

RESEARCH ARTICLE

Suspicious Collaborators: How Governments in Polycentric Systems Monitor Behavior and Enforce Public Good Provision Rules Against One Another

Jeffrey Hanlon¹, Tomás Olivier² and Edella Schlager³¹ Whittier College, US² University of California Los Angeles, US³ University of Arizona, USCorresponding author: Jeffrey Hanlon (jhanlon@whittier.edu)

Monitoring and enforcement have been recognized as keys for sustainable common pool resource governance. With a couple of notable exceptions, however, scholars have not examined how they are deployed when governments are the primary actors devising such agreements and where multiple public goods are provided for – an important level of governance to understand. We explore the design of monitoring and enforcement safeguards that governments adopt to limit opportunism and support compliance in a complex governing arrangement, the New York City Watersheds Memorandum of Agreement. The agreement defines how New York City and a group of watershed jurisdictions jointly manage a shared natural resource. Furthermore, we test how the design of such safeguards vary depending on the type of public good they cover, illuminating how “federal” safeguards may work at the sub-state level, and, ultimately, the particular form of polycentric governance being used. The results indicate that concerns for water quality as well as potential for opportunistic behavior drive institutional design considerations. Monitoring and sanctioning authority for water quality is dominated by state and federal actors, which hold New York City to account, while watershed jurisdictions are held responsible by regional actors for administration of economic development goods.

Keywords: watersheds; polycentricity; safeguards; formal institutions

Introduction

Common pool resource (CPR) scholarship has consistently recognized the importance of monitoring resource characteristics and resource user actions to encourage compliance with institutional arrangements (Ostrom 1990, 1999, 2005; Gibson et al., 2005; Chhatre and Agrawal, 2008). Ostrom (1990) alerted scholars to the importance of monitoring systems characterized by monitors accountable to resource users; graduated sanctioning for rule violators; and low cost conflict resolution mechanisms to settle compliance disputes. Monitoring of resource system performance and rule following behavior of actors provides valuable information, that if appropriately linked with conflict resolution, enforcement, and rule change mechanisms support rule compliance and rule adoption (Schoon & Cox, 2012). Early empirical work, from meta-case analysis (Blomquist 1992; Tang 1992; Schlager, 1994), to lab experiments (Ostrom, Walker, and Gardner, 1992), to large-n field studies (Lam, 1998) provided support for the design principles working in this complementary fashion to produce robust resource governance (Schlager, 2004). In particular, it was not just the presence or absence of such mechanisms that mattered; rather it was the form that the mechanisms took and the types of actors in the role of monitors that affected performance.

CPR governance dilemmas remain popular topics of research. Scholars continue to produce studies identifying the design principles in action (e.g. Quinn et al., 2007; Villamayor-Tomas et al., 2014), exploring how the design principles may be scaled up to explain intergovernmental cooperation (e.g. Heikkila et al., 2011; Dietz et al., 2003), characterizing the conditions in which design principles may emerge in self-governing arrangements (e.g. McCay, 2002; Coleman & Steed, 2009), theorizing about how differences

in the implementation of design principles lead to various resource and user outcomes (e.g. Chhatre and Agrawal, 2008; Anderies et al., 2004), and exploring why self-governing arrangements of irrigation systems, forests, and coastal fisheries exhibit different combinations of design principles (Baggio et al., 2016).

More recently, scholars have analyzed collective action dilemmas among governments, as opposed to individual resource users. Fleischman et al., (2014) find that clearly defined boundaries and monitoring of resource conditions may consistently play a role in governance “success.” Analyses of interstate agreements for species recovery (Schlager and Blomquist, 2008), river basin treaties and organizations (Wolf, Yoffe, and Giordano, 2003) and regional treaties governing transnational common pool resources (Breitmeier, Young, and Zurn, 2006) have explored how CPR design principles may be scaled up to the level of intergovernmental agreements. Heikkila, Schlager, and Davis (2011) found that interstate river compacts, in which collaborators are “suspicious” of each other, contain many cross-scale (governments and resource users at different scales, i.e. state and federal) monitoring linkages designed in their formal institutions. Different layers of governments and resource users monitored the resource and by extension, each other.

We draw on federalism theories to develop hypotheses about the designs of monitoring, sanctioning, and conflict resolution mechanisms, or what some federalism scholars label “safeguards” in intergovernmental governing arrangements (Filippov, Ordeshook, and Shvetsova 2004, Bednar 2009). We hypothesize that the design of these safeguards, that is, who is monitoring and monitored, who triggers review processes and who is reviewed, and the severity of sanctions, will vary by the types of public goods provided for in intergovernmental arrangements. The hypotheses are tested using data coded from the almost four thousand rules that constitute the New York City Watersheds governing arrangement. We find that the patterns of safeguards and their design vary in systematic ways. In particular, local governments who created the intergovernmental task specific jurisdiction are granted the authority to monitor, review, and sanction the actions of the task specific jurisdiction agencies, leaving the monitoring, reviewing, and sanctioning of one another to the task specific jurisdiction. These findings contribute to both common pool resource and federalism literatures, by exploring the institutional designs of specific types of safeguards and how they vary, and also, exploring the roles played by intergovernmental arrangements in a federal system.

The following two sections examine the literature on regional governance, the importance of safeguards in addressing collective action dilemmas, and how the design of safeguards is likely to vary depending on the types of collective goods provided for, concluding with theoretical expectations. The empirical setting is then introduced and used to operationalize the theoretical expectations into a series of hypotheses that reflect the context of the case. This is followed by a methods section explaining the data, and then results and discussion of empirical tests. The paper concludes with limitations and topics for future research.

Regional Governance and the Importance of Safeguards

Many problems and opportunities extend across jurisdictional boundaries (Ostrom et al., 1961; Feiock 2013), whether it is rivers and streams that cross multiple boundaries, the depositing of pollutants into shared airsheds, or the possibility of developing a more efficient transportation system serving a metropolitan region. In many cases, addressing problems and opportunities that spill over jurisdictional boundaries may be resolved via creation of regional governing arrangements (Oakerson 1999; Feiock and Scholz 2010; Feiock 2013), or what Hooghe and Marks (2003) label task specific jurisdictions. Task specific jurisdictions are specialized to address particular policy problems, and support cooperation and coordination among “constituencies who share some geographical or functional space and who have a common need for collective decision making” (Hooghe and Marks, 2003: 240). Their creators and members may include different forms of local governments (villages, municipalities), counties, and states, as well as diverse types of non-governmental organizations.

Governments working together to form and utilize task specific jurisdictions are likely to face similar types of collective action problems as those confronted by governments that together create federal systems, a topic that is well explored and theorized (Elazar 1987, Ostrom 2007; Lutz 1990; Filippov, Ordeshook, and Shvetsova 2004; Bednar 2009). This is so because task specific jurisdictions are created by constitutions, are often autonomous, and exercise authority independently from their members. Following Bednar (2009), governments working together to provide for collective benefits and address shared problems may act in ways that undermine the governing arrangement. To discourage such opportunistic behavior, governing arrangements are laced with different types of safeguards. Safeguards are institutional mechanisms that allow members of a federation, or a task specific jurisdiction, to engage in shared decision-making, monitor one another's behavior and the public goods provided for, review the actions of members to ensure compliance, and to sanction rule-violating behavior. Bednar (2009) points to several types of safeguards, such as decision-making venues that represent different interests and can check one another, political parties that coordinate

officials in diverse and overlapping decision venues, an independent judiciary, and popular safeguards, such as voting or social protests. Each safeguard has “sanctioning” capacity, e.g., elected officials may be voted out of office, or a court may impose a fine.

Bednar’s (2009) safeguards overlap with Ostrom’s design principles (2005), both theoretically in the role that safeguards and design principles play in supporting long term cooperation, and substantively in the form they take, such as decision making venues (design principle three), graduated sanctioning (design principle five), and conflict resolution mechanisms (design principle six). Whereas Bednar (2009) implies that safeguards have the capacity to monitor, Ostrom (2005) includes mutual monitoring as a design principle (design principle four).

Bednar (2009) and Ostrom (2005) theorize that for a federation or local level governing arrangement to be robust, it must exhibit many of the safeguards or design principles, or what Bednar (2009) labels *coverage*. In addition, at least some of the safeguards must act in a *complementary* fashion (Bednar 2009). For example, complementary safeguards provide soft and firm sanctions for transgressions of different severity, so as not to crowd out collective action. In other words, it represents a form of graduated sanctioning among governments. Finally, *redundant* safeguards provide multiple means of observing and correcting behavior that violates the rules (Bednar 2009). Working together, the safeguards make it difficult for actors to engage in opportunistic behavior.

Extending Bednar’s (2009) logic of safeguards to task specific jurisdictions suggests variations in institutional arrangements. Task specific jurisdictions typically provide for several types of public goods, only some of which are core to the mission of the jurisdiction. For example, an irrigation district may provide highly salient (core) public goods such as canal infrastructure and reservoirs as well as less salient (supporting) public goods such as public information and user conservation assistance. For the task specific jurisdiction to hold together and realize its goals, it must ensure that the salient goods are provided for, which requires maintaining cooperation and compliance of member governments. According to Bednar (2009), that means paying attention to coverage, complementarity, and redundancy. As safeguards are costly to develop and implement, investing in safeguards is unlikely to be uniform across the public goods arrangements. Safeguards are likely to differ depending on the saliency of the goods. The most salient public goods reflect the primary purpose of the task specific jurisdiction and are more likely to have incorporated more and diverse forms of safeguards than are public goods that are tangentially related to the goals of the jurisdiction.

The Design of Safeguards for Regional CPR Governance

In federal systems, member governments act opportunistically by shirking their responsibilities to the union (the national government), and by burden shifting, or failing to abide by their commitments to one another (Bednar 2009). The national government may also encroach upon the member governments, exercising its authority in ways that usurp the authority of its members (Bednar 2009). Once again, extending Bednar’s (2009) logic of safeguards to the task specific jurisdiction, these three types of opportunisms, or collective action problems, define three distinct types of interactions among the task specific jurisdiction and its members: shirking referring to the failure of member governments to follow through with their commitments to the task specific jurisdiction; burden shifting representing the failure of member governments to follow through with their commitments to one another; and encroachment referring to the task specific jurisdiction and higher-level authorities usurping the authority of its members.

Safeguards, according to Bednar (2009), are meant to address and mitigate the opportunisms. Thus, we expect that the design of safeguards will vary by the type of opportunism they are intended to address. Safeguards to address shirking should be characterized by a task specific jurisdiction monitoring, reviewing, and enforcing its members’ activities. Safeguards addressing burden shifting should direct member governments to monitor and review each other’s activities. Finally, safeguards for addressing encroachment should allow member governments to hold the task specific jurisdiction to account.

We also expect that highly salient public goods arrangements will exhibit more and different types of safeguards compared to less salient public goods. To test these theoretical expectations, we first introduce the case study, before operationalizing our expectations in a series of hypotheses that reflect the context of the case.

Setting

Our primary data source is the largest watershed governance arrangement in terms of infrastructure and population served in the United States: the New York City Watersheds governing system. The governing system was brought into existence by a memorandum of agreement (MOA), which acts like a constitution. The MOA, created by New York City and the counties, towns, villages and hamlets located in the watersheds

from which the City sources its water supplies, allocates authority to support coordination and cooperation among member governments and to provide for a variety of public goods directed at protecting surface water quality and promoting environmentally-sensitive economic development.

New York City's largest municipal water supply reservoirs are located in the Catskill and Delaware watersheds nearly 100 miles north of the city. Although the reservoirs are located outside of its jurisdictional boundaries, the city enjoys considerable authority to determine land use activities near its reservoirs in order to maintain water quality and avoid the need to chemically filter the water before delivery. Following the 1989 federal Surface Water Treatment Rule, the city was responsible for filtering municipal water, which would require the construction and operation of treatment plants. Or, in lieu of that, the city could receive a filtration avoidance determination (FAD) from the EPA by demonstrating that it had a robust plan for protecting water quality at the source for the long term. The city pursued the latter and negotiated with the governments located in the watersheds a mutually beneficial land use and economic development arrangement. Key to the agreement was that watershed communities allowed the city to acquire and manage undeveloped land and keep it as such in order to protect water quality, and in return, the city funded economic development and other direct benefit programs for upstate communities and residents. The jurisdictions have a long history of non-cooperation and even animosity (Galusha, 2016; Soll, 2013) but have nevertheless produced a robust governing arrangement that appears to have appeased diverse interests.

In 1997, the New York State Department on Environmental Conservation, the New York City Department of Environmental Protection, county and municipal governments in the region encompassing the Catskill and Delaware watersheds of New York, the U.S. EPA, and a number of environmental nonprofits and regional interest groups signed the MOA. The document builds a task specific governing arrangement whereby parties create, fund and administer new projects and programs in the watersheds to, 1) protect the quality of surface waters in and entering New York City's reservoirs, and, 2) protect and improve the economies of upstate communities located in the watersheds. The U.S. EPA, in turn, issued New York City a filtration avoidance determination, recognizing the task specific jurisdiction represents a robust method of protecting water quality.

Together, the rules creating the governing arrangement, its programs, and regulations serve as a reliable source of data for testing hypotheses about patterns of safeguards and their institutional designs.

Operationalizing Research Expectations

Because watershed governance is at once about promoting and protecting a core salient ecosystem service (in this case, water quality) and about protecting against opportunism in general, we develop and empirically test hypotheses about how the formal arrangement may be shaped to those two ends. The primary goal of the arrangement is to preserve New York City's Filtration Avoidance Determination (FAD), which is dependent on objective measures of water quality (USEPA, 2017). If New York City were to lose its FAD, the regional governing arrangement would very likely end as the city would devote its resources to constructing multi-billion dollar chemical treatment infrastructure. Since protecting water quality requires extraordinary measures to prevent human impact on the water bodies in the area, communities were concerned that this would affect their economic vitality (Soll, 2013). As a result, the MOA also includes a series of programs to foster economic development in the watersheds, but in an environmentally sensitive fashion (MOA, Article V). As the goal of the agreement is to protect and produce high quality drinking water without chemical filtering, we expect that the water quality public goods programs (most salient) will exhibit more safeguards than economic development public goods programs (less salient). Likewise, water quality public goods arrangements will provide a greater complementarity of punishments (soft and firm) for rule violations than economic development public goods, and will produce more redundant monitoring, reviewing, and decision-making relationships.

Water quality hypotheses:

H1, coverage: *Water quality public goods will exhibit more safeguards overall than economic development goods.*

H2a, complementarity: *Water quality public goods will exhibit more monitoring, review for compliance, and consequence safeguards than will economic development goods.*

H2b, complementarity of consequences: *Water quality public goods will exhibit more strict consequences (penalties will be greater for rule violations) than economic development goods; and economic development goods will exhibit more mild consequences (penalties will be mild for rule violations) than water quality goods.*

H3, redundancy: *Water quality public goods will exhibit more actors engaged in monitoring, more actors triggering review processes, and more aggregation rules requiring multiple actors to participate in decision processes, compared to economic development goods.*

The MOA defines and divides authority to promote cooperation and guard against the types of opportunism that occur in a classic federal arrangement. Consequently, we may test whether the actors who created the governing arrangement anticipated opportunism through the design of safeguards. By measuring who monitors and is monitored, and who may trigger review processes and who is the subject of review, it is possible to determine when safeguards are meant to deflect encroachment (lower level governments monitor and review higher level governments), burden-shifting (actors at the same level of authority monitoring and reviewing each other), and shirking (higher level government actors monitoring and reviewing lower levels). Furthermore, since New York City is interested in maintaining water quality, and the watershed communities are interested in protecting their economies, we can test how this differs among public goods arrangements.

The New York City Watersheds governing arrangements involve governments at several levels. The U.S. EPA is authorized to issue, monitor, and revoke filtration avoidance determinations (FADs). A FAD allows a water utility to avoid intensive and expensive treatment of water prior to delivery to customers. In New York, the State Department of Health issues New York City's FAD, in consultation with the U.S. EPA. New York State agencies (in particular the Department of Environmental Conservation) also participate in water oversight through the issuance of a water supply permit, which allows the city to draw municipal water from the source watersheds. New York City's water supply permit is conditional on the city meeting its commitments under the MOA.

The watershed governing system exhibits executive, rulemaking, and conflict resolution authorities. The Catskills Watershed Corporation (CWC), a non-profit organization created by the MOA to develop and administer public goods programs (MOA, Article IV) is in charge of executive and rulemaking functions. The parties to the MOA constitute the board of directors. The public goods programs under its direct purview include a variety of water quality and economic development public goods, such as septic system maintenance programs and a program for economic development in the area of the watersheds west of the Hudson River. The conflict resolution authority is carried out by the Watershed Protection Partnership Council (WPPC), which was created by the MOA to address and resolve differences among the parties to the MOA (MOA, Article IV). These two organizations perform the administrative functions of the task specific jurisdiction.

Finally, the city and watershed jurisdictions constitute the final level of governments. They are the jurisdictions that formed the MOA and they are the recipients of the public goods provided for in the governing arrangement. New York City is the central actor in providing for water quality public goods, from financing programs under the purview of the CWC, to engaging in land and easement acquisitions in cooperation with the watershed jurisdictions and landowners. Conversely, the watershed jurisdictions are central in providing for and benefitting from the economic development public goods.

In general, we expect that the multiple levels of government will have different predilections for opportunism, and thus will monitor and review each other in appropriately diverse ways. The federal and state levels will be most interested in making sure no jurisdiction shirks its responsibility to maintain water quality; special purpose jurisdiction entities will be most interested in ensuring that the city doesn't shirk its responsibility to maintaining water quality and that the watershed towns don't shirk their responsibilities to use economic development funds appropriately. Watershed towns will be most interested in ensuring that the city does not shift the burden of water quality protection upon them, while the city will safeguard against watershed towns shifting the burden of maintaining their economic vitality upon the city. Finally, watershed towns and the city will attempt to prevent encroachment of their authorities, especially from state and federal agencies, and less so from the CWC and WPPC as they created those two entities and participate in them.

Opportunism hypotheses:

H4a, shirking: *Federal and state agencies, which issue water quality permits, are more likely to address shirking through monitoring and reviewing safeguards targeting water quality goods.*

H4b, shirking: *The CWC and WPPC will use monitoring and reviewing safeguards to address shirking by the city in relation to water quality public goods; and will use monitoring and reviewing safeguards to address shirking by the watershed towns in relation to economic development public goods.*

H5, burden shifting: Watershed jurisdictions and the city will monitor and review each other to prevent burden shifting for providing economic development and water quality goods.

H6, encroachment: Watershed jurisdictions and the city will monitor and review state and federal agencies more than the CWC or WPPC.

Data

We test the hypotheses using measures based on *configurations of rules* that constitute the governing arrangements of the New York City watersheds. The sources of these rules are the 1997 Memorandum of Agreement, the 2014 New York City Water Supply Permit, the New York City Rules and Regulations, and the Catskill Watershed Corporation program rules. The Catskill Watershed Corporation program rules cover Economic Development, Education, Flood Hazard Mitigation Implementation, Septic Systems, Storm-water Controls, Storm-water Retrofit, Tax Litigation Avoidance, and Community Wastewater Management Program Rules. In total, 444 pages of rules were coded, containing 3,653 rules.

The type of public good arrangement is the dependent variable for the hypotheses. The public goods arrangements were identified by examining each article and subsection of each document for the formal names of programs. Individual programs were identified using section and subsection titles. In some instances, an entire document would consist of institutional statements¹ that structured a program, such as the Tax Litigation Avoidance Program; in others, a document would consist of multiple programs (e.g. Water Supply Permit rules). Public goods programs were further categorized as water quality public goods, which are programs that directly provide for or protect water quality, and economic development public goods, which provide for development activities.

The total number of public goods identified were 71. Of those 71 public goods, 35 included their own safeguards. That is, for 35 of the 71 public goods, the sets of rules creating a public good also created monitoring, compliance, and/or consequence safeguards for the particular public good. Given that the hypotheses focus on variation in safeguards across types of public goods, the 35 public goods with safeguards are the focus of this analysis (see **Table 1**).

The independent variables are the safeguards associated with each public good as well as measures of their institutional design. The three types of safeguards are: 1) monitoring, which allow actors to monitor one another's actions as well as the outputs and outcomes of public goods; 2) reviewing (for compliance), which allow actors to hold one another accountable by providing processes for questioning possible compliance issues and reviewing actions taken by actors; and 3) consequence, which are sanctions for rule violating behavior. The safeguards consist of 1 to N consecutive institutional statements.² For example, in the New York City Agricultural Land Easement public good, these three statements constitute a monitoring mechanism³: "The City will submit copies of its acquisition reports which are submitted to the Primacy Agency, pursuant to the Interim and 1997 FADs, to NYSDEC, and to the Watershed Protection and Partnership Council. Such reports will include the following information for all parcels and easements acquired during the reporting period: address; description of the property, including any easement; county and town where property is located; tax map number; acreage; closing date; and map of property. The acquisition report shall also contain cumulative totals of acreage solicited and acreage acquired identified by Town and Priority Area." (MOA Article II, 84.a). The safeguard creates the means for a state agency (New York State Department of Environmental Conservation) and a task specific jurisdiction entity (the Watershed Protection Partnership Council) to monitor the city and details the means by which they do so.⁴ We use counts of each monitoring, reviewing for compliance, and consequence safeguard by type of good to test hypotheses 1 and 2.

¹ An institutional statement is the unit of analysis in the Institutional Grammar Tool (Crawford and Ostrom 1995, 2005; Basurto et al., 2010; Siddiki et al., 2011). Crawford and Ostrom define institutional statements as "the shared linguistic constraint or opportunity that prescribes, permits, or advises actions or outcomes for actors (both individual and corporate)" (1995: 583). In practice, institutional statements commonly overlap with a sentence in the text, and usually identify a combination of the following elements: an actor who is supposed to do something, an action mandated by the rule; a specification of whether the rule mandates, allows, or forbids an action; a recipient of the action; a series of conditions under which that action should occur or not; and consequences for not abiding with the rule (Siddiki et al., 2011).

² Contact Edella Schlager at schlager@email.arizona.edu for the coding protocols.

³ Throughout the text, we use the terms "safeguard" and "mechanism" interchangeably.

⁴ To assess intercoder reliability for coding safeguards, we distributed a percentage of the statements from each coded documents between three coders. On average, each coder analyzed 70% of the institutional statements within a coded document. The average percentage agreement between all three coders was of 88.7%, with the lowest agreement rate being 73.3%. This assessment was conducted on all the documents coded except for the Storm-water Retrofit Program rules and for the Septic System rules, which were coded and discussed by the authors together.

Table 1: List of 35 public good arrangements.

Economic development	Water quality
CWC Education	Alternate Septic
Good Neighbor Payments	CWC Septic Program
Grant Programs	East of Hudson Watershed Plan
Job Creation	Enhanced Monitoring
Loan Programs	Flood Program
Local Consulting	Forest Conservation Easement
Miscellaneous Payments	MOA New Storm-water
Miscellaneous Procedures	NYC Land/Easement Acquisition
MOA Economic Development	NYC Watershed Agricultural Easement
MOA Education	Riparian Buffers Program
Qualified Economic Development	Sand and Salt Program
State Funding	Septic Program
Tax Consulting Fund	Sewerage Diversion Study
Tax Liability Assessment Program	Small Business Septic Program
	Storm-water Retro Fit
	Stream Corridor
	Total Maximum Daily Loads
	Waste Water Treatment
	Water Conservation
	West of Hudson New Storm-water
	WSP NYC Watershed Easements

Hypotheses 2a and 2b contend that the severity of consequences varies by public good type. We distinguish between severe and mild consequence safeguards. Severe consequences affect the core nature of the agreement (for example, rights to acquire land and those that impose additional restrictions on the production and provision of clean water). Mild consequences are those that affect an actor, but that do not touch the core nature of the agreement (for example, losing the right to raise an objection or being required to pay interest on a debt).⁵

Each of the institutional statements was coded using the rule typology developed by Ostrom (2005). A count of all aggregation rules, which identify actions or decisions that require two or more actors to execute (Ostrom 2005: 202) per public good is used to test hypothesis 3.⁶

Another set of independent variables consists of the types of actors who are engaged in monitoring and compliance reviews as well as actors being monitored and being reviewed. For each safeguard identified, actors that appeared in the safeguard were assigned a category based upon the level of government in which they were located: federal, state, regional, and municipal. Next, the following actor positions within a safeguard were identified: the monitor and the monitored (for monitoring safeguards); the actor triggering a compliance review action and the receiver of a compliance review action (for compliance safeguards); and the imposer of a consequence and the receiver of a consequence (for consequence safeguards). Coded this way, each safeguard may have multiple actors at different levels of government acting upon multiple other actors. The relationships may be horizontal, in that actors may hold other actors of the same level accountable (i.e. a municipal government monitoring another municipality's actions). Or, the relationships may be vertical, with

⁵ Contact Edella Schlager at schlager@email.arizona.edu for the coding protocols.

⁶ To assess intercoder reliability for the application of the Institutional Grammar Tool (IGT), the authors distributed an average of 13% of the statements in each document between three coders. For the coding of rule types (where we identified whether a statement was an aggregation rule or some other rule type), the coders agreed 77% of the times across rule sets. For the remaining IGT components, rates of agreement varied between 76% and 93%. This was done for all of the documents analyzed in this paper, except for the Storm-water Retrofit program rules, where all coding differences were discussed between the coders, and the correct coding agreed upon.

lower level actors (i.e. at the municipal level) holding higher level actors (i.e. at the federal level) accountable, or vice versa. **Table 2** displays variables created to test hypotheses one through six.⁷

Methods

We use a mixed approach to assess the safeguards and their design features associated with water quality and economic development public goods. For hypotheses 1 through 3, we rely on t-tests, which allow a direct comparison of the variable of interest between water quality and economic development goods (i.e., the mean number of safeguards and types of safeguards should be greater for water quality goods compared to economic development goods). We also conduct an additional robustness check to examine whether the variables of interest, taken together, accurately assign predicted observations to the two categories of public goods.⁸

Table 2: Key variables, organized by hypotheses.

<i>Dependent Variable</i>	
Public good type	Economic development public good (1) or water quality public good (0)
<i>Independent variables</i>	
H1: <i>Water quality public goods will exhibit more safeguards overall than economic development goods.</i>	Total number of monitoring, compliance, consequence safeguards
H2a: <i>Water quality public goods will exhibit more monitoring, review, and consequence safeguards than will economic development goods.</i>	Number of each type of safeguard (monitoring, review, consequence)
H2b: <i>Water quality goods will exhibit more strict consequences (penalties will be greater for rule violations) than economic development goods; and economic development goods will exhibit more mild consequences (penalties will be mild for rule violations) than water quality public goods.</i>	Number of each type of consequence safeguard (severe, mild)
H3: <i>Water quality public goods will exhibit more actors engaged in monitoring, more actors triggering review processes, and more aggregation rules which require multiple actors to participate in decision processes compared to economic development goods.</i>	Number of monitors, number of actors triggering compliance reviews, number of aggregation rules
H4a: <i>Federal and state agencies, which issue water quality permits, are more likely to address shirking through monitoring and reviewing safeguards targeting water quality goods.</i>	Counts of public goods in which federal and state actors act as monitors and reviewers regarding water quality and regarding economic development goods
H4b: <i>The CWC and WPPC will use monitoring and reviewing safeguards to address shirking by the city in relation to water quality public goods; and will use monitoring and reviewing safeguards to address shirking by the watershed towns in relation to economic development public goods.</i>	Counts of public goods in which CWC and WPPC act as monitors and reviewers of the city and watershed towns regarding water quality and regarding economic development goods
H5: <i>Watershed jurisdictions and the city will monitor and review each other to prevent burden shifting for providing economic development and water quality goods.</i>	Assessment of public goods in which watershed towns and New York City monitor or review each other regarding water quality and economic development public goods
H6: <i>Watershed jurisdictions and the city will monitor and review state and federal agencies more than the CWC or WPPC.</i>	Assessment of public goods in which the city and watershed jurisdictions monitor or review federal or state agencies and the CWC or WPPC

⁷ The authors conducted several rounds of intercoder reliability tests and codebook revisions to code the elements of a safeguard described here. The first round yielded a reliability of 31%. Subsequent samples were used until the entire safeguards dataset was coded, but the coders never reached an 80% agreement threshold. To attain reliable coding, all differences between coders were discussed and the correct coding agreed upon using the final version of the codebook.

⁸ Robustness analysis is available upon request from Edella Schlager at schlager@email.arizona.edu.

For hypotheses 4 through 6, we rely on a different approach. These hypotheses address whether there is a relationship or association between types of actors engaged in monitoring and reviewing, actors being monitored and reviewed, and types of goods. Given the nature of the variables of interest and the small number of observations within them, we use Fisher's exact test for hypothesis 4, while descriptions of public goods are used to address hypotheses 5 and 6.

Results

For hypotheses 1 through 3, **Table 3** lists t-test coefficients for the variables of interest, comparing water quality and economic development goods. For hypothesis 1, the differences in the average number of safeguards between the two types of goods is in the expected direction, however, that difference is not statistically significant at the .10 level. Hypothesis 2a is partially supported. By breaking out the safeguards into types, we observe a statistically significant difference in the number of monitoring safeguards. Water quality goods are more heavily monitored than economic development goods, as hypothesized. Regarding the other two safeguard types, while water quality public goods have more consequence and reviewing safeguards, the difference in their means is not statistically significant.

T-test results also provide partial support for hypothesis 2b. As expected, economic development goods contain a higher average of mild consequences than water quality public goods, and this difference is statistically significant at the .05 level. Also, as expected, water quality public goods contain more severe consequences than economic development goods, but the difference in means is not statistically significant. Finally, hypothesis 3 is supported. Water quality goods contain more monitors, more aggregation rules, and more actors with the ability to trigger reviewing safeguards, thereby putting "more eyes" on activities. In this case, all three differences were statistically significant at $p < 0.1$.

For hypothesis 4a, **Figure 1** shows the distribution of monitoring mechanisms, identifying instances of the Federal or State governments as monitors (left half of the graph) and as reviewers of actions of the parties to the agreement (right half of the graph). Each half of the graph represents a two-by-two matrix with a cross-tabulation of the variables of interest. Of the ten public goods authorizing federal or state agencies to monitor, nine are water quality public goods, and the difference with economic development public goods is statistically significant ($p < .03$). In addition, all four of the public goods authorizing federal or state agencies to trigger review processes are water quality public goods. This result aligns with our hypothesis but is not statistically significant ($p < .13$), probably because of the small number of mechanisms. In sum, if State or Federal actors are tasked with monitoring or reviewing, those activities focus on water quality goods, as expected.

Figure 2 presents results for hypothesis 4b. It displays frequencies of safeguards mandating task specific actors (the CWC and the WPPC) to monitor (top graphs) and/or review (bottom) actions by watershed jurisdictions (right) and New York City (left) regarding water quality and economic development. In the case of the city (graphs on the left), safeguards have CWC and the WPPC monitoring and reviewing the city's

Table 3: T-test results for saliency hypotheses.

	Variable	Water quality mean (Std. Dev.)	Economic development mean (Std. Dev.)	Probability
H1	Number of safeguards	3.81 (4.5)	1.93 (0.99)	0.135
H2a	Number of monitoring safeguards	1.71 (1.87)	0.79 (0.89)	0.094
	Number of reviewing safeguards	1.14 (1.98)	0.36 (0.84)	0.172
	Number of consequence safeguards	0.95 (1.56)	0.79 (0.58)	0.706
H2b	Number of severe consequence mechanisms	0.14 (0.36)	0	0.147
	Number of mild consequence mechanisms	0.29 (0.46)	0.71 (0.47)	0.012
H3	Number of monitors	1.86 (1.9)	0.86 (0.95)	0.079
	Number of actors triggering compliance reviews	1.38 (1.8)	0.43 (0.85)	0.075
	Number of aggregation rules	1.14 (1.59)	0.143 (0.36)	0.028

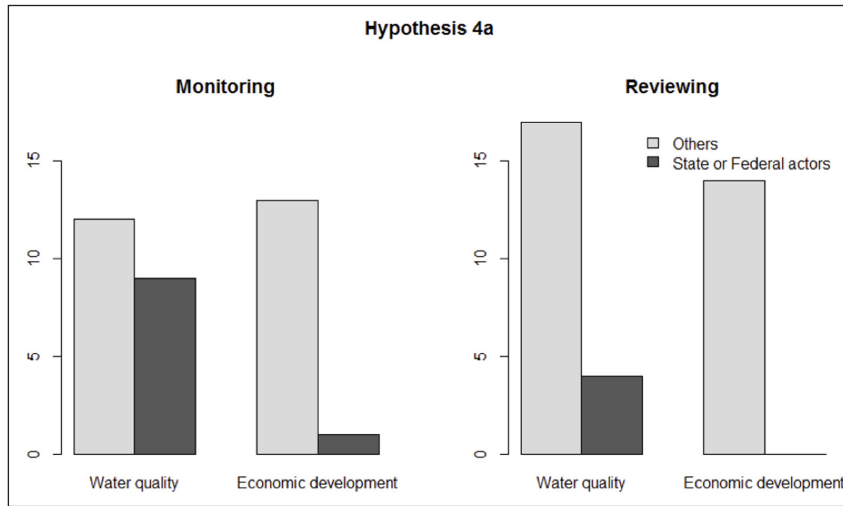


Figure 1: Federal and State as monitors and/or reviewers.

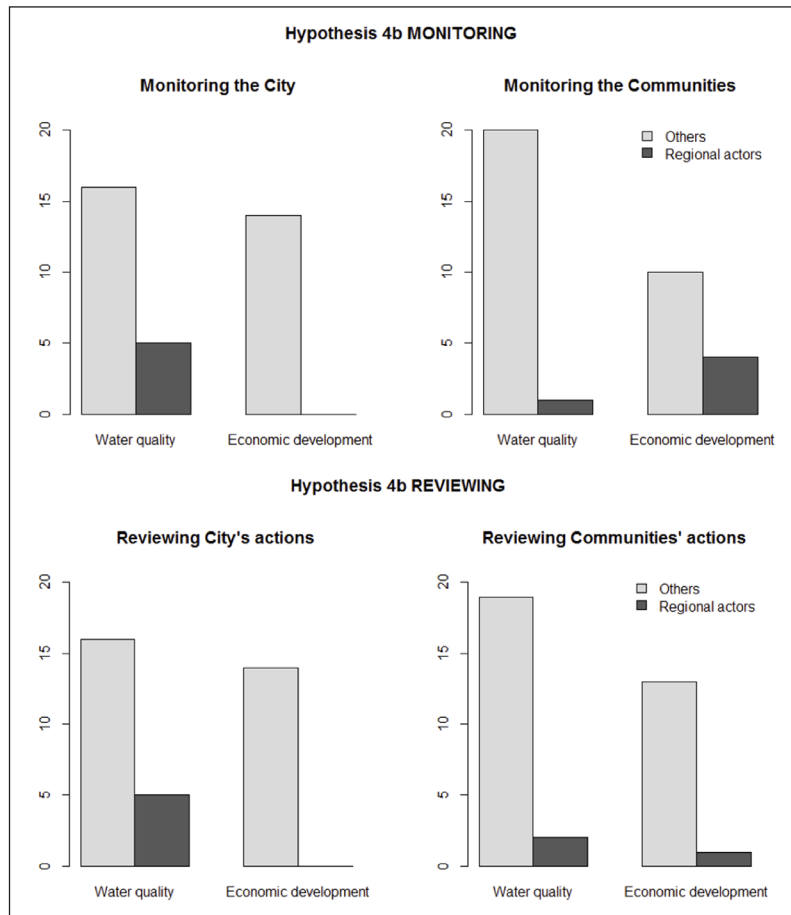


Figure 2: Regional actors as monitors and/or reviewers.

actions for water quality goods. The difference between safeguards regarding water quality and economic development is statistically significant (in both cases the p value is < .07).

In the case of safeguards targeting watershed jurisdictions, the evidence is modest. There are five public goods mandating regional actors monitor watershed jurisdictions. Of these, four are economic development goods and one focuses on water quality. This aligns with the hypothesis, but is not statistically significant ($p < .13$). In addition, there are only three safeguards authorizing task specific actors to trigger review processes, two are water quality goods and one is an economic development good, and thus no conclusions may be drawn. In sum, hypothesis 4b is supported regarding water quality public goods, but too few cases preclude drawing a conclusion for economic development goods.

Hypothesis 5 predicts that watershed jurisdictions and the city will monitor and review each other directly to prevent burden-shifting. We expect the city to monitor and review watershed jurisdictions regarding economic development, and watershed jurisdictions to monitor and review New York City regarding water quality activities. Only four public goods contain monitoring or reviewing for compliance safeguards that authorize the watershed jurisdictions and the City to monitor or trigger review processes regarding the other actor's actions. Two public goods, the tax litigation avoidance program and the enhanced monitoring program provide watershed jurisdictions with the opportunity to review and monitor the city (see **Table 4**). The city monitors watershed jurisdictions' flood control infrastructure projects and the spending of general funds. There are no discernable patterns among the four public goods. In general, the city and watershed jurisdictions do very little monitoring or reviewing for compliance on one another; and, when they are authorized to do so, it occurs to the same extent in water quality public goods as in economic development public goods.

Hypothesis 6 does not focus on whether monitoring and reviewing for compliance differs between types of public goods. Instead, the hypothesis focuses on who the city and the watershed jurisdictions are authorized to monitor and review. As revealed in **Table 5**, the city and the watershed jurisdictions engage in monitoring

Table 4: Burden-shifting directed safeguards.

Public good	Description	Good type	Safeguard type
Tax Litigation Avoidance Program	Watershed community property assessors may object to city choice of property tax assessment experts.	Economic development	Review by watershed jurisdictions
Enhanced Monitoring	The city shall provide watershed jurisdictions (and other parties) rationale if it does not adopt enhanced monitoring recommendations.	Water quality	Monitoring by watershed jurisdictions
Flood Program	Watershed jurisdictions sponsoring city-funded flood control construction projects shall provide the city notice of completion and right to inspect.	Water quality	Monitoring by the city
Good Neighbor Payments	The city may audit project fund spending in watershed jurisdictions at any time.	Economic development	Monitoring by the city

Table 5: Encroachment directed safeguards.

Public good	Description	Good type	Safeguard type
Watershed Agricultural Easements	If the city is prosecuting a violator of easement, the state may not prosecute at the same time, and must give the city adequate time to pursue their case.	Water quality	Review of the state by the city
Tax Litigation Avoidance Program	CWC shall notify the city of choice of tax assessor; the city shall approve any property tax assessor that the CWC hires to do independent appraisals.	Economic development	Review of the CWC by the city; Monitoring of the CWC by the city
Good Neighbor Payments	The city shall review disbursements of the project funds by CWC.	Economic development	Monitoring of the CWC by the city
Enhanced Monitoring	All parties may access and make suggestions on an independent report about enhancing the State-led water quality monitoring program.	Water quality	Monitoring of the state by the city, watershed jurisdictions
State Funding	The city may see which municipalities have applied for state funds and how much they received, to aid in its own funding decisions.	Economic development	Monitoring of the state by the city
Storm-water Retro Fit	Municipalities are notified of the completion of storm-water retrofit projects by CWC and given an opportunity to inspect them.	Water quality	Monitoring of the CWC by the watershed jurisdictions
Total Maximum Daily Loads	Pollutant load allowances are determined collaboratively between city, State and EPA, but city gets substantial input on the process.	Water quality	Monitoring of the state and EPA by the city

and reviewing for compliance of the State and CWC about equally, monitoring the CWC in three public goods and the State in four public goods. In addition, two public goods provide for reviewing, one directed at the State and one at the CWC. Thus, the evidence does not support the hypothesis. The city and watershed jurisdictions are authorized to monitor and review the State and the CWC about equally.

Discussion

The New York City watersheds governing arrangement creates a task specific jurisdiction that provides region-specific public goods (Hooge and Marks 2003). Without the MOA, New York City would not be granted a filtration avoidance determination and would instead have to invest billions of dollars in chemical filtration of its municipal water. In devising the MOA, the signatories linked themselves together in complex ways. Much like a federal arrangement, the MOA reduced the autonomy of its member jurisdictions while at the same time supporting coordination and cooperation among them. New York City surrendered its power of eminent domain in the watersheds (MOA, Section II) for the opportunity to work jointly with watershed jurisdictions to identify environmentally sensitive lands and willing sellers of those properties. And, no longer would watershed jurisdictions and organizations avail themselves of the courts to check the actions of the city; rather, disputes and grievances were to be brought before the WPPC (MOA, Article IV). Furthermore, if the city failed in its commitment to properly fund public goods arrangements, the watershed jurisdictions could have the land acquisition programs suspended; and, if the watershed jurisdictions pursued their grievances in court, the city could suspend its funding of public goods (MOA, Article V).

In this institutional context created by the MOA, we examined whether the specific public goods arrangements included multiple safeguards to ensure their proper implementation, and whether the design of the public goods safeguards varied in theoretically meaningful ways.

The evidence partially supported hypotheses 2a and 2b, and fully supported hypothesis 3. Altogether, these hypotheses proposed that water quality public goods, being highly salient to the task specific arrangement, would contain different safeguards than economic development goods. In particular, water quality public goods are heavily monitored, they include more actors monitoring and triggering compliance review processes, as well as fewer mild consequences for non-compliance. For instance, the Storm-water Retrofit Program involved New York City working with the CWC to fund and construct projects that would limit erosion and pollutant loadings caused by storm water runoff. Monitoring included inspections of projects under construction as well as ongoing monitoring over the life of the projects. In addition, if New York City did not make timely payments for construction and maintenance it would be charged interest on the outstanding payments. Storm-water Project sponsors could appeal to the CWC and have it review the project.

In contrast, economic development public goods have, on average, fewer monitoring mechanisms, as well as fewer monitors and actors in charge of triggering review processes. For instance, the Tax Consulting Fund was created in order to “pay the fees and expenses of professional consultants and/or attorneys retained by counties, towns or villages in [the watersheds] to review, analyze and/or assist in the administration of real property taxes paid by the city on city-owned lands.” (MOA, Section 5, 136(a)). The program was underwritten by the city, and if it did not fund the program in a timely fashion it was subject to interest on its late payments. This was the only safeguard provided for in the rules creating the program and it represents a common safeguard found across both water quality and economic development public goods.

While water quality public goods, on average, contain one more monitoring safeguard per public good compared to economic development goods, both types of public goods have, on average, indistinguishable numbers of reviewing and consequence safeguards. The difference between the public goods types is the severity of the penalties. All consequence safeguards that impose severe penalties are associated with water quality public goods. For instance, if the city misses a payment on a watershed easement and does not cure the violation, the New York State Department of Environmental Conservation may, after consulting with the contesting parties, suspend the ability of the city to acquire watershed land (MOA Article II, Section 85(e)). In contrast, of the fourteen economic development public goods, ten contained consequence safeguards and all imposed mild penalties, which were interest penalties assessed to New York City for late payments for the public goods. This difference, moreover, was statistically significant in relation to water quality public goods.⁹

⁹ As a robustness check for our assessment of hypotheses 1 through 3, we conducted a logistic regression for the first three hypotheses. The logit model allowed us to categorize the public goods by type, and it performed reasonably well, suggesting that

Regarding whether the safeguards were targeting specific types of opportunistic behaviors, the design of the safeguards clearly suggests that shirking, particularly around water quality public goods, was carefully attended to. Federal and state actors as well as the CWC and the WPPC were placed in a position to monitor and review actions regarding water quality public goods compared to economic development public goods. For instance, the Water Supply Permit section 26(e) and MOA section 85(e) gives the State Department of Environmental Conservation an important role as the trigger of a review process that may ultimately lead to severe sanctions for the city. The evidence for economic development public goods, however, was suggestive, but the fewer number of safeguards built into economic development public goods makes it difficult to draw firm conclusions. It appears, though, that the CWC and WPPC are authorized to monitor and review watersheds jurisdictions' actions related to economic development public goods.

We also explored patterns of monitoring and reviewing authority assigned to watershed jurisdictions and New York City (hypotheses 5 and 6). Regarding burden-shifting, there were only four public goods in which the city and watershed jurisdictions monitored or reviewed each other. Having a robust and empowered task specific jurisdiction may do much of the work of preventing burden-shifting, rendering direct safeguards unnecessary. Furthermore, most of the watershed jurisdictions are small towns and villages that lack the capacity to engage in extensive monitoring and reviewing activities, emphasizing the importance of the task specific jurisdiction for taking on these duties.

Hypothesis 6 is unsupported by the data. The city and watershed jurisdictions do not focus monitoring or reviewing efforts more toward the state compared to the task specific jurisdiction. A reading of the public goods arrangements that do provide the opportunity for the city and watershed municipalities to review higher level governments reveals that such safeguards against encroachment are designed to ensure that federal, state and task specific agencies and entities do not make decisions unilaterally.

The evidence provides support for the hypotheses that the creators of the New York Watersheds governing arrangement paid attention to the different types of public goods and their saliency for the arrangement when devising the rules. In particular, these results show how the creators of the agreement homed in on protecting the most salient public good, in the form of utilizing specific mechanism designs and assigning multiple actors responsibilities for implementing such safeguards. Furthermore, the safeguards appeared to be designed in to include coverage and various aspects of complementarity, as well as to address opportunistic behaviors in the form of shirking.

Conclusion

This paper provides a series of theoretical and methodological contributions for the study of regional-scale governance of shared natural resources. First, regional scale common pool resource governance entails a variety of actors, but especially governments who agree to limit their autonomy by creating regional arrangements that provide for collective benefits. Common pool resource theory, which focuses on individual resource users (Ostrom 1990), is suggestive of the design of governing arrangements that are likely to persist, but sheds little insight on the actions of governments and whether they will subject themselves to monitoring or sanctioning by other governments. Bednar's (2009) theory of safeguards, used to explain robust federations, focuses on nation states and their member governments, providing insights into how governments design arrangements that both provide for collective benefits but also support compliance with those arrangements on a national scale. However, the theory of safeguards is not sufficiently fine grained to capture local governments or regional settings. Combined, the two theories provide insights into the design and performance of regional scale governing arrangements, especially task specific jurisdictions.

Second, the manuscript presents a way of coding, organizing and analyzing formal institutions (i.e., rules in form) to explore whether actors anticipate collective action problems and design arrangements to address them. Drawing on the grammar of institutions (Crawford and Ostrom 2005; Basurto, et al., 2010; Siddiki, et al., 2011), we coded the rules constituting the New York City Watershed governing arrangement and used the rules to develop measures of three of Ostrom's design principles/Bednar's safeguards. To our knowledge, this is the first time the grammar has been used to do so. In addition to identifying safeguards, we applied the grammar to identify who was participating in safeguards and who was subject to safeguards. Doing this

the safeguards, the multiple actors involved in decision making, monitoring and reviewing, and the number of aggregation rules distinguish the water quality public goods from the economic development public goods. To see descriptive statistics as well as this analysis, contact Edella Schlager at schlager@email.arizona.edu.

allowed us to develop quantifiable observations that we then used to test hypotheses using configurations of rules and their characteristics.

There are a few clear limitations of this paper. First, it relies on data from a single case of watershed governance, and thus lacks the power to generalize across CPR governance cases of this type generally. However, this is a first step in proposing how such data may be collected and analyzed, and for that it has utility. Second, the results may not be generalizable to cases in other federal arrangements outside the United States that have different political cultures or variations of institutional structure, such as India or Brazil, for example. Also, New York State is not representative of US States generally, and has a strong home rule tradition and particular political culture born out of its colonial heritage. With more analyses in diverse settings, as well as comparative analyses, the research design, methods, and theory tested here may be more rigorously scrutinized.

This approach of combining federalism and common pool resources theory can be further developed and used to address additional questions. For instance, a recent meta-analysis of more than 60 cases of local level self-governing arrangements of fisheries, forests, and irrigation systems that examined the presence of Ostrom's design principles and performance found that a necessary, although not sufficient, condition for success was design principle two (Baggio, et al., 2016). Design principle two focuses on "congruence between appropriation and provision rules and local conditions" (Ostrom 1990). Using the methods developed and applied in this paper, it would be possible to examine how the configurations of rules governing common pool resources determine access to resources and public goods, allocating collective benefits and costs for providing those benefits. Measures of congruence, especially between appropriation and provision rules could be developed to identify and explore how the rules condition who benefits and who pays for benefits. This may also open lines of research on how rules reflect the culture of the rule designers.

Regional governing arrangements are becoming increasingly common and understanding their design and performance is likely to yield important theoretical and applied policy lessons. This is especially true in contexts where nation-states are pushing back from commitments regarding shared natural resource governance, and where local and state governments are assuming those responsibilities in part. It thus becomes critical to better understand the array of institutional tools available for fostering regional-scale collaboration. Our paper constitutes a step in that direction, by exploring the design of governing arrangements involving state and local governments and incorporating insights from common pool resource and federalism theories to explain the complex governing arrangements devised by these organizations to sustainably governing common pool resources and public goods.

Additional File

The additional file for this article can be found as follows:

- **Appendix.** The codebook, descriptive statistics, and robustness check for analyses can be obtained directly from Edella Schlager (schlager@email.arizona.edu) upon request. DOI: <https://doi.org/10.5334/ijc.924.s1>

Competing Interests

The authors have no competing interests to declare.

References

- Anderies, J. M., Janssen, M., & Ostrom, E.** (2004). A Framework to Analyze the Robustness of Social-ecological Systems from an Institutional Perspective. *Ecology and Society*, 9(1), 18. DOI: <https://doi.org/10.5751/ES-00610-090118>
- Baggio, J. A., Barnett, A. J., Perez-Ibarra, I., Brady, U., Ratajczyk, E., Rollins, N., Rubiños, C., Shin, H. C., Yu, D. J., Aggarwal, R., Anderies, J. M., & Janssen, M. A.** (2016). Explaining success and failure in the commons: the configural nature of Ostrom's institutional design principles. *International Journal of the Commons*, 10(2), 417–439. DOI: <https://doi.org/10.18352/ijc.634>
- Basurto, X., Kingsley, G., McQueen, K., Smith, M., & Weible, C. M.** (2010). A Systematic Approach to Institutional Analysis: Applying Crawford and Ostrom's Grammatical Syntax. *Political Research Quarterly*, 63, 523–37. DOI: <https://doi.org/10.1177/1065912909334430>
- Bednar, J.** (2009). *The Robust Federation: Principles of design*. New York, NY: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511819445>

- Blomquist, W.** (1992). *Dividing the Waters: Governing Groundwater in Southern California*. San Francisco, CA: ICS Press.
- Breitmeier, H., Young, O. R., & Zürn, M.** (2006). *Analyzing International Environmental Regimes: From Case Study to Database*. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/1264.001.0001>
- Chhatre, A., & Agrawal, A.** (2008). Forest commons and local enforcement. *Proceedings of the National Academy of Sciences of the United States of America*, 105(36), 13286–13291. DOI: <https://doi.org/10.1073/pnas.0803399105>
- Coleman, E. A., & Steed, B.** (2009). Monitoring and sanctioning in the commons: An application to forestry. *Ecological Economics*, 68(7), 2106–2113. DOI: <https://doi.org/10.1016/j.ecolecon.2009.02.006>
- Crawford, S. E., & Ostrom, E.** (1995). A grammar of Institutions. *American Political Science Review*, 89(03), 582–600. DOI: <https://doi.org/10.2307/2082975>
- Crawford, S. E., & Ostrom, E.** (2005). A grammar of Institutions. In E. Ostrom (Ed.), *Understanding Institutional Diversity* (pp. 137–174). Princeton, NJ: Princeton University Press.
- Dietz, T., Ostrom, E., & Stern, P. C.** (2003). The Struggle to Govern the Commons. *Science*, 302(5652), 1907–1912. DOI: <https://doi.org/10.1126/science.1091015>
- Elazar, D. J.** (1987). *Exploring federalism*. Tuscaloosa, AL: The University of Alabama Press.
- Fiock, R. C.** (2013). The Institutional Collective Action Framework. *Policy Studies Journal*, 41(3), 397–425. DOI: <https://doi.org/10.1111/psj.12023>
- Fiock, R. C., & Scholz, J. T.** (2010). *Self-organizing federalism: Collaborative Mechanisms to Mitigate Institutional Collective Action Dilemmas*. New York, NY: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511642319>
- Filippov, M., Ordeshook, P. C., & Shvetsova, O.** (2004). *Designing federalism: A theory of self-sustainable federal institutions*. New York, NY: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511610875>
- Fleischman, F., Ban, N. C., Evans, L. S., Epstein, G., García-Lopez, D., & Villamayor-Tomás, S.** (2014). Governing large-scale social-ecological systems: Lessons from five cases. *International Journal of the Commons*, 8(2), 428–456. DOI: <https://doi.org/10.18352/ijc.416>
- Galusha, D.** (2016). *Liquid Assets: A History of New York City's Water System. Expanded Edition*. Fleischmanns, NY: Purple Mountain Press.
- Gibson, C. C., Williams, J. T., & Ostrom, E.** (2005). Local Enforcement and Better Forests. *World Development*, 33(2), 273–284. DOI: <https://doi.org/10.1016/j.worlddev.2004.07.013>
- Heikkila, T., Schlager, E., & Davis, M. W.** (2011). The Role of Cross-Scale Institutional Linkages in Common Pool Resource Management: Assessing Interstate River Compacts. *Policy Studies Journal*, 39(1), 121–145. DOI: <https://doi.org/10.1111/j.1541-0072.2010.00399.x>
- Hooghe, L., & Marks, G.** (2003). Unraveling the Central State, but How? Types of Multi-level Governance. *American Political Science Review*, 97(2), 233–243. DOI: <https://doi.org/10.1017/S0003055403000649>
- Lutz, D. S.** (1990). From Covenant to Constitution in American Political Thought. *Publius: The Journal of Federalism*, 10(4), 101–133. DOI: <https://doi.org/10.2307/3329746>
- McCay, B.** (2002). Emergence of Institutions for the Commons: Contexts, Situations, and Events. In E. Ostrom, T. Dietz, N. Dolšak, P. C. Stern, S. Stonich, & E. U. Weber (Eds.), *The Drama of the Commons*. Washington, DC: National Academies Press.
- Oakerson, R.** (1999). *Governing Local Public Economies: Creating the Civic Metropolis*. Oakland, CA: ICS Press.
- Ostrom, E.** (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. New York, NY: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511807763>
- Ostrom, E.** (1999). Coping with tragedies of the commons. *Annual Review of Political Science*, 2(1), 493–535. DOI: <https://doi.org/10.1146/annurev.polisci.2.1.493>
- Ostrom, E.** (2005). *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.
- Ostrom, E., Walker, J., & Gardner, R.** (1992). Covenants With and Without a Sword: Self-Governance is Possible. *American Political Science Review*, 86(2), 404–417. DOI: <https://doi.org/10.2307/1964229>
- Ostrom, V.** (2007). *The Political Theory of a Compound Republic*. 3rd Edition. Lexington Books.
- Ostrom, V., Tiebout, C. M., & Warren, R.** (1961). The organization of government in metropolitan areas: a theoretical inquiry. *American Political Science Review*, 55(4), 831–842. DOI: <https://doi.org/10.1017/S0003055400125973>
- Quinn, C. H., Huby, M., Kiwasila, H., & Lovett, J. C.** (2007). Design principles and common pool resource management: An institutional approach to evaluating community management in semi-arid

- Tanzania. *Journal of Environmental Management*, 84(1), 100–113. DOI: <https://doi.org/10.1016/j.jenvman.2006.05.008>
- Schlager, E.** (1994). Coastal Fishery Dilemmas and Fishers' Institutional Responses, In E. Ostrom, R. Gardner, & J. Walker (Eds.), *Rules, Games and Common Pool Resources*. Ann Arbor: University of Michigan Press.
- Schlager, E.** (2004). Common-Pool Resource Theory. In R. R. Durant, D. J. Fiorino, & R. O'Leary (Eds.), *Environmental Governance Reconsidered: Challenges, Choices, and Opportunities* (pp. 145–175). Cambridge, MA: MIT Press.
- Schlager, E., & Blomquist, W. A.** (2008). *Embracing watershed politics*. Boulder, CO: University Press of Colorado. DOI: <https://doi.org/10.2307/j.ctt46nvqs>
- Schoon, M. L., & Cox, M. E.** (2012). Understanding Disturbances and Responses in Social-Ecological Systems. *Society & Natural Resources*, 25(2), 141–155. DOI: <https://doi.org/10.1080/08941920.2010.549933>
- Siddiki, S. N., Weible, C. M., Basurto, X., & Calanni, J.** (2011). Dissecting policy designs: An application of the institutional Grammar tool. *Policy Studies Journal*, 39(1), 79–103. DOI: <https://doi.org/10.1111/j.1541-0072.2010.00397.x>
- Soll, D.** (2013). *Empire of Water: An Environmental and Political History of the New York City Water Supply*. Ithaca, New York: Cornell University Press.
- Tang, S. Y.** (1992). *Institutions and collective action: Self-governance in irrigation*. San Francisco, CA: ICS Press.
- United States Environmental Protection Agency.** (2017). *New York City Filtration Avoidance Determination: Surface Water Treatment Rule Determination for New York City's Catskill/Delaware Water Supply System*.
- Villamayor-Tomas, S., Fleischman, F. D., Perez Ibarra, I., Thiel, A., & van Laerhoven, F.** (2014). From Sandoz to Salmon: Conceptualizing resource and institutional dynamics in the Rhine watershed through the SES Framework. *International Journal for the Study of the Commons*, 8(2), 361–395. DOI: <https://doi.org/10.18352/ijc.411>
- Wolf, A. T., Yoffe, S. B., & Giordano, M.** (2003). International waters: identifying basins at risk. *Water Policy*, 5(1), 29–60. DOI: <https://doi.org/10.2166/wp.2003.0002>

How to cite this article: Hanlon, J., Olivier, T., & Schlager, E. (2019). Suspicious Collaborators: How Governments in Polycentric Systems Monitor Behavior and Enforce Public Good Provision Rules Against One Another. *International Journal of the Commons*, 13(2), pp. 977–992. DOI: <https://doi.org/10.5334/ijc.924>

Submitted: 19 July 2018

Accepted: 17 March 2019

Published: 30 October 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



International Journal of the Commons is a peer-reviewed open access journal published by Ubiquity Press.

