundary river basins because a focus on clear upstream–downstream problems allows us to control for problem structure. Managing transboundary rivers is among the most difficult types of conflict in international environmental politics as upstream countries may have incentives to free ride on downstream countries (Mitchell and Keilbach 2001). Third, the selection was restricted to Europe because the data offered by the European Environment Agency are likely to be of more homogenous quality as compared to global data sets. Finally, we used the river dyad year level as the unit of analysis to retain dyad-specific variation per river while using the basin level as the unit of analysis rests on the implausible assumption that all riparian countries pollute the basin to the same extent (Bernauer and Kalbhenn 2010, 7, 8). Furthermore, some agreements are specified at the level of the individual river rather than the overall river basin. A dyadic approach is used because the basic argument against a monadic approach is that cooperation is a relational phenomenon.[6]

By drawing on European river basins over the past decades, we are assured of comparable standards in the measurement of our dependent variable (water quality), good time-series data coverage, and control for unobserved cultural and other social factors that might explain cross-sectional variation.

## *Dependent Variable: Water Quality*

Several pollutants could potentially serve as indicators of water quality. We selected Biological Oxygen Demand ($BOD_5$) as the indicator for four reasons. First, $BOD_5$ is a commonly used measurement to determine water quality and is produced by human activity (Bernauer and Kuhn 2010). It is a measure of how much dissolved oxygen is consumed as microbes break down organic matter that usually takes place over five days at an incubation temperature of 20°C.[7] Flows from sewage discharges or industrial processes such as from paper production result in an increase in $BOD_5$. $BOD_5$ levels are easily measured by standard procedures that enhance data quality. Moreover, $BOD_5$ can travel far downstream, and background values as well as levels of natural variation are low (Meybeck, Chapman, and Helmer 1990, 17-20). Second, this type of pollution can be impacted by policies. The discharge of untreated or poorly treated sewage is the major source of organic pollution that leads to high $BOD_5$ levels. Abatement technologies are available: organic pollution can be curbed by the installation of wastewater treatment plants or by reducing the amount of sewage discharged into the river. Yet most of these remedies are costly, even for developed countries. Third, this indicator measures central goals of the transboundary water agreements analyzed and is available in time-series format. Finally, $BOD_5$ allows for the measurement of point-source pollution that is relatively easy to identify and quantify.

In our statistical analyses, we *invert* the scale such that higher values of the dependent variable represent higher water quality.[8]

$$\text{lBODquality} = (-1) * \log(\text{BOD}_5).$$

## *Independent Variables: Effects of International Agreements*

For analytical purposes, the form of an agreement is understood as dichotomous: an agreement is either legally binding or legally nonbinding. The United Nations

Environment Programme's (UNEP 2002) Atlas of International Freshwater Agreements as well as data taken from the Transboundary Freshwater Dispute Database revealed that of the sixty-nine river basins in Europe, forty-five river basins are regulated by 199 LBAs. Only 47 of the 199 LBAs cover water *quality* issues. Moreover, five LBAs cover more than one river basin. Omitting double counting, we arrive at thirty-six LBAs. Additional research detected nine LNBAs (Köppel 2014). In summary, we include thirty-six LBAs and nine LNBAs covering transboundary water quality in Europe. Due to additional criteria for inclusion in our data set (see Data Set section), we analyzed fifteen LBAs and nine LNBAs.

Our sample includes transboundary water quality agreements and protocols for river basins but excludes amendments to agreements or protocols. There are considerable institutional design differences between the majority of agreements and protocols, on the one hand, and between amendments to those, on the other. Amendments to protocols or agreements in most cases contain only minor adjustments that, if included, would not change the values of the key explanatory variables significantly (Bernauer et al. 2013). We include agreements as of the "date of signature" (see Neumayer 2003; Von Stein 2005) and remove agreements upon expiry.

Besides the form of an agreement, we use a more nuanced approach by combining the form of an agreement each with an index of "precision" (combining depth and clarity of obligations) and an index of "delegation" (combining a scale on monitoring provisions with a scale on dispute resolution mechanisms). Precision and delegation are coded based on a content analysis of the agreement text(s). Table S2 in the Online Supplemental Materials offers an overview of the codebook for agreements. Each design element is coded as a dummy variable indicating whether or not an agreement is precise and whether or not an agreement includes delegation mechanisms.

Instead of employing a simple additive index of form and substance, we explored an alternative composition of all three components. We separated the form (LBA vs. LNBA) from the sum of the indices for precision and delegation (high vs. low scores). This affords us four combinations: *LBA+* with high precision and delegation, *LBA−* with low precision and delegation, *LNBA+* with high precision and delegation, and *LNBA−* with low precision and delegation.

## Control Variables[9]

Beyond the influence of an agreement's institutional design, scholars have argued that other factors influence environmental quality. Hence, we control for these factors as well. The selection of control variables is mainly driven by what other researchers have identified as important determinants for water quality (Bernauer and Kuhn 2010; Köppel 2012; Sigman 2004). The first three variables make the dependent variable comparable across locations, whereas the remaining control variables control for social and economic aspects.

*Flow rate: (lflow).* Water quality is affected strongly by fluctuating river flow since the flow rate has an effect on the dilution of the pollutants. It is the volume of water which passes through a given cross section of the river channel per unit of time. The flow rate is measured in $m^3/s$. We hypothesize a positive effect of the flow rate on water quality.

*Water temperature.* The concentration of dissolved oxygen varies considerably due to temperature fluctuations. We follow past research which used the time rate of exponential decay of $BOD_5$ (known as the “deoxygenation rate,” $k$) to control for the speed of natural attenuation (Sigman 2004, 10; Bernauer and Kuhn 2010). The value for $k$ is calculated by using a nonlinear function from the hydrologic literature (Bowie et al. 1985, 139).[10] We hypothesize a negative effect of the deoxygenation rate $k$ on water quality.

*Population density.* We control for the intensity of the polluting activity. This factor is measured by a population density (*lpopdensus*) proxy. The discharge of untreated or poorly treated sewage is the major source of organic pollution that leads to low $BOD_5$ levels. We hypothesize that population density covaries negatively with $BOD_5$.

*Democracy.* Democracies often behave differently than autocracies. Besides the dyadic claim that democracies rarely, if ever, fight each other, some authors (e.g., Povitkina 2018) suggest that the level of democracy may impact pollution levels. We include the Polity IV scores for the upstream country (*polityus*) as well as the downstream country (*polityds*) and expect a positive effect of democracy on water quality.

*Level of economic development.* Much of the literature on the Environmental Kuznets Curve (EKC) suggests an inverted U-shaped relationship between pollution and economic development, that is, as Mitchell and Deane (2008, 25) argued “‘[r]icher’ means ‘greener.’” We account for the level of economic development of the upstream location by including real income per capita (*lrgdpus*) and its squared term ($lrgdpus^2$). It was suggested that water quality improves at higher levels of economic development (Ringquist and Kostadinova 2005, 94).

*Trade intensity.* Some authors have suggested a beneficial relationship between trade and the environment, that is, “freer” means “cleaner” (Antweiler, Copeland, and Taylor 2001; Frankel and Rose 2005).

We measure the level of trade intensity (*trade_intens*) between the upstream and the downstream country as follows:

$$\text{Trade intensity} = \frac{\text{Exports}_{\text{usds}} + \text{Imports}_{\text{usds}}}{\text{GDP}_{\text{us}}}$$

$\text{Exports}_{\text{usds}}$ ($\text{Imports}_{\text{usds}}$) is the exports (imports) from the upstream (downstream) to the downstream (upstream) country in the dyad, and $\text{GDP}_{\text{us}}$ is the GDP of the upstream country (Sigman 2004, 8). We expect a positive effect of trade intensity on water quality.

*European Union (EU) membership.* Past research argued that membership in an international or supranational institution, and the EU in particular, may affect state behavior and therefore water quality levels (Bernauer and Kuhn 2010, 81-82). Being a member of the EU may have benign effects on pollution levels due to the EU's binding environmental standards.[11] We control whether EU membership is upstream only (*euusonly*), downstream only (*eudsonly*), or on both ends of the river-basin dyad (*eu_both*); we expect a positive effect of EU membership on water quality.

### *Data Set*

Our data set (see Online Supplemental Material for sources) only includes transboundary waterways that fulfill the following requirements:

- falls into geographical Europe after WWII (due to single source availability and homogenous data standards),
- all measuring stations fall within 5 km of the international border (Bernauer and Kuhn 2010, 84),
- have at least ten observation points on the dependent variable during 1954 to 2007, and
- clear upstream and downstream configurations to control for problem structure.

For most agreements, data are available over long periods of time. In particular, for the majority of cases, data are available before agreements were signed. Moreover, some countries share different rivers with different neighbors leading to cross-sectional variation even within states. This justifies employing the river dyad year as the unit of analysis.

The data set includes up to 692 observations covering 39 transboundary rivers and 43 country dyads from 1954 until 2007. The dependent variable ($BOD_5$) measures the levels of transboundary water quality and is based on the annual mean of point-source pollution data provided by the European Environment Agency (2010).

The resulting data set is an (unbalanced) panel data set containing observations for countries and rivers over time (annual averages).

## Findings

Did water quality in European upstream/downstream river-basin dyads improve over the past decades? A first glimpse is offered by Figure 1. For all river-basin dyads included in the data set, whether regulated or not, the lowess analysis of water quality (as measured by the inverted log of $BOD_5$) points to an improvement over the second half century of the 20th century (Figure 1), yet we also find improvements in

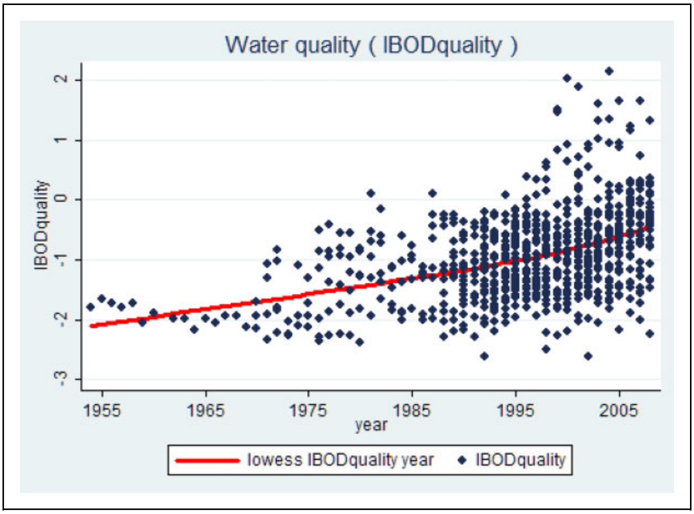


**Figure 1.** Water quality biological oxygen demand in Europe's transboundary rivers (all dyads).

water quality for legally binding and LNBAs (Figures S1 to S3). Did legally binding (LBAs) or LNBAs account for additional improvements in river water quality—or are other political and economic factors at work?

In our analyses, we employed time and panel (river dyad) dummies in order to arrive at a difference-in-difference estimator that controls for potentially secular improvements in water quality even in unregulated river basins. We will proceed in two steps. First, we will focus purely on the legal form—LBAs versus LNBAs. Second, we combine the form of agreement with indices of precision and delegation to arrive at more nuanced results (LBA+, LBA−, LNBA+, and LNBA−). We also probed the results for different possibilities of clustering, especially across country dyads and within river basins. In addition, we will contrast various subperiods.

Before we proceed with the analysis, a look at the temporal evolution of regulatory efforts over time is in order. As Figure 2 demonstrates, among the dyads included in our analysis, few cases of regulation existed until 1987. Given the end of the Cold War in Europe around 1990 and the growth of river regulation at that point in time, we will concentrate on the period beginning in 1990 until 2007. Figure 2 also indicates that we have data for a sufficiently large number of control ("unregulated") cases since 1990. Furthermore, we witness a substantial degree of variation in the number of nonbinding agreements during 1990 to 2007 in both directions, including close to demise of nonbinding agreements by the end of our temporal domain—while the number of LBAs remained reasonably constant in the aggregate beginning 1994. Our main analyses will concentrate on the period 1990 to

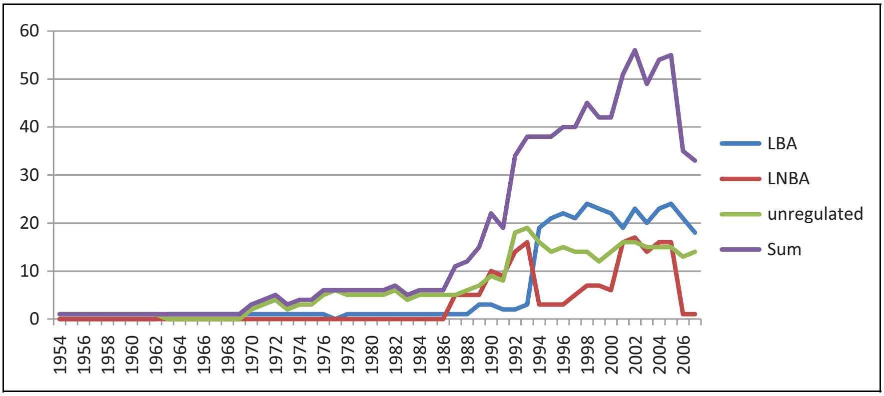


**Figure 2.** Distribution of legally binding and legally nonbinding agreements over time.

2007, with robustness checks encompassing either all periods with data availability or 1994 to 2007 (see Online Supplemental Material).

In the baseline model (model 1.1 in Table 1), we control for the effects of water quality–related aspects to afford greater comparability across our dyads and analyses. The results indicate that the rate of water flow (lflow) increases water quality, whereas water temperature (k) has no effect on average water quality. Increasing population density on the upstream side (lpopdenus) of a river dyad improves water quality. While the first two variables neither reach statistical significance, population density of the upstream part of the dyad has a consistently benign and often statistically discernible positive effect on water quality in our analyses without clustering (see Tables 1 and 3).

Once we add the core variables of interest, namely, whether the *form* of international agreements has an impact on water quality, model 1.2 (Table 1) offers a first glimpse that LBAs foster improvements of water quality, while LNBAs have the opposite, yet statistically insignificant effects. Once all political and economic control variables are included in model 1.3, this wedge between LBAs and LNBAs is maintained: binding agreements have a benign and nonbinding agreements no effect on water quality in upstream/downstream transboundary river dyads in post-WWII Europe.

Could some of the effects, especially the malign effects, differ across subgroups? From our database, we know that all river dyad years of the Danube river basin comprised the only dyads where the *first* agreement was of a legally nonbinding nature. Thus, we exclude all dyads of the Danube river basin from the analysis of model 1.4 and pitch the remaining cases against all unregulated dyads, while in model 1.5, we contrast all dyads falling in the Danube river basin with all unregulated dyads. Comparing the results for models 1.4 and 1.5 illustrates that LBAs maintain their benign effect on water quality, while LNBAs have no effects on water

**Table 1.** Explaining Change in Water Quality During 1990 to 2007.

| lBODquality | (1.1) | (1.2) | (1.3) | (1.4) | (1.5) |
|---|---|---|---|---|---|
| lflow | 0.122 (0.149) | 0.135 (0.149) | 0.119 (0.147) | 0.192 (0.169) | 0.0528 (0.197) |
| *k* | 0.00585 (0.585) | 0.0495 (0.583) | −0.0188 (0.576) | 0.189 (0.615) | −0.192 (0.717) |
| lpopdensus | 1.450** (0.673) | 1.342** (0.672) | 2.219** (0.970) | 2.202** (1.087) | 2.141* (1.165) |
| LBA | | 0.183** (0.0835) | 0.302*** (0.0888) | 0.354*** (0.121) | 0.313** (0.139) |
| LNBA | | −0.0304 (0.0508) | −0.00237 (0.0510) | −0.0119 (0.0864) | −0.00184 (0.0885) |
| polityus | | | 0.0312*** (0.00966) | 0.0362*** (0.00941) | 0.0244 (0.0164) |
| polityds | | | −0.0107 (0.00746) | −0.0505*** (0.0158) | −0.0131 (0.00879) |
| lrgdpus | | | −4.077** (1.613) | −9.149*** (1.942) | −4.468** (1.997) |
| $lrgdpus^2$ | | | 70.84** (28.63) | 164.7*** (34.49) | 74.70** (34.70) |
| trade_intens | | | −0.0859 (0.0625) | −0.0876 (0.0659) | −0.195* (0.102) |
| euusonly | | | −0.149* (0.0764) | 0.0485 (0.177) | −0.161* (0.0911) |
| eudsonly | | | −0.231*** (0.0856) | −0.270*** (0.0925) | −0.292*** (0.109) |
| eu_both | | | −0.115 (0.0959) | −0.156 (0.115) | −0.153 (0.125) |
| _cons | −10.53*** (3.887) | −10.20*** (3.876) | −55.94*** (21.06) | −113.7*** (24.89) | −46.94** (22.18) |
| $N$ | 577 | 577 | 564 | 345 | 420 |
| $R^2$ | .771 | .774 | .791 | .796 | .761 |

*Note*: Standard errors in parentheses; river-dyad dummies and yearly time dummies omitted. Model 1.1 to 1.3 includes all river-basin dyads, whether regulated or unregulated. Model 1.4 compares all regulated river-basin dyads—excluding the Danube river-basin dyads—with unregulated river-basin dyads. Model 1.5 compares all Danube river-basin dyads with unregulated river-basin dyads. LBA = legally binding agreement; LNBA = legally nonbinding agreement.

*$p < .10$.

**$p < .05$.

***$p < .01$.

quality. In summary, LBAs have a benign effect on water quality and LNBAs have no effect. These findings support Hypotheses 1 and 2.

A few notes on our control variables merit mentioning. In general, adding them to the equations does not change the overall findings. The degree of democracy versus autocracy for the upstream part of the dyad (polityus) has a positive effect on water quality, whereas the degree of democracy of the downstream country (polityds) has a negative effect.

Economic wealth of the upstream dyad member (lrgdpus and $lrgdpus^2$) consistently exhibits EKC-like behavior: increases in GDP improve water quality. By contrast, trade intensity has a consistently negative impact on water quality.

During the early decades of our observation period, only few, mostly Western European countries were members of the EU. EU membership increased after 1990, most notably with the inclusion of ten Central and East European countries in 2004 and later additions. All newcomers have to abide by the existing regulation of the EU ("acquis communautaire") in order to become members. As the EU is expected to have a benign effect on those countries with the lowest regulatory and enforcement standards (Avrami and Sprinz 2019), we expect a positive effect of EU membership on water quality. While the reference case refers to none of the dyad members being an EU member, we included dummies for EU upstream members (euusonly), EU downstream members (eudsonly), and if both parts of the dyad consisted of EU members (eu_both). At variance with expectations, we find that EU membership has no or a detrimental effect on water quality, and some of the coefficients are statistically significant, especially if only the downstream member of a river dyad was an EU member. These results cast doubts on whether the EU is necessarily a benign actor on water quality in Europe, especially since eu_both has no statistically significant effect.

Our analyses assumed statistical independence between all dyads in a river basin and across country dyads. So far, we have focused on the particular measurement station in a specific river basin and included river dyad dummies to focus on changes in water quality over time. If clustering were indeed occurring, the reported standard errors would be too small, and we might draw too optimistic conclusions from our findings in Table 1. We replicated Table 1 with clustering by country dyads (to account for potential equal treatment across the same river-basin dyads, Table S4), by river-basin dyads (to account for correlated behavior across dyads falling into the same river basin, Table S5), or both (Table 2). As one-way clustering delivers essentially the same results as two-way clustering, we focus on the latter (Table 2). Taking account of clustering upholds the findings of Table 1 that LBAs have a benign effect on water quality, whereas LNBAs have no statistically discernible effect.[12]

Could our findings originate from too narrow a focus on purely the form of an agreement rather than a combination of form and *substance*? To shed light on this question, we combined the form of an agreement (legally binding vs. legally nonbinding), each with an index of precision and an index of delegation (LBA+, LBA−, LNBA+, LNBA−; see Table S2 for the detailed coding procedures; LBA+

**Table 2.** Explaining Change in Water Quality During 1990 to 2007 with Clustering for Country Dyads and for River Basins.

| lBODquality | (2.1) | (2.2) | (2.3) | (2.4) | (2.5) |
|---|---|---|---|---|---|
| lflow | 0.122 (0.0994) | 0.135 (0.0964) | 0.119 (0.101) | 0.192 (0.160) | 0.0528 (0.112) |
| *k* | 0.00585 (0.353) | 0.0495 (0.358) | −0.0188 (0.369) | 0.189 (0.611) | −0.192 (0.427) |
| lpopdensus | 1.450 (1.606) | 1.342 (1.622) | 2.219** (0.897) | 2.202 (1.290) | 2.141** (0.836) |
| LBA | | 0.183* (0.105) | 0.302*** (0.0922) | 0.354** (0.151) | 0.313** (0.129) |
| LNBA | | −0.0304 (0.0439) | −0.00237 (0.0421) | −0.0119 (0.0522) | −0.00184 (0.0710) |
| polityus | | | 0.0312*** (0.0106) | 0.0362*** (0.0116) | 0.0244 (0.0285) |
| polityds | | | −0.0107* (0.00600) | −0.0505** (0.0226) | −0.0131** (0.00626) |
| lrgdpus | | | −4.077* (2.029) | −9.149*** (2.427) | −4.468* (2.334) |
| $lrgdpus^2$ | | | 70.84* (38.41) | 164.7*** (42.86) | 74.70* (41.59) |
| trade_intens | | | −0.0859 (0.0718) | −0.0876 (0.0725) | −0.195** (0.0773) |
| euusonly | | | −0.149** (0.0591) | 0.0485 (0.120) | −0.161* (0.0923) |
| eudsonly | | | −0.231* (0.125) | −0.270* (0.148) | −0.292** (0.116) |
| eu_both | | | −0.115 (0.120) | −0.156 (0.170) | −0.153 (0.188) |
| _cons | −10.53 (8.709) | −10.20 (8.718) | −47.04* (24.01) | −113.6*** (28.20) | −46.94* (25.21) |
| *N* | 577 | 577 | 564 | 345 | 420 |
| $R^2$ | .771 | .774 | .791 | .796 | .761 |

*Note*: Standard errors in parentheses; with clustering for country dyads and for river basins; river-dyad dummies and yearly time dummies omitted. Model 2.1 to 2.3 includes all river-basin dyads, whether regulated or unregulated. Model 1.4 compares all regulated river-basin dyads—excluding the Danube river-basin dyads—with unregulated river-basin dyads. Model 2.5 compares all Danube river-basin dyads with unregulated river-basin dyads. LBA = legally binding agreement; LNBA = legally nonbinding agreement.

*$p < .10$.

**$p < .05$.

***$p < 0.01$.

represents a LBA with high precision and delegation, etc.). Table 3 replicates the models of Table 1 with our more finely grained core variables of interest.

Our analysis of all cases indicates that a combination of LBAs with high precision and delegation (LBA+) improves water quality (model 3.3); we find support for all our five hypotheses. In effect, only LBA+ shows positive and statistically significant effects—which echoes our earlier findings. Once the Danube river-basin dyads are excluded (model 3.4), we cannot evaluate Hypotheses 3 through 5 due to missing LBA+ cases, yet none of the remaining combinations of bindingness and high or low degrees of precision and delegation show a discernible effect at conventional levels of statistical significance.

Once we concentrate on the dyads of the Danube river basin (model 3.5), entries for LBA− and LNBA+ are lacking cases, while the remaining coefficients are not statistically different from zero. The coefficient for LBA+ is however somewhat smaller than that for LNBA−, pointing into the direction of Hypothesis 4. We replicated Table 3 with clustering by country dyads (Table S6) or by river-basin dyads (Table S7), or both (Table 4). One-way clustering delivers essentially the same results as two-way clustering. Hence, we focus on Table 4 and find that replicating the previous analyses with two-way clustering yields the additional insight that LBA+ has roughly twice the effect of LBA− (model 4.3), and the coefficient reaches standard levels of statistical significance.[13]

An additional perspective of the effect of LBA versus LNBAs is offered by contrasting the coefficients of LBA+ with LNBA+ and LBA− with LNBA−. Wherever such comparisons are possible (see Tables 3 and 4), the LBA outperformed the LNBA, lending support to Hypotheses 3 through 5.

The previous analyses suggest that LBAs outperform LNBAs regardless of the specification of the particular operationalization of the variable of core interest. To provide a direct test whether LBAs outperform LNBAs, we restrict our case domain to river basins that have been regulated (see Table 5). We find that LBAs outperform LNBAs (models 5.2 and 5.3), which is also upheld in case of two-way clustering (models 5.5 and 5.6) as well as in one-way clustering by country dyads (Table S8) or by river-basin dyads (Table S9). In summary, the main argument brought forward in the hypotheses remains supported: LBAs beat LNBAs in terms of improving water quality.

## *Robustness Checks*

To check the robustness of our findings, we focused on a range of time periods and report only unclustered findings. We lengthened the period under consideration to all cases ("over time"), on the one hand, and shortened it to the period 1994 to 2007, on the other. As the Tables S10 to S13 show, all our findings from Tables 1, 3, and 5 are upheld: LBAs have a benign effect on water quality, and LNBAs have no effect or even a negative effect. In addition, we undertook further robustness checks (see Online Supplemental Material for details).

**Table 3.** Explaining Change in Water Quality During 1990 to 2007.

| lBODquality | (3.1) | (3.2) | (3.3) | (3.4) | (3.5) |
|---|---|---|---|---|---|
| lflow | 0.122 (0.149) | 0.105 (0.149) | 0.0831 (0.149) | 0.106 (0.169) | 0.0530 (0.196) |
| *k* | 0.00585 (0.585) | −0.0261 (0.583) | −0.136 (0.579) | −0.0681 (0.616) | −0.191 (0.714) |
| lpopdensus | 1.450** (0.673) | 1.377** (0.679) | 2.395** (0.975) | 2.554** (1.097) | 2.140* (1.163) |
| LBA+ | | 0.338** (0.145) | 0.354** (0.147) | | −0.531 (0.661) |
| LBA− | | 0.215 (0.150) | 0.182 (0.149) | 0.159 (0.132) | |
| LNBA+ | | 0.0706 (0.181) | 0.129 (0.179) | 0.0155 (0.160) | |
| LNBA− | | 0.0780 (0.109) | 0.0367 (0.108) | 0.0690 (0.0972) | −0.845 (0.681) |
| polityus | | | 0.0204** (0.00946) | 0.0236*** (0.00857) | 0.0244 (0.0164) |
| polityds | | | −0.0129* (0.00758) | −0.0497*** (0.0161) | −0.0132 (0.00869) |
| lrgdpus | | | −4.293*** (1.618) | −9.136*** (1.971) | −4.469** (1.994) |
| $lrgdpus^2$ | | | 75.17*** (28.71) | 165.3*** (34.99) | 74.73** (34.62) |
| trade_intens | | | −0.0972 (0.0634) | −0.105 (0.0667) | −0.195* (0.101) |
| euusonly | | | −0.166** (0.0776) | −0.0212 (0.178) | −0.161* (0.0909) |
| eudsonly | | | −0.242*** (0.0867) | −0.265*** (0.0939) | −0.292*** (0.109) |
| eu_both | | | −0.147 (0.0997) | −0.148 (0.117) | −0.154 (0.125) |
| _cons | −10.53*** (3.887) | −10.42*** (3.912) | −59.36*** (21.10) | −115.2*** (25.25) | −52.90** (24.54) |
| *N* | 577 | 577 | 564 | 345 | 420 |
| $R^2$ | .771 | .776 | .791 | .791 | .761 |

*Note*: Standard errors in parentheses; river-dyad dummies and yearly time dummies omitted. Model 3.1 to 3.3 includes all river-basin dyads, whether regulated or unregulated. Model 3.4 compares all regulated river-basin dyads—excluding the Danube river-basin dyads—with unregulated river-basin dyads. Model 3.5 compares all Danube river-basin dyads with unregulated river-basin dyads. LBA = legally binding agreement; LNBA = legally nonbinding agreement.

*$p < .10$.

**$p < .05$.

***$p < .01$.

**Table 4.** Explaining Change in Water Quality During 1990 to 2007 with Clustering for Country Dyads and for River Basins.

| lBODquality | (4.1) | (4.2) | (4.3) | (4.4) | (4.5) |
|---|---|---|---|---|---|
| lflow | 0.122 (0.0994) | 0.105 (0.0971) | 0.0831 (0.0960) | 0.106 (0.150) | 0.0530 (0.113) |
| *k* | 0.00585 (0.353) | −0.0261 (0.353) | −0.136 (0.355) | −0.0681 (0.573) | −0.191 (0.417) |
| lpopdensus | 1.450 (1.606) | 1.377 (1.620) | 2.395** (0.941) | 2.554 (1.509) | 2.140** (0.832) |
| LBA+ | | 0.338** (0.165) | 0.354** (0.168) | 0 (0.000000290) | −0.531 (0.439) |
| LBA− | | 0.215*** (0.0706) | 0.182** (0.0814) | 0.159 (0.111) | 0 (0.000000141) |
| LNBA+ | | 0.0706 (0.0678) | 0.129 (0.0866) | 0.0155 (0.0607) | 0 (9.84e-08) |
| LNBA− | | 0.0780 (0.120) | 0.0367 (0.113) | 0.0690 (0.0741) | −0.845 (0.508) |
| polityus | | | 0.0204 (0.0133) | 0.0236 (0.0139) | 0.0244 (0.0277) |
| polityds | | | −0.0129** (0.00594) | −0.0497** (0.0230) | −0.0132* (0.00699) |
| lrgdpus | | | −4.293** (2.020) | −9.136*** (2.467) | −4.469* (2.313) |
| $\text{lrgdpus}^2$ | | | 75.17* (37.97) | 165.3*** (43.75) | 74.73* (41.13) |
| trade_intens | | | −0.0972 (0.0732) | −0.105 (0.0788) | −0.195** (0.0760) |
| euusonly | | | −0.166** (0.0614) | −0.0212 (0.130) | −0.161* (0.0939) |
| eudsonly | | | −0.242* (0.126) | −0.265* (0.149) | −0.292** (0.119) |
| eu_both | | | −0.147 (0.128) | −0.148 (0.171) | −0.154 (0.194) |
| _cons | −10.53 (8.709) | −10.42 (8.682) | −49.92** (23.65) | −115.2*** (29.29) | −52.72** (24.86) |
| *N* | 577 | 577 | 564 | 345 | 420 |
| $R^2$ | .771 | .776 | .791 | .791 | .761 |

*Note*: Standard errors in parentheses; with clustering for country dyads and for river basins; river-dyad dummies and yearly time dummies omitted. Model 4.1 to 4.3 includes all river-basin dyads, whether regulated or unregulated. Model 4.4 compares all regulated river-basin dyads—excluding the Danube river-basin dyads—with unregulated river-basin dyads. Model 4.5 compares all Danube river-basin dyads with unregulated river-basin dyads. LBA = legally binding agreement; LNBA = legally nonbinding agreement.

*$p < .10$.

**$p < .05$.

***$p < .01$.

**Table 5.** Explaining Change in Water Quality During 1990 to 2007 (without Unregulated Cases).

| lBODquality | (5.1) | (5.2) | (5.3) | (5.4) | (5.5) | (5.6) |
|---|---|---|---|---|---|---|
| Lflow | 0.229 (0.224) | 0.229 (0.221) | 0.148 (0.230) | 0.229* (0.112) | 0.229*** (0.0759) | 0.148* (0.0825) |
| *K* | 1.495 (2.163) | 0.903 (2.156) | 0.461 (2.191) | 1.495 (1.665) | 0.903 (1.191) | 0.461 (1.391) |
| lpopdensus | 4.470*** (1.170) | 5.104*** (1.187) | 1.079 (2.488) | 4.470*** (0.925) | 5.104*** (0.465) | 1.079 (1.308) |
| LBA | | 0.456*** (0.162) | 0.528*** (0.175) | | 0.456*** (0.0269) | 0.528*** (0.0492) |
| LNBA | | −0.0241 (0.0604) | −0.0121 (0.0624) | | −0.0241 (0.0608) | −0.0121 (0.0682) |
| polityus | | | −0.0202 (0.0312) | | | −0.0202 (0.0222) |
| Polityds | | | −0.0101 (0.00931) | | | −0.0101 (0.00799) |
| Lrgdpus | | | 2.956 (3.827) | | | 2.956*** (0.869) |
| $\text{lrgdpus}^2$ | | | −65.68 (71.34) | | | −65.68*** (18.60) |
| trade_intens | | | −0.0444 (0.0784) | | | −0.0444 (0.104) |
| euusonly | | | −0.162* (0.0968) | | | −0.162*** (0.0450) |
| eudsonly | | | −0.0794 (0.133) | | | −0.0794 (0.0951) |
| eu_both | | | −0.249* (0.144) | | | −0.249** (0.119) |
| _cons | −19.74*** (4.945) | −22.10*** (5.071) | 39.36 (52.41) | −19.74*** (4.194) | −22.10*** (2.220) | 39.36** (17.67) |
| *N* | 340 | 340 | 330 | 340 | 340 | 330 |
| $R^2$ | .822 | .827 | .835 | .822 | .827 | .835 |

*Note*: Standard errors in parentheses; river-dyad dummies and yearly time dummies omitted. Models 5.1 to 5.3: without clustering, models 5.4 to 5.6: with clustering for country dyads and for river basins. LBA = legally binding agreement; LNBA = legally nonbinding agreement.

*$p < .10$.

**$p < .05$.

***$p < .01$.

First, if we omit the river dyads governed by general cooperation agreements on transboundary waters (i.e., remove the rivers Daugava, Douoro, and Neman from the analyses),[14] the general pattern of support for LBA and LBA+ is maintained; this result holds for both forms of operationalizing the form and substance of agreements (see Tables S14 and S15).

Second, we added the installation of water treatment plants by the upstream country as a further control variable for the first building bloc of water quality–related variables. Replicating Tables 1 and 3 in Tables S16 and S17, we find that LBAs have a positive and statistically discernible effect; however, our results of Table S17 indicate a strong negative effect of LBA+ on water quality. These results are, however, driven by considerable data attrition due to the availability of data for wastewater treatment plants as compared to previous analyses: Nearly 50 percent of cases were lost!

Third, we checked for the robustness of our findings to the inclusion of military capacity as a control variable, specifically the ratio of upstream to downstream of the Composite Indicator of National Capability. The substantive conclusions remain the same, and military capacity itself has no discernible effect on water quality (Tables S18 to S19).

## Conclusions

This article has examined whether LBAs are more effective in improving environmental conditions than international LNBAs. First, we purely focused on the form of LBAs versus LNBAs. Second, we used a more nuanced approach by combining the form of agreement with indices of precision and delegation (LBA+, LBA−, LNBA+, LNBA−).

Overall, we are confident that LBAs and LBA+s nearly universally lead to improved water quality (supporting hypothesis 1); LNBAs have no discernible effect on water quality in the configuration of cases analyzed in this article (supporting Hypothesis 2); and if decision makers have the opportunity to pursue LBAs with high degrees of precision and delegation (LBA+), they will beat the effect of any other combination of form of agreement with an index of precision and delegation (lending support for Hypotheses 3 to 5).

Our findings support an earlier quantitative study by Breitmeier, Young, and Zürn (2006) which showed that legalization of international institutions raises their effectiveness. However, it lends no support to the results of qualitative case studies by Gurtner-Zimmermann (1998), Holtrup (1999), and Tschanz (2001), indicating that soft law can be at least as effective as binding treaties. Our findings also do not support the study by Victor, Raustiala, and Skolnikoff (1998), who argued that soft law has proven more effective in changing the behavior of relevant actors than binding treaties. This might be explained by our article employing the impact or effectiveness dimensions rather than the output or outcome dimension favored by other studies that emphasize behavioral change.

Taken together, three broad conclusions emerge. First, from an academic perspective, LBAs, either alone or in combination with high degrees of precision and delegation, have discernible, positive effects on river water quality in post-WWII Europe if clear upstream/downstream configurations are chosen as the case domain.

Second, from a practitioner's perspective, pursuing LBAs may be worth the effort. Given our case domain, LNBAs are not the benign option some have hoped for.

Third, the additional increment of explanatory power added by social variables, both the type of agreement and the control variables, over simple models that adjust the dependent variable to enhance the comparability of water quality across locations is very modest. This begs the overall question under which more precise constellations a river water quality agreement may be worth the effort to be agreed upon and implemented.

With a view to generalization, we suggest to compare our results on clear upstream/downstream configurations with other types of international river water quality constellations, both inside Europe and across all continents, to extend the issue under consideration from water quality to water quantity and navigation, as well as to include related issue areas such as shared lakes.

## Acknowledgments

We appreciate comments received on earlier versions on occasion of the Second Environmental Politics & Governance Conference, June 16-19, 2016, Seminarhotel Gerzensee, Gerzensee, Switzerland, the Potsdam Center for Quantitative Research (PCQR), May 6, 2015, Campus Griebnitzsee, University of Potsdam, Potsdam, Germany, the International Studies Association Annual Convention 2011, Montreal, March 16-19, 2011, and the annual meeting of the Austrian, German, and Swiss Political Science Associations (Dreiländertagung) 2011, Basel, Switzerland, January 13-14, 2011, Michaël Aklin, and Jonathan Olmsted, as well as two reviewers of the *Journal of Conflict Resolution*.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Martin Köppel gratefully acknowledges research support from the German Federal Environmental Foundation (Deutsche Bundesstiftung Umwelt).

## Supplemental Material

Supplemental material for this article is available online.

## Notes

1. Agreements are defined here synonymously with regimes following Krasner's (1983, 2) broadly accepted definition.
   A legally binding treaty is defined as an intergovernmental agreement in written form and governed by international law "whether embodied in a single instrument or in two or more related instruments and whatever its particular designation" (United Nations [UN] 1969, Article 2(1a)). Examples of binding treaties include the 2010 New START agreement between the USA and Russia, the 1997 Kyoto Protocol, or the 1985 Vienna Convention. See also Köppel (2012, 18-19).
   A legally nonbinding agreement is used for an intergovernmental agreement in written form that is not governed by international law, "whether embodied in a single instrument or in two or more related instruments and whatever its particular designation" (United Nations 1969, Article 2(1a)). Particularly important is the requirement for an agreement to be in written form. Numerous nonbinding pledges have not been published and are no more than "oral" or "informal" promises (Lipson 1991; Aust 1986). Nonbinding instruments, as used in this article, include documented action plans, codes of conduct, agreed measures, resolutions, and similar policies (Mitchell 2003, 431-34; Thürer 2000). Well-known examples of nonbinding agreements are the 1975 Helsinki Final Act, the North Sea Declarations, Agenda 21, as well as the Forest Principles adopted by the 1992 Rio Conference on Environment and Development. For analytical purposes, the terms binding and nonbinding are understood as dichotomous. An agreement is either legally binding or legally nonbinding.
2. In contrast to other valuable contributions to the field, we employ multiple operationalizations of the core variables of interest. While, for example, Bernauer et al. (2013) use several treaty characteristics as independent variables, these are quite different from our second operationalization of our core variables.
3. For a more thorough explanation of these causal mechanisms, see Köppel (2012, 21-27).
4. Past research found that delegating the authority to monitor and enforce obligations of an agreement to third parties is likely to enhance regime effectiveness (Downs, Rocke, and Barsoom 1996; Böhmelt and Pilster 2010).
5. A transboundary river is defined as a river that either crosses or demarcates international political boundaries (United Nations Environment Programme 2002, 6).
6. While some of the agreements included in the analysis are bilateral, most of them are multilateral. Including a multilateral agreement (MEA) implies that there is more than one dyad involved. Hence, in all instances where MEAs are involved, we included—according to the case selection criteria outlined in the Data Set section—all possible dyads in the data set. In a subsequent analysis, we offer clustering by river basins (and country dyads).
7. Concentrations normally increase as a result of organic pollution. $BOD_5$ levels between 1 and 2 mg/l signify a relatively clean river, $BOD_5$ concentrations of more than 5 mg/l indicate a relatively polluted river (Meybeck, Chapman, and Helmer 1990, 80; European Environment Agency 2009).

8. An “l” in front of the operational variable name points to logarithmic transformation. “us” refers to upstream, and “ds” refers to downstream in a variable name.
9. We did not include a proxy to control for the EU Water Framework Directive (WFD). The WFD (2000/60/EC) was enacted on October 23, 2000, and its goals are focused on achieving a good water status for all surface waters in Europe by the year 2015. We omitted the WFD because most of the analyzed agreements were signed and implemented before the WFD was negotiated in 2000. However, we include EU membership dummies to capture potential EU effects more generally.
10. We are indebted to Hillary Sigman for assistance with calculating the deoxygenation rate.
11. Several EU environmental standards may have contributed to a change in country’s pollution behavior. In particular, the EU’s WFD (2000/60/EC) has been in effect since the year 2000, but earlier environmental regulations, such as the Urban Waste Water Treatment Directive, provide for secondary (biological) wastewater treatment as well.
12. The two-way clustering accounted for clustering of 22 river basins and 34 country dyads. We note that, counter to expectations, not all standard errors in Table 2 are larger than those found in Table 1, although the magnitude in difference is small. Two-way clustering resulted in an initial variance–covariance matrix that was not positive semidefinite; for the analyses, negative eigenvalues were replaced with zero values (see http://fmwww.bc.edu/RePEc/bocode/v/vce2way.sthlp for details).
13. In both cases, the standard errors of both coefficients of LBA− shrank considerably in models 4.2 and 4.3 as compared to models 3.2 and 3.3. The caveats of the clustering method mentioned in the previous analysis also apply here, yielding extremely minimal standard errors for variables where there are no cases.
14. These general agreements are different than the other river-specific agreements. They concern all rivers these countries share and are accompanied by considerable institutional design differences that could lead to biased results.

## References

Abbott, Kenneth W., Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter, and Duncan Snidal. 2000. “The Concept of Legalization.” *International Organization* 54 (3): 401-19.

Abbott, Kenneth W., and Duncan Snidal. 2000. “Hard and Soft Law in International Governance.” *International Organization* 54 (3): 421-56.

Antweiler, Werner, Brain R. Copeland, and M. Scott Taylor. 2001. “Is Free Trade Good for the Environment?” *American Economic Review* 91 (4): 877-908.

Aust, Anthony. 1986. “The Theory and Practice of Informal International Instruments.” *International & Comparative Law Quarterly* 35 (4): 787-812.

Avrami, Lydia, and Detlef F. Sprinz. 2019. “Measuring and Explaining the EU’s Effect on National Climate Performance.” *Environmental Politics* 28 (5): 822-846. doi: 10.1080/09644016.2018.1494945

Bernauer, Thomas, Tobias Böhmelt, and Vally Koubi. 2013. “Is There a Democracy–Civil Society Paradox in Global Environmental Governance?” *Global Environmental Politics* 13 (1): 88-107.

Bernauer, Thomas, and Anna Kalbhenn. 2010. "The Politics of International Freshwater Resources." In *International Studies Encyclopedia Online*, edited by R. A. Denemark, 5800-21. Oxford: Wiley-Blackwell Publishers.

Bernauer, Thomas, Anna Kalbhenn, Vally Koubi, and Gabriele Spilker. 2013. "Is There a "Depth Versus Participation" Dilemma in International Cooperation?" *The Review of International Organizations* 8 (4): 477-97.

Bernauer, Thomas, and Patrick M. Kuhn. 2010. "Is There an Environmental Version of the Kantian Peace? Insights from Water Pollution in Europe." *European Journal of International Relations* 16 (1): 77-102.

Bernauer, Thomas, and Peter Moser. 1996. "Reducing Pollution of the Rhine River: The Influence of International Cooperation." *Journal of Environment and Development* 5 (4): 391-417.

Bodansky, Daniel. 2010. *The Art and Craft of International Environmental Law*. Cambridge, MA: Harvard University Press.

Böhmelt, Tobias, and Ulrich H. Pilster. 2010. "International Environmental Regimes: Legalisation, Flexibility and Effectiveness." *Australian Journal of Political Science* 45 (2): 245-60.

Bowie, G. L., W. B. Mills, D. B. Porcella, C. L. Campbell, and J. R. Pagenkopf. 1985. *Rates, Constants, and Kinetics Formulations in Surface Water Quality Modelling*. Athens, GA: US EPA.

Breitmeier, Helmut, Arild Underdal, and Oran R. Young. 2011. "The Effectiveness of International Environmental Regimes: Comparing and Contrasting Findings from Quantitative Research." *International Studies Review* 4 (13): 579-605.

Breitmeier, Helmut, Oran R. Young, and Michael Zürn. 2006. *Analyzing International Environmental Regimes. From Case Study to Database*. Cambridge, MA: MIT Press.

Brown Weiss, Edith. 2007. "Conclusions: Understanding Compliance with Soft Law." In *Commitment and Compliance: The Role of Non-binding Norms in the International Legal System*, edited by Dinah Shelton, 535-53. Oxford, UK: Oxford University Press.

Chayes, Abram, and Antonia Handler Chayes. 1993. "On Compliance." *International Organization* 47 (2): 175.

Conca, Ken. 2006. *Governing Water: Contentious Transnational Politics and Global Institution Building, Global Environmental Accords*. Cambridge, MA: MIT Press.

Downs, George W., David Rocke, and Peter N. Barsoom. 1996. "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50 (3): 379-406.

Easton, David. 1965. *A Systems Analysis of Political Life*. New York: Wiley.

European Environment Agency. 2009. EEA Core Set Indicator on Oxygen Consuming Substances in Rivers (Csi 019). Copenhagen.

European Environment Agency. 2010. Waterbase Rivers. Version 10. Quality Control Documentation. Copenhagen.

Fischhendler, Itay. 2008. "When Ambiguity in Treaty Design Becomes Destructive: A Study of Transboundary Water." *Global Environmental Politics* 8 (1): 111-36.

Frankel, Jeffrey A., and Andrew K. Rose. 2005. "Is Trade Good or Bad for the Environment? Sorting out the Causality." *Review of Economics and Statistics* 87 (1): 85-91.

Goldstein, Judith L., Douglas Rivers, and Michael Tomz. 2007. "Institutions in International Relations: Understanding the Effects of the GATT and the WTO on World Trade." *International Organization* 61 (1): 37-67.

Gurtner-Zimmermann, Arnold. 1998. "The Effectiveness of the Rhine Action Program. Methodology and Results of an Evaluation of the Impacts of International Cooperation." *International Environmental Affairs* 10 (1): 241-67.

Guzman, Andrew T. 2005. "The Design of International Agreements." *European Journal of International Law* 16 (4): 579-612.

Holtrup, Petra. 1999. "Der Schutz Grenzüberschreitender Flüsse in Europa - Zur Effektivität Internationaler Umweltregime." *Dissertation no. 3642*, Berichte des Forschungszentrums Jülich, Jülich.

Jacobson, Harold K., and Edith Brown Weiss. 1998. "Assessing the Record and Designing Strategies to Engage Countries." In *Engaging Countries: Strengthening Compliance with International Environmental Accords*, edited by Edith Brown Weiss and Harold K. Jacobson, 511-54. Cambridge, MA: MIT Press.

Köppel, Martin. 2009a. *Does Regime Design Matter? Analyzing Success and Failure of International Cooperation*. Munich: AVM.

Köppel, Martin. 2009b. "The Effectiveness of Soft Law: First Insights from Comparing Legally Binding Agreements with Flexible Action Programs." *Georgetown International Environmental Law Review* 21 (4): 821-35.

Köppel, Martin. 2012. *Explaining the Effectiveness of Binding and Non-binding Agreements. Lessons from Transboundary Water Quality in Europe's River Basins*. Unpublished dissertation, University of Tuebingen, Tuebingen.

Köppel, Martin. 2014. "Nonbinding Agreements as an Effective Alternative? A First Assessment of Transboundary River Pollution in Europe." *International Journal of Water* 8 (1): 1-16.

Krasner, Stephen D. 1983. *International Regimes*. Ithaca, NY: Cornell University Press.

Lipson, Charles. 1991. "Why Are Some International Agreements Informal?" *International Organization* 45 (4): 495-538.

Meybeck, M., Deborah V. Chapman, and Richard Helmer. 1990. *Global Freshwater Quality: A First Assessment*. Cambridge, MA: Published on Behalf of the World Health Organization and the United Nations Environment Programme by Blackwell Reference.

Miles, Edward L., Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjaerseth, and Elaine M. Carlin. 2002. *Environmental Regime Effectiveness: Confronting Theory with Evidence*. Cambridge, MA: MIT Press.

Mitchell, Ronald B. 2003. "International Environmental Agreements: A Survey of Their Features, Formation, and Effects." *Annual Review of Environment and Resources* 28:429-61.

Mitchell, Ronald B. 2007. "Compliance Theory: Compliance, Effectiveness, and Behavior Change in International Environmental Law." In *Oxford Handbook of International Environmental Law*, edited by Jutta Brunnee, Daniel Bodansky, and Ellen Hey, 893-921. Oxford, UK: Oxford University Press.

Mitchell, Ronald B. 2010. *International Politics and the Environment. 1st ed, Sage Series on the Foundations of International Relations*. Thousand Oaks, CA: Sage.

Mitchell, Ronald, and Glenn D. Deane. 2008. "Comparing Institutional Influence: The Relative Effectiveness of Three Environmental Agreements." Paper presented at the Annual Meeting of the International Studies Association, Bridging Multiple Divides, San Francisco, Mar 26-29.

Mitchell, Ronald B., and Patricia Keilbach. 2001. "Reciprocity, Coercion, or Exchange: Symmetry, Asymmetry and Power in Institutional Design." *International Organization* 55 (4): 891-917.

Neumayer, Eric. 2003. "Are Left-wing Party Strength and Corporatism Good for the Environment? Evidence from Panel Analysis of Air Pollution in OECD Countries." *Ecological Economics* 45 (2): 203-20.

Povitkina, Marina. 2018. "The Limits of Democracy in Tackling Climate Change." *Environmental Politics* 27 (3): 411-32.

Prins, Gwyn, and Steve Rayner. 2007. "Time to Ditch Kyoto." *Nature* 449 (7165): 973-75.

Raustiala, Kal. 2005. "Form and Substance in International Agreements." *The American Journal of International Law* 99 (3): 581-614.

Ringquist, Evan J., and Tatiana Kostadinova. 2005. "Assessing the Effectiveness of International Environmental Agreements: The Case of the 1985 Helsinki Protocol." *American Journal of Political Science* 49 (1): 86-102.

Roberts, J. Timmons, Bradley C. Parks, and Alexis A. Vásquez. 2004. "Who Ratifies Environmental Treaties and Why? Institutionalism, Structuralism and Participation by 192 Nations in 22 Treaties." *Global Environmental Politics* 4 (3): 22-65.

Salamé, Léna, and Pieter Van der Zaag. 2010. "Enhanced Knowledge and Education Systems for Strengthening the Capacity of Transboundary Water Management." In *Transboundary Water Management: Principles and Practice*, edited by Anton Earle, Anders Jägerskog, and Joakim Öjendal, 171-86. Washington, DC: Earthscan.

Sigman, Hilary. 2004. "Does Trade Promote Environmental Coordination? Pollution in International Rivers." *Contributions to Economic Analysis & Policy* 3 (2): Article 2.

Simmons, Beth. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13 (1): 273-96.

Skjærseth, Jon Birger, Olav Schram Stokke, and Jørgen Wettestad. 2006. "Soft Law, Hard Law, and Effective Implementation of International Environmental Norms." *Global Environmental Politics* 6 (3): 118-19.

Social Learning Group. 2001a. *Learning to Manage Global Environmental Risks. A Comparative History of Social Responses to Climate Change, Ozone Depletion, and Acid Rain,* Vol. 1. Cambridge, MA: MIT Press.

Social Learning Group. 2001b. *Learning to Manage Global Environmental Risks. A Functional Analysis of Social Responses to Climate Change, Ozone Depletion, and Acid Rain*, Vol. 2. Cambridge, MA: MIT Press.

Spilker, Gabriele, and Vally Koubi. 2016. "The Effects of Treaty Legality and Domestic Institutional Hurdles on Environmental Treaty Ratification." *International Environmental Agreements: Politics, Law and Economics* 16 (2): 223-38.

Sprinz, Detlef, and Tapani Vaahtoranta. 1994. "The Interest-based Explanation of International Environmental Policy." *International Organization* 48 (1): 77-105.

Stinnett, Douglas M., and Jaroslav Tir. 2009. "The Institutionalization of River Treaties." *International Negotiation* 14:229-51.

Thürer, Daniel. 2000. "Soft Law." In *Encyclopedia of Public International Law*, edited by R. Bernhardt, 452-60. New York: Elsevier.

Tir, Jaroslav, and Douglas M. Stinnett. 2012. "Weathering Climate Change: Can Institutions Mitigate International Water Conflict?" *Journal of Peace Research* 49 (1): 211-25.

Tschanz, R. 2001. "Gewässerschutz Am Rhein: Zur Lösung Von Grenzüberschreitenden Umweltschutzproblemen [Rhine Water Protection: Solving Transboundary Environmental Problems]." MA thesis, Swiss Federal Institute of Technology (ETH), Center for International Studies, Zurich.

United Nations. 1969. "Vienna Convention on the Law of Treaties." opened for signature May 23, 1969. United Nations Conference on the Law of Treaties, Regis. No. 18,232, *UN Doc. A/CONF.39/27*.

United Nations Environment Programme. 2002. *Atlas of International Freshwater Agreements*. Nairobi: United Nations Organization.

Underdal, Arild. 2002. "One Question, Two Answers." In *Environmental Regime Effectiveness: Confronting Theory with Evidence*, edited by Edward L. Miles, Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjærseth, and Elaine M. Carlin, 3-46. Cambridge, MA: MIT Press.

Victor, David G., Kal Raustiala, and Eugene B. Skolnikoff. 1998. *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. Cambridge, MA: International Institute for Applied Systems Analysis; MIT Press.

Von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99 (4): 611-22.

Von Stein, Jana. 2008. "The International Law and Politics of Climate Change." *Journal of Conflict Resolution* 52 (2): 243-68.

Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S. E. Bunn, C. A. Sullivan, and C. R. Liermann. 2010. "Global Threats to Human Water Security and River Biodiversity." *Nature* 467 (7315): 555-61.

Wettestad, Jørgen. 1995. *"Nuts and Bolts for Environmental Negotiators?" Designing Effective International Regimes: A Conceptual Framework*. Oslo: Fridtjof Nansen Institutt.

Young, Oran R. 2004. "The Consequences of International Regimes." In *Regime Consequences. Methodological Challenges and Research Strategies*, edited by Arild Underdal and Oran Young, 3-24. London: Kluwer Academic.

Young, Oran R. 2011. "Effectiveness of International Environmental Regimes: Existing Knowledge, Cutting-edge Themes, and Research Strategies." *Proceedings of the National Academy of Sciences* 108 (50): 19853-60.

Young, Oran R., and Marc A. Levy. 1999. "The Effectiveness of International Environmental Regimes." In *Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms*, edited by Oran R. Young, 1-32. Cambridge, MA: MIT Press.