

Measuring the effect of water sharing treaties between riparian states

Abstract

Due to climate change, cooperation over freshwater resources is going to become more critical to human security and economic development in the coming years. Most analyses of treaties governing river basins operate under the assumption that such agreements can change states' behavior and are enforceable; however, this is not a settled debate within the wider International Relations Literature. I combine data from the Transboundary Freshwater Dispute Database and the Global Runoff Data Center to estimate the effect of treaties on actual cross-border water flow using an event study design, and find no significant effects.

Author: Laurens Hof, affiliation: AUC. laurens.hof@student.auc.nl

Major: Social Sciences

Supervisor: dr. Erik Ansink, affiliation: VU. erik.ansink@vu.nl

Submission date: 31-01-2020

Keywords: water agreements, event study, river discharge, realism, institutionalism

1: Introduction

In the past century, population growth and rising water consumption (fueled by industrialization) have multiplied and intensified pressures over water access. This trend is predicted to continue due to a higher prevalence of droughts, floods and other weather anomalies. However, instead of water stresses causing widespread conflict, the past half century has seen widespread proliferation in international treaties regulating shared access to this water (Hamner & Wolf, 1998). Since they seemingly have contributed so much to reducing militarized conflicts over water, scholars have been interested in studying their role. However, in this, they often get caught up in a debate within the field of International Relations.

2a: Theoretical background: realism

On the one hand, there is realist theory, which paints a grim picture of the world where the world of international relations is governed only by the most naked of power interests (Mearsheimer, 1994). This is drawn from five relatively basic assumptions about the nature of the international system: firstly, that it is anarchic, meaning that all sovereignty lies with the state without any higher authority that can exercise power over them; secondly, that states are inherently dangerous to each other because they always possess *some* kind of offensive military potential; thirdly, that states can never know other states' intentions for certain; fourthly, that the most fundamental motive of a state is self-preservation; and finally, that they are instrumentally rational in pursuing this motive.

These assumptions lead to a pattern of behavior where states are very suspicious of each other, view their military capabilities as the only thing standing between them and destruction (and are therefore susceptible to getting caught up in arms races). Furthermore, it leads to a mindset with a very strong emphasis on the positionality of power. After all, if all parties gain from a situation, but a rival (which potentially means any other state) gains more than you, the security of your position deteriorates. It is easy to see how this would lead to a zero-sum view of the world: if all relevant expressions of power are fundamentally positional, any gains made must come at the expense of someone else.

If all states are self-serving, suspicious of one another and hold a zero-sum worldview, it would be extremely difficult to reach lasting cooperation, since negotiation processes are deeply affected by the participants' views of the process, both in when looking at negotiation in river basins (Madani, 2010) and in more general contexts (Axelrod & Keohane, 1985). Two obstacles stand out in particular: the aforementioned considerations of relative positional gains and worries that the other party might cheat (Grieco, 1988). The latter is already hard enough to deal with in an anarchic environment like the international system where there is no higher authority that can enforce contracts. The former adds the complications of distributing the already insecure gains of cooperation in a way that has both sides feeling like they are the relative winners of the deal.

In the context of international river water management, realist theory re-evaluates the proliferation of new treaties. While treaties can sometimes be effective means to prevent costly wars, they typically self-select for relatively unimportant issues. In other words, the only agreements that get signed are the ones that do not meaningfully conflict with both parties' vital interests. Furthermore, treaties (and international institutions) only function by codifying existing power relationships (Mearsheimer, 1994), and do not significantly affect a state's incentives or actions. The treaties that change states' behavior are not part of their core security interests, while the treaties that affect vital interests do not change

any behavior since they just put into writing a pre-existing situation that arose from regional power relations. This is supported by the findings by Brochmann (2012) that large power imbalances within a basin make cooperation more likely, and by Zawahri, Dinar and Nigatu (2014) that bilateral treaties (which are often used by preponderant states to magnify the power imbalance) are more likely to address issues of water allocation and other vitally important concerns.

2b: Theoretical background: institutionalism

On the other hand, there is institutionalism, sometimes called liberal institutionalism. It can be seen as an extension of the realist model, since it accepts the core assumptions. However, its conclusions are significantly different. The institutionalist framework makes extensive use of game theoretical models to explain cooperation between states and make policy recommendations to ease this cooperation. The main factors that promote cooperation are that the states are able to look forward along long time horizons where payoffs are regular, and that reliable information about other state's actions is quickly available (Axelrod & Keohane, 1985). These factors are not likely to be present in the sphere of military security, since theoretically, a well-timed surprise strike could wholly eliminate one of the states (therefore making the payoffs in the future a lot less regular) and military activity is kept secret relatively easily (armaments can be hidden, and asymmetric attacks can give plausible deniability). However, there is a lot of potential for cooperation in other fields less sensitive to unilateral defection, provided that states have enough ability to assess and make such decisions (Chayes & Chayes, 1993). In the context of freshwater sharing therefore, emphasis is placed on the reduction of negotiation frictions and the cost of political maintenance (Dinar, De Stefano, Nigatu, & Zawahri, 2019; Zawahri, Dinar, & Nigatu, 2014).

Realist scholars criticize approach on the fact that it is based on arbitrary distinctions between issues that are security-sensitive, and issues that are not. According to Mearseimer (1994), there is no basis for considering such areas as "economic development" not security-sensitive, since a nation's economic prosperity is one of the main determinants of its military capabilities. Furthermore, an argument could be made that, since policymakers are used to operating in a more zero-sum environment, they would place too much emphasis on the positional components of the payoffs, which in turn complicates cooperation, makes defections more likely and thus build further distrust. Nonetheless, empirical work regarding river systems mostly operates in an institutionalist framework, implicitly if not explicitly. Some examples of this include the emphasis placed by Dinar et al. (2019) on transaction costs,

There are two other effects that are also stressed by institutionalist scholars, both of which are grounded in the concept of reciprocity (Axelrod & Keohane, 1985). The first is that linking issues can expand the possibilities for collaboration, both in a general context (Haas, 1980) and specifically regarding river basin negotiation (Eleftheriadou & Mylopoulos, 2008). By "making trade-offs on different chessboards" (Axelrod & Keohane, 1985) solutions that are more socially optimal but were previously hard to reach due to distributional disagreements between states can be achieved. However, there is a downside to this: linkages can make agreements significantly harder to renegotiate, which can exacerbate the effects of unexpected disasters (Fischhendler, 2003). With climate change making freak weather events more common, this can prove to be a severe limitation to existing freshwater treaties (Cooley & Gleick, 2011). The second effect is that establishing and honoring agreements improves the relationship between the signatory states overall, making them more likely to reach deals in the future. While the theoretical basis for this is relatively limited, being mostly based on the surprising effectiveness of tit-for-tat strategies,

there have been multiple empirical findings supporting when applied to river negotiations (Brochmann, 2012; Dinar, Katz, De Stefano, & Blankespoor, 2019; Wolf, Stahl, & Macomber, 2003).

3a: Hypothesis

However, even though treaties seem to improve subsequent relations between riparians, it has to my knowledge not been established to what extent transboundary river agreements achieve their stated purpose. This is relevant in the debate between realists and liberal institutionalists, especially when considering water allocation treaties. In this case, a specific and falsifiable prediction can be derived from the realist framework: signing a water allocation treaty will, in general, not affect the actual water flow an upstream nation permits its downstream riparian. After all, access to enough water of good quality can be considered part of the core security interest of a nation because of how crucial it is for both economic development and the health of its citizens. Water allocation treaties therefore fall in the category of “primary interests” and will not easily sway states’ intentions. Furthermore, since local treaties and institutionalization are a formalization of pre-existing local power relations, the treaty will not produce meaningful change: it will only commit to paper a previously unspoken status quo. Therefore, my research will test the hypothesis that water flows to downstream riparians are changed by the presence of a water allocation treaty.

3b: Data

I have catalogued treaties from the Transboundary Freshwater Dispute Database (TFDD), which is part of the Basins At Risk (BAR) project. This aims to systematically collect data on events related to cooperation and conflict on water use between countries in order to learn more about the relationships between conflict and freshwater resources. Part of this project is a database on international freshwater agreements from 1820 to 2007. I have selected all treaties from this database that concern water quantity, excluding entries such as re-affirmation (though not revision) of older treaties and minutes of negotiation meetings.

Then, I coupled these treaties to entries from the Global Runoff Data Centre (GRDC), which collects the data from a large collection of discharge measurement stations across the globe. While in many cases it will not be possible to find the exact border crossing flow, I can use the data from the closest measuring station to the relevant border. In almost all cases, the measuring station was situated in the downstream country, with some exceptions made for instances where there was no station downstream of the border but one very close (less than 5 kilometers) upstream. I requested data from five years before and five years after the signing of the treaty, to capture as many of the surrounding years as possible while not having to discard too many treaties due to non-availability of stations. Some of the stations have daily data available, some have monthly data, and some have both. I extrapolate both to yearly flow, because often the exact date a treaty goes into force is not available (only the date of signing) and because a measurement at the yearly level will be less affected by seasonality.

I also requested flow data for a “control group” of sorts, to increase the accuracy of the precipitation-based benchmark model. These were selected through selecting all stations in the same World Meteorological Organization region, excluding those in the same catchment area, selecting those that had high-quality coverage for the same time period as its corresponding “treatment” station, and then selecting the station that has a long-term average discharge closest to its counterpart.

I also use precipitation data, since precipitation is one of the most important determinants of water flow at a specific point. I get precipitation data from the Centre for Environmental Data Analysis (CEDA). They provide the Climate Research Unit Time Series database (CRU-TS), of which I use the 4.03 version. This database is gridded, with each 0.5-degree by 0.5-degree square having data on several meteorological variables, of which precipitation is by far the most available. I download shapefiles from www.HydroSHEDS.com, which is a project that uses satellite imaging of altitude profiles to construct a watershed allocation for almost all the landmass on earth (Lehner & Grill, 2013). I select the ones in my treatment and control groups and overlay them on the CRU-TS grid. I kept all the fields that had some level of overlap with the catchments, rather than excluding all those that partially fell outside of them. This does have the drawback of using the entirety of the catchment, where ideally only the parts of the catchments upstream from the measuring stations would be used. The GRDC does have a selection of such maps available, but only for 35 of the 52 rivers in my extended sample. Therefore, in the interest of consistency and due to lack of expertise to manually draw in the necessary subdelineations, I opted to use the HydroSHEDS basin definitions for the entire sample. In some high-latitude regions, especially where the river in question is small and located in an otherwise flat and forested area, the HydroSHEDS basin detection program can have trouble outlining the relevant basin. In these cases, since the river is small anyway, I instead use the precipitation that has been imputed for the raster block in which the measuring station is present.

I received the GRDC discharge data in either a monthly or daily format. After picking the one with the lowest proportion of missing values for the period surrounding the treaty dates, I used the `imputeTS` package in R to fill in missing values by removing seasonality, applying linear interpolation, and adding the seasonality back on (Moritz & Bartz-Beielstein, 2017). After that, I averaged the discharge for each year, since it is common to still use cubic meters per second even when talking about yearly discharge. Similarly, I average the precipitation numbers both spatially and temporally, since there is precedent for both.

3c: Model

The relevant question is not whether allocation treaties are followed to their letter: as McCaffrey (2003) points out, it can be very difficult to accommodate the variety of possible future world states properly, not to mention the fact that it is difficult to draw the line between disregarding a treaty and observing it good faith while being constrained by external factors (Chayes & Chayes, 1993). On top of that, while treaties with more explicitly detailed sharing, measurement and compliance mechanisms may be more easily enforced, they also fail more easily if complicating circumstances arise (Adger, Arnell, & Tompkins, 2005). Keeping an agreement intentionally vague may serve as a way to give future diplomats leeway to devise a solution better tailored to the future situation, although this can backfire in the situation where either of the parties does not have the political will to enter into renegotiations (Fischhendler, 2003). Instead, in the spirit of placing this research in the realist-institutionalist debate, I will only try to estimate a change in state behavior compared to before the agreement. The way I implement this into my measurement is to measure the deviation from the pre-treaty “normal” flow before and after the treaty. This also makes it so I can ignore any specific amounts that countries share. This is helpful with the analysis, since different treaties use different definitions (some treaties use percentage sharing, some use a minimum amount of water to be allocated to the downstream party, while others use a water depth at the border that needs to be maintained).

However, even though the data is obtainable, a good identification strategy is hard to come by. Since decision-making capacity is a scarce good for states, they will only “invest” it in areas where there is already a high chance of changing behavior. This leads to strong selection bias, and the large number of hard-to-quantify ways in which states (and riparian pairs) differ will make it extremely difficult to construct a valid control group. I have also considered using an instrumental variable, one candidate for which is a democracy score according to the Polity IV database, which has been widely used both as a variable of interest and a control variable (Zawahri, Dinar, & Nigatu, 2014; Dinar, De Stefano, Nigatu, & Zawahri, 2019; Dinar, Katz, De Stefano, & Blankespoor, 2018; Wolf, Stahl, & Macomber, 2003). However, the evidence on the effect this has on treaty formation is often indirect (e.g. Brochmann (2012)) and sometimes conflicting (e.g. Dinar et al. (2019)), making this a less than ideal candidate. Furthermore, my sample turned out to be relatively restricted, making it less likely to get meaningful results, especially since instrumental variable estimation has larger standard errors, which will have an especially debilitating effect in my small sample. In essence, these challenges are similar to those encountered in event studies, which have a wide application in analyzing stock price movements (Kothari & Warner, 2009).

In the financial literature, this is most often done by defining some measure of “abnormal returns”, or distance from a benchmark model that has been estimated over a different timeframe (the “estimation window”) than the periods directly preceding and following the treatment (the “event window”) (Kothari & Warner, 2009). The benchmark can be estimated in different ways: for stock prices different methods include the average price over the estimation period, the market index or a prediction of price based on the Fama-French three- or five-factor models, which give an estimation of the value of a stock based on certain structural determinants. I try to implement two versions of such tests, based on two different benchmarks: a mean-trend framework and a baseline model derived from yearly average precipitation.

The first method I use is to compute the average flow in the five years before the treaty, and for each year, calculate deviation from this average for each year. To dampen the relative impact of unexpectedly large deviations in rivers with large discharges, I use percentual deviation, and since I do not have an a priori reason to suspect the effect will work in a specific direction, I have taken the absolute value of the deviation. After all, if the upstream riparian is the one giving concessions in the negotiation process, one might expect the effect on cross-border flow (if there indeed is one) increase, while it would decrease if it is the downstream country giving up more ground. Since it is virtually impossible to glean this information from the treaty text and other descriptions of agreement, our best bet is to allow the effect to go either way. This will also prevent differently signed effects in different river systems from cancelling each other out. I control for precipitation, making the model look like this:

$$y_{it} = \hat{\beta}_0 + \hat{\beta}_1 t_i + \hat{\beta}_2 p_{it} + \hat{\beta}_3 z_i + \varepsilon$$

Here, t_i is our binary treatment variable, which is split into two specifications. Firstly, I use a basic treatment dummy stating whether a treaty has been signed or not. Secondly, I insert battery of relative time fixed effects, indicating time effects relative to $t-1$, the last year before the treaty is signed. I use an either/or approach, since in event studies, using all of a treatment indicator and two-way fixed effects at the same time leads to perfect multicollinearity, even when omitting reference categories (Borusyak & Jaravel, 2017). The first specification is identical to a t-test comparing the samples before and after the event and has more power than an evaluation of each year. The second test, however, might allow us to better notice effects of the treaty that inconsistent over time, such as an especially strong deviation in

the first years, or compliance to the treaty only picking up in later years after an implementation period (although it should be noted that longer-term effects are discounted in a static canonical regression like this (Borusyak & Jaravel, 2017)). y_{it} is my outcome variable, the absolute percentual deviation from the pre-treaty average:

$$y_{it} = \frac{|f_{it} - \overline{f_{i,pre}}|}{\overline{f_{i,pre}}}$$

z_i is a collection of fixed effects for the river involved. It could be argued that it would be better to assign the unit fixed effects based on the country pair the border runs between, since that more effectively captures idiosyncrasies of different countries, and of the relations between them. However, in cases where one pair of countries has different treaties on different river basins, both of which are in my sample, it does not catch the more natural differences per basin, such as to what extent precipitation is absorbed into the groundwater stock, how much water is “lost” to evapotranspiration (the CRU-TS has some data available on this, but not enough to feasibly account for my entire sample) and characteristics of the river that make it more or less difficult to constrain the flow to downstream areas. On the other hand, using river fixed effects only accounts for country features through geography, and most likely doesn’t capture the quality of relations at all. It also misses the countries in the case of a single pair signing agreements on multiple river basins. Since this latter case is more prevalent than the case where one river is treated on by multiple country pairs (the only instance of that is for the Orange River, with both treaties between South Africa and Lesotho and between South Africa and Namibia), I have decided to implement the river fixed effects.

I employ effects for relative time rather than absolute time since the treaties are spread out over the better part of a century, and therefore implementing absolute time would introduce so many different years to control for that I would start to encroach on estimating each year individually.

The other method uses a different benchmark model, where I instead measure a deviation from the value that would be predicted based on precipitation. I achieve this using a three-step process. In the first step, I regress flow on precipitation on all periods that I have data on, and for all periods for the rivers in my dataset without treaty. This corresponds to the simple equation:

$$f_{it} = \hat{\beta}_0 + \hat{\beta}_1 p_{it} + \hat{\beta}_2 z_i + \varepsilon$$

Where f_{it} is the observed flow at time t and station i and p_{it} the precipitation. I control for river fixed effects. Then, using the estimated coefficients obtained in the first step, I compute:

$$\hat{f}_{it}(p_{it}) = \hat{\beta}_0 + \hat{\beta}_1 p_{it} + \hat{\beta}_2 z_i$$

To obtain an estimation of the water flow based only on the precipitation for that year. Finally, I use the observed deviation from this benchmark as my dependent variable:

$$y_{it} = \frac{|f_{it} - \hat{f}_{it}(p_{it})|}{\hat{f}_{it}(p_{it})}$$

Which I incorporate into my final estimation:

$$y_{it} = \hat{\beta}_0 + \hat{\beta}_1 t_i + \hat{\beta}_2 z_i + \varepsilon$$

Where t_i is again the treatment variable specified both as a dummy and as a year indicator, and z_i the collection of river fixed effects. This estimation strategy allows me to more closely approximate the variations in water flow that result from human actions. It also has the benefit of a larger estimation window, which also includes different, not-treated rivers as well, granting the freedom of sample size expansion while not directly being foiled by the plausibly present selection biases.

A large empirical literature that employs event studies to understand stock price fluctuations already exists. Kothari and Warner (2009) write that in just the leading five journals in financial studies, the number of published event studies exceeded 550 as of 2000, with presumably thousands of other studies being published since then and in other journals. In other fields of economics, they are often used as an enhancement of difference-in-differences estimators (Borusyak & Jaravel, 2017; Słoczyński, 2018) or to capture dynamic treatment effects (Dobkin, Finkelstein, Kluender, & Notowidigdo, 2018). According to Corrado (2010), they are also often used in legal proceedings involving allegations of insider trading. Nevertheless, the methodology is still evolving, and there is relatively limited information of formal identification conditions for more simple cases like my study. Abraham and Sun (2019) identify three assumptions: parallel trends in baseline outcomes, lack of anticipatory action and treatment effect homogeneity.

The parallel trend assumption is very hard to test for in a difference-in-differences setting; in an event study it is impossible. This is because there is no control group trend that can be compared to the treatment group trend in a separate window: all sample entries are in the treatment group. However, a substitute for this would be assuming that there is no trend at all before the event time. This can be tested using the Mann-Kendall test, which has a null hypothesis of an independently and identically distributed population, and an alternative hypothesis of the data following some monotonic trend. This assumption is doubly important, since my independent variable is fully a function of time, either implicitly or explicitly. Therefore, the presence of a trend could introduce correlation between the regressor and the error term, violating the more fundamental exogeneity assumption. In the mean-trend model, there is no theoretical basis to expect a specific trend *a priori*, but in the precipitation benchmark model one could make the argument that as entries in the later 20th century are considered, it becomes more feasible for the states involved to artificially alter river flows, as technology has improved over time and financial backing for such projects has become more readily available. However, the structuring of the time indicators as fixed effects for relative rather than absolute year will likely ameliorate this, since the time between event windows is generally larger than the time within them.

Sadly, the lack of anticipatory behavior cannot be formally established. It is rather obvious that the signatory states have prior knowledge of the agreement, but it can instead be argued that this assumption is less stringent than in traditional applications of event studies. When evaluating stock prices, even the slightest hint of information leakage can invalidate any result due to the mass of arbitrageurs, and in studies on, for example, the marginal propensity to consume, it is also important that the consumers do not expect their income shocks. However, there are no meaningful opportunities to arbitrage in river systems negotiations, and there is no clear direction an anticipatory action would take. In some situations, a country might allow for more water to flow downstream in order to try to signal their cooperative intentions, while in others they might constrain the flow rate more than usual in order to put pressure on the other parties. This combined with the fact that I measure absolute percentual deviation and look at flows up to five years in advance of the treaty in the worst case, will minimize the issues coming from the non-fulfillment of this assumption. Furthermore, even in some of

the cases where the assumption is violated, this fact can be interpreted as a useful result: if an upstream state restricts flow in order to force an agreement, this will show up as a lack of deviation due to the treaty. That would still be exactly what I am looking for, since the true situation is indeed that the treaty formalizes the existing status quo.

Treatment effect heterogeneity is a less stringent assumption, only making the results slightly harder to interpret (Abraham & Sun, 2019). I account for potential heteroskedasticity by testing the significance of my model using White standard errors.

4a: Results

Something stands out from the geographic distribution of my selected rivers and stations is that no entries from South or East Asia are present in my sample. This is not to say there is a lack of major rivers there, or that they lack importance, or that no treaties have been signed governing them. Rather, it is due to the fact that the discharge measuring stations in that area typically do not lend themselves for my research question. On the Indian subcontinent, the stations along the Ganges, Brahmaputra and Indus rivers do not cover time periods overlapping with the treaties signed: they typically start operating several years after the signing of a treaty. South-East Asia similarly has immensely important rivers in the Mekong Irrawaddy systems, which many treaties have been established over. However, in this region the relevant measuring stations are more centrally located, dozens if not hundreds of kilometers away from the nearest border. Finally, the major rivers in East Asia do not often cross borders, and when they do, they are typically not important enough to warrant specific treaties.

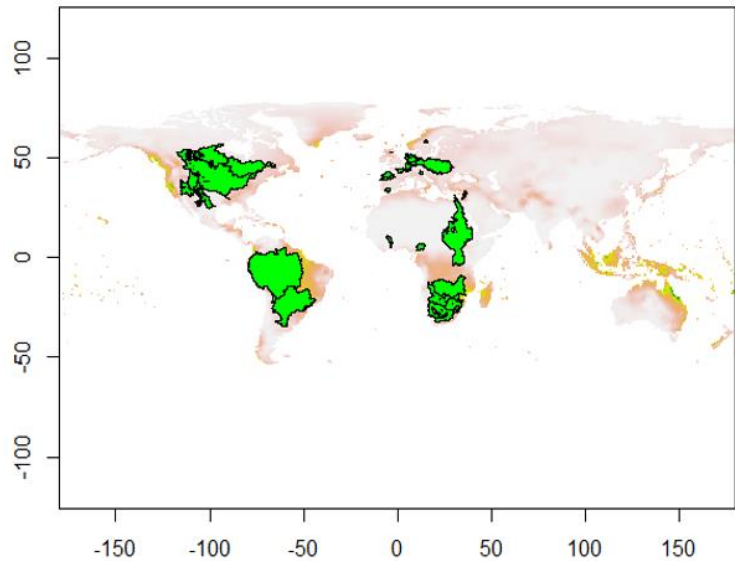


Figure 1: A world map of the precipitation data, with the outlines of my used basins overlaid

I ran into some difficulties with the implementation of my model. Firstly, more precipitation data was missing than expected, which led to a further reduction of my already small sample size, leaving only 26 treated entries. Another problem I ran into was that some rivers had multiple stations. Since I used river-level fixed effects, and there was some overlap in time between the corresponding two measuring periods. This caused the treatment variable to be decomposed into a linear combination of the fixed effects, causing uninterpretable results due to perfect multicollinearity. The two rivers causing this problem were the Orange (A treaty between South Africa and Lesotho plus a treaty between South Africa and Namibia) and Okavango rivers (A treaty between Botswana and Namibia and one between Namibia and Angola). If the problem had been caused by only a single river, I could have solved it by setting that river as the reference category, but in this case, I had to adapt my analysis to it more thoroughly. In the end, especially since the sample size was quite small, I could not keep the separate binary “treaty” indicator and had to resign to using just the relative year fixed effects.

The Mann-Kendall test on the precipitation benchmark sample returned a p-value of 0.23, and an estimated coefficient of 0.03. Even if the trend would be stronger than picked up by the test, this would

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.0235639	1.0800594	2.7994	0.005382	**
precipitation	-0.0076077	0.0058497	-1.3005	0.194214	
relativeyear0	1.3491653	0.5263384	2.5633	0.010754	*
relativeyear1	0.0463520	0.3715645	0.1247	0.900789	
relativeyear2	0.8974184	0.8232308	1.0901	0.276356	
relativeyear3	0.3017923	0.5144458	0.5866	0.557799	
relativeyear4	0.5335796	0.4399193	1.2129	0.225924	
relativeyear5	1.1975974	0.8638201	1.3864	0.166443	

Table 1: Deviation from pre-treaty average, controlling for precipitation

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-0.065640	0.249479	-0.2631	0.7925263	
relativeyearminus5	-0.121694	0.853592	-0.1426	0.8866628	
relativeyearminus4	-1.234399	0.857235	-1.4400	0.1502082	
relativeyearminus3	0.443634	0.211781	2.0948	0.0364588	*
relativeyearminus2	-0.778736	1.203720	-0.6469	0.5178285	
relativeyear0	0.081490	0.435084	0.1873	0.8514679	
relativeyear1	-0.610893	0.444899	-1.3731	0.1700489	
relativeyear2	-0.855919	0.644100	-1.3289	0.1842175	
relativeyear3	0.419316	0.261328	1.6046	0.1089282	
relativeyear4	0.356870	0.284961	1.2523	0.2107566	
relativeyear5	0.356496	0.299343	1.1909	0.2339830	

Table 2: Precipitation Benchmark results

indicate that there is a slight upward trend in deviation from the discharge predicted by the precipitation values. This would be directly opposite to the more plausible effect of the deviation decreasing over time due to technological improvements, as detailed in the assumptions section.

The main tests seem to suggest no meaningful change around the treaty year. While the mean trend analysis shows an increase in the deviation that is significant at the 0.05 level, it is not at all backed up by the precipitation benchmark analysis for the same period. Furthermore, the effect disappears in all other periods, and the two analyses frequently yield opposite-signed coefficients. Interestingly, the precipitation control in the mean trend analysis is not significant on its own. When taken at face value, the effect of signing a water sharing agreement being pronounced in the first year and then abating in later years would not be an implausible result. However, given the fact that the year-by-year evaluation of the strength of the deviation creates some multiple-testing problems, I do not think that the 0.05 significance threshold is strict enough to confidently say the effect exists.

4b: Discussion

Overall, water sharing treaties do not seem to meaningfully affect state actions. Although this is in line with the testable predictions of realist theory, it does not imply that such agreements are mere “paper tigers” whose only value is making the geopolitical status quo in the region more easily visible to outside

observers by committing it to a piece of paper. Treaties can still be a valuable tool to keep dialogue between countries open, and to use as a steppingstone for other cooperation. This paper does not in itself contradict the findings supporting the role of water treaties as a tool against militarization of conflicts (Wolf, Stahl, & Macomber, 2003) or their role in facilitating future cooperation (Zawahri, Dinar, & Nigatu, 2014). Their role in creating more goodwill between states might indeed be the most important role these agreements play, especially when considering the effects of climate change. Since climate change will make the availability of river freshwater increasingly volatile, flexibility in adapting to extreme situations is more valuable than strong agreements on allocation (Adger, Arnell, & Tompkins, 2005). This is a role that is better filled by water sharing treaties if they consist of loosely binding cooperation channels rather than strict sharing rules (Fischhendler, 2003).

Further research can be done on investigating the role that water treaties play. If their role indeed is to provide a reference point for cooperation, they would ease the impact of unexpectedly harsh circumstances. This could be empirically tested by comparing the effect of a severe drought on, for example, local GDP figures for basins with and without a treaty. The role of treaties as enabling devices for future cooperation can also be more closely examined. One could do this by for example, trying to estimate the effect of a cooperative treaty over water on the likelihood of treaties such as demilitarization deals or free trade agreements in the following years. Brochmann (2012) carried out an analysis resembling this, but only estimated the likelihood of cooperative events on the BAR scales as a function of the number of water treaties present.

This study can also be extended. Obvious respecifications include altering the relative dating. For example, instead of taking the signing of the treaty as the event date, one could use the date a treaty goes into effect. Another potentially interesting option is looking at the effect of *entering* treaty negotiations rather than concluding them. This would shed light on the extent to which upstream states use the water they allow to the downstream states as a negotiation tool. They could either constrain the flow to put pressure on the downstream state or allow more to pass through as a sign of good intentions. This could potentially be extended to negotiations on other issues as well. If the right data can be acquired, it would also be interesting to perform a more traditional difference-in-differences analysis comparing the water flow at the border between states that entered negotiations and ended up signing a treaty, and states that started negotiations but broke them off before reaching an agreement. This would somewhat ameliorate the selection bias issues that made such an analysis unfeasible in this paper. Finally, an interesting study would be whether the quantity of water permitted to the downstream country changes if a party with a particularly hawkish (or dovish) foreign policy unexpectedly comes to power in the upstream country.

5: Conclusion

I have not found any evidence supporting the hypothesis that water-sharing treaties have any effect on the actions of the countries signing them. A small effect was found in the specific year the treaty was signed, but this disappeared with an alternate model specification. This has implications for the way freshwater disputes are analyzed, since scholars typically implicitly assume an institutionalist starting point, while this paper posits evidence indicating that a more realist framework could be more appropriate. However, further research is needed to determine whether this is indeed the case, and to better understand what the role of water sharing treaties is if they do not alter allocation.

Bibliography

- Abraham, S., & Sun, L. (2019). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. doi:<http://dx.doi.org/10.2139/ssrn.3158747>
- Adger, W. N., Arnell, N. W., & Tompkins, E. L. (2005). Successful adaptation to climate change across scales. *Global Environmental Change*, 77-86.
- Axelrod, R., & Keohane, R. O. (1985). Achieving Cooperation under Anarchy: Strategies and Institutions. *World Politics*, 226-254.
- Borusyak, K., & Jaravel, X. (2017, May 8). Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume. doi:<http://dx.doi.org/10.2139/ssrn.2826228>
- Brochmann, M. (2012). Signing River Treaties- Does It Improve River Cooperation? *International Interactions*, 141-163.
- Chayes, A., & Chayes, A. H. (1993). On compliance. *International Organization*, 175-205.
- Cooley, H., & Gleick, P. H. (2011). Climate-proofing transboundary water agreements. *Hydrological Sciences Journal*, 711-718.
- Corrado, J. C. (2010). Event studies: a methodology review. *Accounting & Finance*, 207-234.
- Dinar, A., De Stefano, L., Nigatu, G., & Zawahri, N. (2019). Why are there so few basin-wide treaties? Economics and politics of coalition formation in multilateral international river basins. *Water International*, 463-485.
- Dinar, S., Katz, D., De Stefano, L., & Blankespoor, B. (2018). Do treaties matter? Climate change, water variability, and cooperation along transboundary river basins. *Political Geography*, 162-172.
- Dobkin, C., Finkelstein, A., Kluender, R., & Notowidigdo, M. J. (2018). The Economic Consequences of Hospital Admissions. *American Economic Review*, 308-352.
- Eleftheriadou, E., & Mylopoulos, Y. (2008). Game Theoretical Approach to Conflict Resolution in Transboundary Water Resources Management. *Journal of Water Resources Planning and Management*, 466-474.
- Fischhendler, I. (2003). Legal and institutional adaptation to climate uncertainty: a study of international rivers. *Water Policy*, 281-302.
- Global Runoff Data Center. (2019). River Discharge Data. 56068 Koblenz, Rhineland-Palatinate, Germany: GRDC in the Bundesanstalt fuer Gewaesserkunde.
- Grieco, J. M. (1988). Anarchy and the limits of cooperation: a realist critique of the newest liberal institutionalism. *International Organization*, 485-507.
- Haas, E. B. (1980). Why Collaborate? Issue-Linkage and International Regimes. *World Politics*, 357-405.
- Hamner, J., & Wolf, A. (1998). Patterns in International Water Resource Treaties: The Transboundary Freshwater Dispute Database. *Colorado Journal International Environmental Law and Policy*.

- Kothari, S. P., & Warner, J. B. (2009). Econometrics of Event Studies. In E. Eckbo (Ed.), *Handbook of Empirical Corporate Finance* (pp. 3-36). North Holland: Elsevier.
doi:<https://doi.org/10.1016/B978-0-444-53265-7.50015-9>
- Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, 2171-2186.
- Madani, K. (2010). Game theory and water resources. *Journal of Hydrology*, 225-238.
- Mearsheimer, J. J. (1994). The False Promise of International Institutions. *International Security*, 5-49.
- Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS. *imputeTS: Time Series Missing Value Imputation in R*. The R Journal.
- Sloczynski, T. (2018). A General Weighted Average Representation of the Ordinary and Two-Stage Least Square Estimands. doi:1810.01576
- University of East Anglia Climate Research Unit; Harris, I. C.; Jones, P. D. (2019). CRU TS4.03: Climatic Research Unit (CRU) Time-Series (TS) version 4.03 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2018).
doi:<http://dx.doi.org/10.5285/10d3e3640f004c578403419aac167d82>
- Wolf, A., Stahl, K., & Macomber, M. F. (2003). Conflict and cooperation within international river basins: the importance of institutional capacity. *Water Resources Update*.
- Zawahri, N. A., Dinar, A., & Nigatu, G. (2014). Governing international freshwater resources: an analysis of treaty design. *International Environmental Agreements*, 307-331.

Appendix: list of stations used in the treatment group.

Countries	Date	GRDC station number	River Name	Station Name
Mexico, USA	1906-05-21	4151800	Rio Grande	Laredo, TX
Canada, USA	1909-01-11	4143300	St. Lawrence River	Ogdensburg, NY
Canada, USA	1938-04-09	4213210	Rainy River	Fort Frances
Canada, USA	1941-10-09	4143550	St. Lawrence River	Cornwall (ONT)
Argentina, Paraguay	1945-06-01	3268500	Pilcomay Rio	La Paz
Canada, USA	1950-02-27	4236010	Niagara River	Queenston
FRG, Austria	1950-10-16	6342928	Isar	Mittenwald
Norway, Finland	1951-04-25	6730330	Neidenelva	Neiden
Italy, Yugoslavia	1957-07-18	6549100	Soca	Solkani
Egypt, Sudan	1959-11-08	1662100	Nile River	Dongola
Belgium, Netherlands	1961-02-24	6421500	Maas	Borgharen
France, Spain	1963-07-29	6125730	Maudan	Fos
Spain, Portugal	1964-07-16	6212420	Douro Rio	Puente Pino
FRG, Switzerland	1966-04-30	6935500	Rhine	Diepoldsau, Rietbruecke
Mexico, USA	1972-07-14	4352100	Colorado Rio	Lindero Internacional Norte
Brazil, Paraguay, Argentina	1979-10-19	3264500	Parana Rio	Posadas
South Africa, Mozambique, Swaziland	1983-02-17	1896502	Limpopo River	Combomune
South Africa, Mozambique, Swaziland	1983-02-18	1898501	Umbeluzi Rio	Goba Montante
South Africa, Mozambique, Swaziland	1983-02-19	1897501	Incomati Rio	Ressano Garcia C.F.
South Africa, Mozambique, Swaziland	1983-02-20	1599100	Maputo	Siphofaheni
South Africa, Lesotho	1986-10-24	1159651	Orange River	Oranjedraai
Canada, USA	1989-10-24	4213594	Souris River	Near Westhope, ND
Finland, USSR	1989-10-26	6855400	Vuoksi	Tainionkoski
South Africa, Namibia	1992-09-14	1159100	Orange River	Vioolsdrif
Namibia, Botswana, Angola	1994-09-16	1357100	Okavango River	Mohembo/Mtaembo
Israel, Jordan	1995-09-28	6594080	Jordan River	Agan Naharayim