Physical constraints and appropriative water rights: Implications for the efficiency of water markets

Joey Blumberg[1], Chris Goemans[2], and Dale Manning[3]

[1] *USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO 80526,* joey.blumberg@gmail.com

[2] *Agricultural and Resource Economics, Colorado State University, Fort Collins, CO 80526,* chris.goemans@gmail.com

[3] *Baker School for Public Policy; Agricultural and Resource Economics, University of Tennessee, Knoxville, TN 37996,* dtmanning@utk.edu

November 10, 2025

Abstract

Water demand in many arid regions exceeds available supplies, and the growing gap between the two has drawn increased attention to the institutional efficiency of water governance systems. This article evaluates the economic efficiency of different water allocation mechanisms and how market efficiency is impacted by water shortages, return flows, and heterogeneity in the initial allocation of appropriative water rights. We develop analytical and numerical water allocation models for a hypothetical river basin and use simulations of surface water supply conditions to evaluate changes in economic value under different basin compositions and institutional frameworks. Prior appropriation, the predominant institution in the western United States, is compared to alternative, share-based water allocation regimes. We find that due to return flows—the positive externalities associated with water use—the initial allocation of appropriative rights can distort the effective availability of the resource. When senior rights are held downstream, total water use can be constrained, particularly under shortage conditions. Moreover, competitive water markets do not always outperform share-based systems in mitigating these inefficiencies. Importantly, and contrary to previous literature, these results suggest that even well-functioning markets may be unable to fully offset inefficiencies rooted in the initial allocation of water rights.

Keywords: prior appropriation, water sharing, water trading, water allocation, optimization

1 Introduction

Efficient water management is a growing global challenge as climate change and population growth place increasing pressure on finite freshwater resources. These challenges are particularly acute in the western United States (WUS), where rapid population growth (Brown et al., 2019), intensifying droughts (Gutzler and Robbins, 2011), and new and competing demands for water (Dieter et al., 2018; Schwabe et al., 2020) strain existing systems (Blumberg and Warziniack, 2025). The extent to which water scarcity affects arid river basins depends on the ability of human systems to adapt. Although technological innovations have led to marked improvements in water use efficiency (Donnelly and Cooley, 2015), institutional changes in the governance of western water are occurring at a slower pace (Brewer et al., 2006; Libecap, 2011; Leonard et al., 2019). Achieving outcomes close to a social optimum through technological improvement alone is unlikely if the mechanisms governing water use inherently introduce inefficiencies. As such, water users and policymakers are questioning whether alternative allocation institutions to those traditionally used throughout the WUS should be considered (UNM, 2023; USU, 2023). In this study, we assess how institutional mechanisms for allocating water rights, both with and without markets, influence the efficiency of water use. We further examine how factors such as shortages, the spatial arrangement of users, and return flows shape outcomes within these institutional frameworks—a set of interactions that has received limited attention in prior economic research.

In the humid regions of the eastern United States, water is typically allocated under the riparian water rights doctrine. This system grants landowners abutting water bodies the right to use adjacent water "reasonably," without significantly diminishing its supply for downstream users (Jones and Cech, 2009). In contrast, in the arid regions of the WUS, where the prevalence of irrigated agriculture requires large diversions away from waterways and high consumptive use, a riparian doctrine is insufficient to ensure the rights of downstream users to divert water. Consequently, the WUS has adopted a doctrine of prior appropriation (Jones and Cech, 2009), which assigns water rights based on the act of physically diverting and putting water to "beneficial use." Under this system, users must obtain a legal right to divert water from any natural system, and these rights are ordered by their date of establishment, which places them within a hierarchy of seniority.

When basin water supplies are insufficient to fulfill all allotments, older (senior) rights are sat-

isfied first, while newer (junior) rights may receive nothing. This priority structure gives rise to the defining feature of prior appropriation: the "first in time, first in right" principle. By linking rights to beneficial use and prioritizing seniority, the system is designed to encourage investment in diversion infrastructure and deter speculative claims on undeveloped water. Indeed, there is evidence that appropriative rights facilitated collective investment in irrigation systems across the WUS by reducing risk for senior users (Libecap, 2007; Leonard and Libecap, 2019). At the same time, economists widely agree that prior appropriation leads to inefficient outcomes when not accompanied by a functioning water market (Burness and Quirk, 1979, 1980; Huffaker et al., 2000). In their seminal work, Burness and Quirk (1979, 1980) analyze this inefficiency by modeling the long-run equilibrium of expected diversions and profits for firms sharing a waterway. They demonstrate that efficient resource use can be achieved under prior appropriation given the presence of a competitive leasing market. However, the institutional and hydrological landscape of the WUS has evolved substantially since their analysis.

Modern water markets, while increasingly common, often remain highly localized and face considerable transaction costs—including legal, informational, and infrastructure-related barriers to trade (Brewer et al., 2006; Libecap, 2011; Leonard et al., 2019; Womble and Hanemann, 2020; Rafey, 2023). This renders current water allocations relatively inflexible, hindering the transfer of water to more beneficial uses as new and potentially higher-value demands emerge (Brewer et al., 2008). The initial allocation of water rights therefore becomes an important determinant of water use efficiency, as high transaction costs may prevent markets from correcting allocation inefficiencies. Additionally, many basins in the WUS are now over-appropriated, with average supplies insufficient to satisfy all legal claims (Spivak, 2021). In such settings, junior rights holders (who may have high marginal values for water) can be entirely curtailed, and if located upstream, this curtailment can stifle the ability to reuse water via return flows. Under a strict priority regime, these uses by junior users may be foregone even when they impose minimal consumptive loss.

In general, the current body of economic research offers important perspectives on optimizing water use under institutional constraints, particularly emphasizing the role of market mechanisms (Chong and Sunding, 2006; Schwabe et al., 2020). The Coase Theorem (Coase, 1960), a cornerstone in environmental economics, posits that if property rights are clearly defined and transaction costs are negligible, markets can drive an efficient allocation of resources regardless of the initial

distribution of rights. Coase stresses the potential for private negotiations to address externalities and achieve optimal resource allocation. However, the application of Coasean bargaining to water rights governed by a doctrine of prior appropriation is not clear-cut. In practice, the unique physical characteristics of water and the structure of contemporary water allocation mechanisms make it difficult to design markets that meet Coase's assumptions (Debaere et al., 2014; Bruno and Jessoe, 2024). For example, when high-priority rights are held downstream, upstream junior users may be unable to divert and generate return flows, reducing the total water available for reallocation.[1] Such hydrologic constraints are directly influenced by the initial allocation of water rights, and when combined with the significant transaction costs common in real-world water markets, they limit the extent to which actual outcomes can align with the Coasean ideal.

In contrast to fixed-allocation mechanisms like prior appropriation, share-based regimes—where water rights are allocated as proportional shares of the available supply rather than as fixed quantities—offer a different paradigm. For example, some farmers have formed cooperative organizations called "irrigation districts," pooling the priority of their water rights and redistributing risk evenly across shares.[2] Water is then allocated among members according to the number of shares they own, ensuring nonzero water allocations each year for each shareholder. In practice, Ji and Cobourn (2018) find that irrigation districts in Colorado allocate larger portions of land to drought-sensitive, high-value crops compared to areas with similar overall priority without a share-based system, resulting in a 6% average increase in profit. Using a simulation-based approach, He et al. (2012) find a potential gain in producer surplus if irrigation districts were to switch from seniority-based allocations to proportional-sharing water policies. Using a natural experiment along the Colorado–New Mexico border, Smith (2021) finds that proportional-sharing systems generate higher water productivity than prior-appropriation regimes but may also encourage excessive investment when initial shares are poorly calibrated. Overall, these studies align with Ostrom and Gardner (1993), who suggest that irrigation systems endure when rights and obligations are proportional and users participate in rule-making.

Such sharing mechanisms provide two key advantages when markets are limited. First, they

[1] Water diverted is not a perfectly rivalrous good, and most water uses necessitate diverting more water than what is consumed, with return flows ranging from 5–90% depending on the use (Burness and Quirk, 1980).

[2] Globally, water-sharing plans have been adopted in parts of the United States, Canada, Australia, and Chile (Grafton et al., 2011; Tingey-Holyoak, 2014).

guarantee some upstream diversion, which can increase total return flows. Second, they reduce asymmetry in risk across users, producing an initial allocation closer to efficiency. Share-based systems may thus lessen dependence on markets to reach optimal outcomes. The initial allocation of rights therefore influences not only the total water available through return-flow externalities but also the degree of market activity required to achieve efficiency.

The influence of initial property rights on market efficiency is a topic of research not unique to water rights. In the context of land demarcation, Libecap and Lueck (2011) show that different property allocation regimes have enduring impacts on land values, transaction costs, and the structure of land markets. They find that the rectangular system, by providing standardized, secure property rights, enhances economic outcomes compared to the more flexible metes-and-bounds system. In the context of sulfur dioxide cap-and-trade systems, Henry III et al. (2011) demonstrate that the initial allocation of pollution permits can lead to increased social costs by shifting emissions to areas of higher marginal damage. Such examples in the broader economics literature highlight the principle that the initial distribution of property rights can influence the efficiency of resource allocation, even in the presence of markets.

The present study contributes to this broad economic literature by examining two relatively understudied determinants of water allocation outcomes: (1) the initial assignment of rights and (2) the physical characteristics of river basins. We show that the interaction of these two features can either exacerbate or mitigate institutional inefficiencies. We begin with an analytical model to demonstrate how the spatial ordering of users under priority-based allocation regimes can distort efficiency by limiting the return flows available for downstream use. We then turn to the role of leasing markets, which in principle could reallocate water to mitigate these inefficiencies, and derive conditions under which socially efficient trades deviate from privately profitable ones. Finally, we use Monte Carlo simulations of water supply and user placement to resolve analytical ambiguities and identify the basin conditions under which each institution performs best.

Results reveal that system-wide efficiency under prior appropriation is especially sensitive to the spatial ordering of user priorities and the hydrologic potential for return flows during shortage conditions—even when markets operate without friction. Because return flows create positive externalities, assigning senior rights downstream can reduce the volume of water available to the system as a whole. When water is abundant, leasing markets achieve efficient allocations. However,

under moderate to severe shortages—an increasingly common occurrence in the WUS—the performance of prior appropriation depends on whether upstream users hold senior rights and whether they generate high return flows. Much of the literature suggests that if transaction costs could be substantially reduced, water markets could eliminate the inefficiencies associated with prior appropriation (Garrick et al., 2013; Erfani et al., 2014). However, our analysis shows that even in the complete absence of transaction costs, markets cannot fully resolve inefficiencies arising from return-flow externalities. By examining the physical factors that impact institutional efficiency, this study provides valuable insights for policymakers and water managers in the development and implementation of effective and efficient water allocation strategies.

2 Analytical model

Consider a linear waterway with three users located sequentially along the stream, indexed from furthest upstream ($i = 1$) to furthest downstream ($i = 3$). Let w_i denote the diversion quantity of water of user i in a given year, and let $v_i(w_i)$ be the associated value of water use. Available surface water in that year is given by $\tilde{W} \in \boldsymbol{W}$, where $\boldsymbol{W}$ is the distribution of possible supply conditions. A proportion $\theta_i \in (0, 1)$ of water diverted by user i returns to the stream and becomes available for all downstream users, while the remaining $(1 - \theta_i)$ is consumed or otherwise lost from the system.

In practice, return flows can originate from both conveyance losses (e.g., seepage during transport) and from water that is used and then partially returned to the stream (e.g., irrigation). Let $\kappa_i \in [0, 1)$ denote the fraction of water lost during conveyance to user i, and $\phi_i \in [0, 1)$ the fraction of the diversion that is consumed during use. Then the effective diversion reaching user i is $w_i = (1 - \kappa_i)\tilde{W}_i$, and the fraction returning to the stream is $(1 - \phi_i)$. Combining these gives an effective return-flow parameter $\theta_i = (1 - \kappa_i)(1 - \phi_i)$, so that $\theta_i w_i$ re-enters the stream while $(1 - \theta_i)w_i$ is lost. Both sources of loss can be subsumed into the single parameter θ_i without loss of generality.

2.1 The social planner

To illustrate the basic allocation problem, suppose a central decision-maker seeks to maximize the total value of water use across all three users:

$$\begin{aligned} \max_{w_1,w_2,w_3\geq 0} \quad & v_1(w_1) + v_2(w_2) + v_3(w_3) \\ \text{s.t.} \quad & w_1 \leq \tilde{W}, \\ & w_2 \leq \tilde{W} - (1-\theta_1)w_1, \\ & w_3 \leq \tilde{W} - (1-\theta_1)w_1 - (1-\theta_2)w_2. \end{aligned} \tag{1}$$

The first constraint states that the most upstream user cannot divert more than the total water available at the head of the system. The second constraint accounts for consumptive and conveyance losses from user 1's diversion, leaving $\theta_1 w_1$ plus any unused water for user 2. The third constraint similarly accounts for losses from both upstream users before water reaches user 3. This formulation provides a general framework for describing water allocation in a connected river system with return flows, and will serve as the foundation for introducing institutional rules in subsequent sections. Figure A1 provides a visual representation of this set-up.

Suppose all three users have value functions that are twice differentiable, strictly increasing and concave, and return flow parameters are the same for all users, $\theta_i = \theta \in (0, 1)$. Let the constraints in Equation 1 be written in slack form:

$$\begin{aligned} g_1(\mathbf{w}) &= \tilde{W} - w_1 \geq 0, \\ g_2(\mathbf{w}) &= \tilde{W} - (1-\theta)w_1 - w_2 \geq 0, \\ g_3(\mathbf{w}) &= \tilde{W} - (1-\theta)w_1 - (1-\theta)w_2 - w_3 \geq 0. \end{aligned} \tag{2}$$

The Lagrangian is

$$\mathcal{L} = v_1(w_1) + v_2(w_2) + v_3(w_3) + \lambda_1 g_1(\mathbf{w}) + \lambda_2 g_2(\mathbf{w}) + \lambda_3 g_3(\mathbf{w}), \tag{3}$$

with $\lambda_i \geq 0$ and $\lambda_i g_i(w) = 0$ for all i, and the first-order conditions (FOCs) are:

$$\begin{aligned} v_1'(w_1) &= \lambda_1 + (1-\theta)\big(\lambda_2 + \lambda_3\big), \\ v_2'(w_2) &= \lambda_2 + (1-\theta)\lambda_3, \\ v_3'(w_3) &= \lambda_3. \end{aligned} \tag{4}$$

In the interior case with positive allocations and intermediate scarcity—the downstream availability constraint binds ($\lambda_3 > 0$) while the headwater constraint is slack ($\lambda_1 = 0$)—the FOCs imply

$$v_1'(\hat{w}_1) < v_2'(\hat{w}_2) < v_3'(\hat{w}_3). \tag{5}$$

Observation 1. Due to the positive externalities created through return flows, the first best allocation under the social planner results in upstream users using water to a lower marginal value than downstream users. Furthermore, when all users are identical (i.e., $v_i(\cdot) = v(\cdot)\ \forall\ i$), the social planner unambiguously allocates more water to upstream users ($\hat{w}_1 > \hat{w}_2 > \hat{w}_3$).[3]

2.2 Prior appropriation

In contrast to the social planner's allocation, which seeks to maximize total basin value, a system of prior appropriation (PA) allocates water to satisfy the rights of users in order of seniority, without regard to marginal values. To represent priority, let $m \in \{1, 2, 3\}$ index seniority, with $m = 1$ the most senior and $m = 3$ the most junior. Under PA, diversions are made sequentially in increasing order of m: the most senior right ($m = 1$) must be fully satisfied before any junior rights can be exercised, and junior users may divert water only insofar as they do not injure more senior users. We again assume that all three users have strictly increasing and concave value functions $v_{i,m}(\cdot)$, return flow parameters $\theta_{i,m} = \theta \in (0, 1)$, and indices $i \in \{1, 2, 3\}$ that denote stream position from furthest upstream to furthest downstream.

We focus on two polar cases to illustrate how PA outcomes depend on the alignment of priority with stream position. In the first case, priority and stream position are perfectly *negatively* correlated: the furthest downstream user ($i = 3$) holds the most senior right ($m = 1$), the middle user ($i = 2$) holds the intermediate right ($m = 2$), and the furthest upstream user ($i = 1$) holds the

[3]This is further demonstrated for n users in the Appendix (subsection A.2).

most junior right ($m = 3$). In the second case, priority and stream position are perfectly *positively* correlated: the furthest upstream user ($i = 1$) is most senior ($m = 1$), the middle user ($i = 2$) is intermediate ($m = 2$), and the furthest downstream user ($i = 3$) is most junior ($m = 3$). Each user holds a maximum diversion right of size $x > 0$, and total inflow equals $\tilde{W} = x$, so there is just enough water in the stream to satisfy the most senior right.

Case 1: Downstream seniority. Meeting the senior right requires $w_{i=3,m=1} = x$. To deliver x units to this user, any upstream diversion would reduce the inflow by $(1 - \theta)$ and therefore injure the senior. Thus, junior rights must be fully curtailed, leading to allocations of $\{\hat{w}_{13}, \hat{w}_{22}, \hat{w}_{31}\} = \{0, 0, x\}$, total water use of x, and a total system value of $v_{31}(x)$.

Case 2: Upstream seniority. Meeting the senior right requires $w_{i=1,m=1} = x$. Because the senior is upstream, junior diversions from return flows do not injure the senior (whose entitlement is already satisfied). Hence the feasible PA allocation is $\{\hat{w}_{11}, \hat{w}_{22}, \hat{w}_{33}\} = \{x, \theta x, \theta^2 x\}$, leading to total water use of $x(1 + \theta + \theta^2)$ and a total system value of $v_{11}(x) + v_{22}(\theta x) + v_{33}(\theta^2 x)$.

These cases illustrate how the spatial ordering of users interacts with PA's seniority rule during scarcity: placing seniority upstream allows return flows to support junior uses without injuring the senior, increasing total water use; placing seniority downstream curtails upstream diversions and foregoes those downstream benefits. It can also be shown that when the senior user is located in the middle ($i = 2, m = 1$), total water use falls between Case 1 and Case 2. Moreover, in Case 1, the allocations would be unchanged if priority between users $i = 1$ and $i = 2$ were reversed, since the decisive factor is simply that the most senior right lies furthest downstream. In Case 2, allocations differ if users $i = 2$ and $i = 3$ switch positions in the priority ranking, but total diversions remain unambiguously larger than in Case 1.

Observation 2. Under PA, more total water is diverted for beneficial use when senior users are located upstream and inflow is insufficient to fulfill all entitlements. Downstream seniority represents a clear departure from the social planner's strategy, which favors upstream diversions to exploit the partially non-rivalrous nature of water use through return flows. Furthermore, if value functions are identical across seniority levels (i.e., $v_{1,m} = v_{2,m} = v_{3,m} \; \forall \; m$), Case 2 unambiguously

results in higher total value generated.

2.3 Proportional and equal sharing

Proportional sharing (PS) refers to a family of allocation rules designed to spread risk among water users, in contrast to the strictly hierarchical ordering of prior appropriation. In practice, PS can take many forms. Some irrigation districts allocate water proportionally to acres planted, while others pool appropriative rights and reassign shares based on seniority and diversion entitlements. In other cases, such as the Colorado–Big Thompson project, water is supplied through transbasin diversion and allocated in proportion to the number of shares purchased and water availability in a given year, with no consideration of return flows. These institutional variants share the common feature that each user's allocation is tied to a proportion of total available supply, but they differ in how shares are defined and assigned.

Because of this empirical diversity, our analytical exercise focuses on a limiting case: equal sharing (ES), in which all users receive identical allocations regardless of stream position or prior claims. This setup removes hierarchy entirely and provides a useful benchmark for comparing efficiency outcomes when priority rules are abandoned. We return to more nuanced proportional sharing rules in our simulation exercises, where alternative share definitions can be incorporated explicitly.

Formally, with three users and identical return flow parameters $\theta_i = \theta \in (0, 1)$, the ES allocation problem is:

$$\begin{aligned} \max_{w_1, w_2, w_3 \geq 0} \quad & v_1(w_1) + v_2(w_2) + v_3(w_3) \\ \text{s.t.} \quad & w_1 \leq \tilde{W}, \\ & w_2 \leq \tilde{W} - (1-\theta) w_1, \\ & w_3 \leq \tilde{W} - (1-\theta) w_1 - (1-\theta) w_2, \\ & w_1 = w_2 = w_3. \end{aligned} \tag{6}$$

The first three constraints ensure physical feasibility given return flows, while the final constraint enforces equal diversions. Let q denote the common diversion amount, so Equation 6 can be

expressed as:

$$\begin{aligned} q &\leq \tilde{W}, \\ q &\leq \tilde{W} - (1-\theta)q, \\ q &\leq \tilde{W} - (1-\theta)q - (1-\theta)q. \end{aligned} \tag{7}$$

The binding condition is the third constraint, which yields:

$$q = \frac{\tilde{W}}{1 + 2(1-\theta)}. \tag{8}$$

Total diversions under ES are therefore:

$$3q = \frac{3\tilde{W}}{1 + 2(1-\theta)}. \tag{9}$$

To facilitate comparison with PA, we focus on the same scarcity cases introduced above, where $\tilde{W} = x$ and each user holds a right of size $x > 0$.

Case 1: Downstream seniority under PA. Under PA, the allocation is $\{0, 0, x\}$, yielding total diversions of x. Under ES, the allocation is $\{q, q, q\}$ with $q = x/[1 + 2(1-\theta)]$, yielding total diversions of $3x/[1 + 2(1-\theta)]$. Since $\theta \in (0, 1)$, we have $1 + 2(1-\theta) < 3$, which implies $3x/[1 + 2(1-\theta)] > x$. Thus, ES always results in greater total diversions than PA in this case.

Case 2: Upstream seniority under PA. Under PA, the allocation is $\{x, \theta x, \theta^2 x\}$, yielding total diversions of $x(1+\theta+\theta^2)$. Under ES, total diversions are again $3x/[1+2(1-\theta)]$. In this case, $x(1 + \theta + \theta^2) > 3x/[1 + 2(1-\theta)]$ for all $\theta \in (0, 1)$, so PA results in strictly greater total diversions than ES.

Observation 3. Under ES, some amount of water is always diverted upstream, so some return flow capacity is internalized. When seniority and stream position are perfectly negatively correlated (Case 1), ES always yields higher total diversions than PA. When seniority and stream position are perfectly positively correlated (Case 2), PA always yields higher total diversions than ES. In terms of total value, the ranking generally depends on users' value functions. If value functions are identical, strictly increasing, and concave, then ES strictly dominates PA in Case 1, while the

value comparison in Case 2 remains ambiguous: PA concentrates more water upstream and exploits return-flow reuse, whereas ES equalizes allocations but forgoes some reuse. Which effect dominates depends on θ and the curvature of $v(\cdot)$.

Observation 4. When priority is not perfectly aligned with stream position (e.g., if the most senior user is located in the middle of the stream or if junior and senior priorities are interleaved), the comparison between ES and PA no longer admits a simple ranking. The relative performance of the two institutions depends jointly on the size of return flows, the number of users, and the exact spatial ordering of seniority. Larger return flows tend to amplify the advantage of PA when senior users are upstream, because more water can be reused as it cascades downstream. By contrast, as the number of users increases, ES tends to capture more of the available water through guaranteed upstream diversions, which narrows the gap with PA. Because closed-form thresholds in the mixed-ordering cases are not tidy, we revisit the proposition in our numerical simulations.

Observation 5. With the exception of the polar case where inflow is just sufficient to satisfy a single senior right located furthest downstream (Case 1 under PA), all of the institutional mechanisms considered here generate total diversions that exceed the physical inflow $\tilde{W}$. This arises because upstream diversions return a positive fraction of water to the stream, allowing the same physical units to be re-used downstream.

2.4 Markets

So far, we have evaluated only the allocation rules themselves, each of which determines who may divert water, without any regard to value. In contrast, the social planner always allocates the next unit of water to the user with the greatest marginal value. Allowing users to voluntarily trade their allocations in a market can, in principle, approach the first best outcome by moving water from lower- to higher-valued uses.

Here, using the same three-user framework above, we examine when trading can mitigate these inefficiencies. For clarity, we drop the m subscripts for priority and simply denote users by their stream position $i \in \{1, 2, 3\}$. Our objective here is to highlight how stream position and return flows shape the outcomes of trade. We assume that water rights are fully transferable and that

users may sell the entirety of their allocated diversion quantity $\hat{w}_i$ to other users. As before, we focus on the extreme scarcity case where total inflow equals $\tilde{W} = x$, so there is only enough water in the stream to satisfy the most senior right.

Case 1: Downstream seniority (trade from $i = 3 \to i = 1$). Consider a transfer of $t > 0$ from the most downstream user ($i = 3$) to the most upstream user ($i = 1$) when initial allocations are $\hat{w}_1 = \hat{w}_2 = 0$ and $\hat{w}_3 = x$. A trade is socially efficient if the total value generated after the trade exceeds the value without the trade, or:

$$v_1(t) + v_2(\theta t) + v_3(x - t + \theta^2 t) > v_3(x). \tag{10}$$

The left-hand side includes (i) the value to user 1 from diverting t units, (ii) the value to user 2 from the return flow θt, and (iii) the value to user 3 from their remaining allocation plus the return flow from user 2, $\theta^2 t$.

The condition for a socially efficient trade differs from the privately efficient condition, i.e., the middle user is not involved in the transaction. If neither trader accounts for return flows, the trade occurs only if

$$v_1(t) + v_3(x - t) > v_3(x), \tag{11}$$

which compares the gains to user 1 with the seller's direct loss in value. If user 3 accounts for the direct return flow from user 1's diversion but not for benefits to user 2, the trade occurs only if

$$v_1(t) + v_3(x - t + \theta t) > v_3(x). \tag{12}$$

Here, the seller recognizes that some of the diverted water will return to them via θt, but still ignores the positive spillover to the middle user.

Because $v_2(\theta t) \geq 0$ for any $t > 0$ and $\theta \in (0, 1)$, the left-hand side of the socially efficient condition (Equation 10) is always at least as large as the partial-internalization case (Equation 12), which in turn is at least as large as the no-internalization case (Equation 11). This ordering means the set of trades satisfying the social condition is always weakly larger than the set satisfying the private conditions.

The "wedge" between socially efficient and privately profitable trades arises because the trading parties ignore part or all of the value created by return flows to non-participants. In Case 1, the middle user benefits from return flows generated by the trade, but these spillovers are not accounted for by the buyer or seller. The more intervening users between the trading parties, and the smaller the relevant θ values, the larger the uninternalized component of social benefit—and thus the greater the gap between socially efficient and privately chosen trades.

Case 2: Upstream seniority (trade from $i = 1 \to i = 3$). Consider a transfer of $t > 0$ from the most upstream user ($i = 1$) to the most downstream user ($i = 3$) when initial allocations are $\hat{w}_1 = x$, $\hat{w}_2 = \theta x$, and $\hat{w}_3 = \theta^2 x$. That is, downstream users are only eligible to divert the return flows from upstream users. Under PA, the upstream senior user may reduce their own diversion and directly deliver water to a downstream buyer, even if this reduces flows available to intermediate, junior users.

A trade is socially efficient if the total value generated after the trade exceeds the value without the trade:

$$v_1(x-t) + v_2\big(\theta(x-t)\big) + v_3\big(t + \theta^2(x-t)\big) > v_1(x) + v_2(\theta x) + v_3(\theta^2 x). \tag{13}$$

The left-hand side includes (i) the reduced value to user 1 from keeping $x - t$ units, (ii) the value to user 2 from the reduced return flow from user 1, and (iii) the value to user 3 from the purchased water t plus the remaining return flow from user 2. The right-hand side is the total value without trade. Let $v_2\big(\theta(x-t)\big) - v_2(\theta x) = \gamma$, so Equation 13 can be rewritten as

$$v_1(x-t) + v_3\big(t + \theta^2(x-t)\big) + \gamma > v_1(x) + v_3(\theta^2 x). \tag{14}$$

If the trade is evaluated privately between users 1 and 3, user 2's loss is ignored. The trade occurs if

$$v_1(x-t) + v_3\big(t + \theta^2(x-t)\big) > v_1(x) + v_3(\theta^2 x), \tag{15}$$

which compares the change in value to the seller and buyer only. The buyer still internalizes the return flow from the seller's remaining diversion, $\theta^2(x-t)$, but the loss of value from the reduction

in flow to the intermediate user is not considered.

Because $v_2\big(\theta(x-t)\big) < v_2(\theta x)$ for $t > 0$ (γ is negative), the left-hand side of the socially efficient condition (Equation 14) is always less than the left-hand side of the private condition (Equation 15). This ordering means that the set of trades satisfying the private condition is always weakly larger than the set satisfying the social condition. The wedge between these cases arises because the trading parties do not compensate the intermediate user for the reduced availability of return flows. The larger the marginal value $v_2'(\cdot)$ at θx and the greater the reduction in return flow θt, the larger the gap between socially efficient and privately chosen trades.

Observation 6. In general, positive return flows to a non-trading intermediate user create a wedge between the social-efficiency and private-profitability conditions. Social efficiency accounts for the gains to all affected users, whereas private profitability accounts only for the trading parties. As shown in Case 1, this wedge arises because intermediate users may benefit from return flows generated by the trade (a positive externality), while in Case 2 the wedge arises because an intermediate user may lose return flows (a negative externality). The magnitude and direction of these wedges depend on stream position, return flow parameters, and the placement of seniority. We return to this proposition in our numerical simulations, which allow for n users.

Observation 7. Both priority- and share-based allocation mechanisms are subject to the same fundamental limitation in market settings. With zero transaction costs, the market reallocates water so that marginal values are equalized across participants. Under these assumptions, the only channel through which institutions affect final basin value is the total volume of diversions available prior to trade. Institutions that allocate more water upstream, and therefore generate greater return flows, expand the set of tradable water and yield higher post-market total value.

2.4.1 Consumptive use

For the cases above, water rights and eligible trades are defined in terms of diversion quantities. Several previous studies suggest that economic efficiency may be improved if such rights and trades are instead defined in terms of consumptive use (Huffaker et al., 2000; Leonard et al., 2019; Pérez-Blanco et al., 2020; Wobus et al., 2022). We explore this formally in the Appendix (A.3) and show

that some market failures arising from return flows can persist under both definitions of rights. In practice, the seller (buyer) must divert less (more) than the traded consumptive volume to account for conveyance losses or other inefficiencies (e.g., irrigation technology), and these diversions still generate return flows that affect non-participants. Intuitively, if c_i denotes consumptive use for user i and $\eta_i \in (0, 1]$ is their diversion/application efficiency, then a trade of c units of consumptive use requires a change in diversion of $t = c/\eta_i$ where $t > 0$ can be expressed in the same way as in the diversion-based cases above. The remaining $(1 - \eta_i)t$ is returned to the stream as potential externalities, where $(1 - \eta_i)$ can be redefined as θ_i.

2.4.2 Trades in practice

In practice in the WUS, full transfers or temporary leases in PA systems are generally subject to review by a state engineer or equivalent water authority. The review focuses on preventing injury to other right holders, meaning that trades cannot legally reduce the quantity or timing of water available to more senior users. This system is designed to avoid negative externalities to senior users but does not create any mechanism for recognizing or incentivizing the positive externalities that may accrue to intermediate users through return flows. In Case 1–type situations, this means the law prevents injury but not missed opportunity: where a trade that would create positive spillovers for non-trading users is foregone because those gains are not internalized by the trading parties. As a result, parameter regions exist in which a trade would be socially efficient (because it generates positive spillovers to non-trading users) but is not privately profitable for the trading parties, and thus does not occur. Conversely, trades like Case 2, which impose uncompensated costs on junior intermediate users, can be privately profitable and legally permissible, provided that no senior user is injured.[4] The net effect is that actual trading outcomes in leasing markets may deviate systematically from the socially optimal pattern of trades described in the theoretical cases above.

While the stylized cases above isolate the mechanisms by which return-flow externalities create wedges between socially efficient and privately profitable trades, they do not exhaustively enumerate all possible trading configurations. We choose to focus on bilateral trades between two parties to illustrate how real-world market outcomes can diverge from the social optimum when return-flow

[4]Some states in the WUS do however allow private parties to protest trades (Chong and Sunding, 2006).

benefits or costs to non-participants are unpriced. Modern trades may involve more complex exchanges or sequential reallocations, but the central insight carries over: without a mechanism to internalize these spillovers, market activity may fail to realize the full potential gains from trade.

2.5 Analytical ambiguities and propositions

The polar cases outlined above provide analytical insight into how institutional rules interact with the physical characteristics of river systems—stream-priority positioning and return flows—to shape both allocative efficiency and subsequent market outcomes. However, as the number of users increases, closed-form analysis becomes unwieldy and clear rankings break down. Thresholds at which one mechanism begins to outperform another depend jointly on value functions, return flow parameters, the spatial distribution of seniority, and the degree of scarcity. For example, how does relative efficiency change as n grows large? How much seniority must be concentrated upstream for PA to perform well? What happens when scarcity is not extreme but merely low? These kinds of questions motivate our simulation exercises, and the propositions below summarize hypotheses suggested by the analytical framework that we will test with a numerical approach.

Proposition 1 (Scale with n). With strictly increasing, concave $v_i(\cdot)$ and no trading, share-based mechanisms (PS/ES) tend to generate higher total basin value than PA as the number of users n increases. The intuition is that PS/ES guarantee a positive allocation to more users, placing additional units on the steep (high–marginal value) portion of $v_i(\cdot)$, whereas PA leaves some users at zero. Additionally, with trading, the amount of water that must be reallocated to approach the efficient outcome is largest under PA and smallest under ES, because PA's initial allocations are farthest from equalized marginal values.

Proposition 2 (Priority–location misalignment). As the correlation between stream position and priority declines, PA captures fewer return flows, lowering total diversions pre-market. In the market, this loss further widens the wedge between the privately profitable and socially efficient outcome (defined by the first best allocation under the SP).

Proposition 3 (Shortage amplifies PA losses). As shortage increases (defined as the sum of unment diversion entitlements), PA performs worse relative to share rules: junior users are fully

curtailed more often, so less water is diverted and reused downstream.

3 Numerical models

In this section, we extend the linear waterway to accompany n users and present optimization frameworks that emulate the water allocation institutions discussed above.[5] We compare four institutional mechanisms—prior appropriation, proportional sharing, equal sharing, and leasing markets—against the benchmark social planner solution. The leasing market mechanism is applied as a secondary layer, allowing users to trade allocations assigned under each of the prior mechanisms. We then explore how variation in key parameters affects the magnitude of the disparities in efficiency and total diversions highlighted by the analytical models.

3.1 Water allocation mechanisms

3.1.1 Social planner

To establish a benchmark for institutional performance, we begin with the social planner (SP) problem, which represents the first-best allocation of water in the absence of institutional or informational constraints. The SP solution serves as an upper bound on economic efficiency, against which the performance of other institutions can be compared. The objective of the SP is to maximize the total value of production within a river basin by allocating water w_i to users $i = 1, ..., n$ with total value functions $v_i(w_i)$, where i indexes stream position from furthest upstream (1) to furthest downstream (n). Available water in a given year is a random draw $\tilde{W} \in \boldsymbol{W}$, where $\boldsymbol{W}$ is the distribution of possible water supply conditions. Some proportion of water $\theta \in (0, 1)$ diverted to the upstream user returns to the stream and becomes available for the downstream user as a

[5] Stylized two-user cases for each institutional mechanism are also presented in the Appendix (subsection A.1).

return flow. The SP's deterministic maximization problem in a given year is thus:

$$\begin{aligned} \max_{w_1,\ldots,w_n \geq 0} \quad & \sum_{i=1}^{n} v_i(w_i) & & \\ \text{s.t.} \quad & w_1 \leq \tilde{W} & & \\ & w_2 \leq \tilde{W} - (1-\theta_1)w_1 & & \\ & \vdots & & \\ & w_n \leq \tilde{W} - \textstyle\sum_{i=1}^{n-1}(1-\theta_i)w_i. & & \end{aligned} \qquad \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \text{Physical constraints} \tag{16}$$

There are n constraints, each imposing the maximum amount of water that can be physically delivered to the i^{th} user based on the location and return flow efficiency of upstream users. Importantly, return flows generated from user i are available to all downstream users.

3.1.2 Prior appropriation

We emulate the institution of PA using a constrained linear programming problem that sequentially allocates water based on user priority. This optimization problem will hereinafter be referred to as the PA allocation function:

$$\begin{aligned} \max_{w_{1,m},\ldots,w_{n,m} \geq 0} \quad & \sum_{i=1}^{n} \beta_{i,m} w_{i,m} & & \\ \text{s.t.} \quad & w_{1,m} \leq \tilde{W} & & \\ & w_{2,m} \leq \tilde{W} - (1-\theta_1)w_{1,m} & & \\ & \vdots & & \\ & w_{n,m} \leq \tilde{W} - \textstyle\sum_{i=1}^{n-1}(1-\theta_i)w_{i,m} & & \\ & w_{i,m} \leq R_{i,m} \;\; \forall i \in \{1,\ldots,n\}. & & \end{aligned} \qquad \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \text{Physical constraints} \\ \left.\begin{array}{l} \\ \end{array}\right\} \text{Institutional constraints} \end{array} \tag{17}$$

Here, $m \in \{1,\ldots,M\}$ indexes user priority, with $m = 1$ denoting the most senior right. The coefficient $\beta_{i,m}$ is used to enforce priority ordering in the solution: it must be sufficiently large such that users with higher priority (lower m) are allocated water before junior users. $R_{i,m}$ denotes the maximum quantity user i is legally entitled to divert under their water right. The institutional constraints ensure no user is allocated more than their entitlement, while the physical constraints

(identical to those in the SP problem) impose hydrologic feasibility.

Note that i still indexes stream position, but it is independent of the user's priority rank m. For example, the most senior right may belong to any stream position, not necessarily the most upstream or downstream. In the simulation section that follows, both i and m are randomly assigned to users to analyze how efficiency varies across different stream-priority configurations. Once the vector of allocations $\hat{w}_{i,m}$ is obtained from the PA allocation function, the total value generated can be evaluated as $\sum_{i=1}^{n} v_i(\hat{w}_{i,m})$.

3.1.3 Proportional sharing

As discussed prior, proportional sharing (PS) rules take many different forms in practice. Here, we adopt a system that acts as a hybrid between strict priority and equal sharing. Suppose users under a PA system elect to pool their risk and assign shares according to priority levels. Under this PS system, water is allocated proportionally according to each user's number of shares until all endowments are satisfied, in descending order from the user with the most shares. To formalize this, suppose shares are defined as a decreasing function of priority rank m, so that more senior users hold more shares but junior users are not left at zero. One convenient formulation is:

$$\gamma_{i,m} = \sum_{j=m}^{n} \frac{1}{j}, \tag{18}$$

$$\delta_{i,m} = \frac{\gamma_{i,m}}{\sum_{m=1}^{n} \gamma_{i,m}}, \tag{19}$$

where $\delta_{i,m}$ denotes the normalized share of user i with seniority rank m. This weighting scheme creates a declining but non-zero share profile across users.[6] Conceptually, this version of PS lies between the polar cases of PA (strict hierarchy) and ES (no hierarchy), and provides a natural way to examine how outcomes evolve when institutions gradually move from one extreme to the other.

The PS mechanism can be modeled with an iterative linear optimization approach. Again, let $R_{i,m}$ be the water right that determines the maximum amount of water allowable for user i with priority m. PS can be emulated as follows:

[6]For example, when $n = 4$, the shares implied by Equation 19 are $\delta_1 = 0.5208$, $\delta_2 = 0.2768$, $\delta_3 = 0.1458$, and $\delta_4 = 0.0625$, meaning that senior users have substantially more weight while junior users still retain some guaranteed allocation.

1. Calculate the minimum amount of water necessary, $\grave{w}_1$, to satisfy the maximum allowable quantity for the highest priority user ($R_{i,1}$) while abiding by the physical and institutional constraints below.

$$\begin{aligned} \min_{\grave{w}} \quad & \grave{w} \\ \text{s.t.} \quad & \left.\begin{aligned} & w_{1,m} \leq \grave{w} \\ & w_{2,m} \leq \grave{w} - (1-\theta_1)w_{1,m} \\ & \vdots \\ & w_{n,m} \leq \grave{w} - \textstyle\sum_{i=1}^{n-1}(1-\theta_i)w_{i,m} \end{aligned}\right\} \text{Physical constraints} \\ & \left.\begin{aligned} & w_{i,1} = R_{i,1} \\ & w_{i,m} = \delta_{i,m}\left(\textstyle\sum_{j=1}^{n} w_j\right) \end{aligned}\right\} \text{Institutional constraints} \end{aligned} \tag{20}$$

2. If $\tilde{W} \leq \grave{w}_1$, use the following maximization problem to determine allocations.

$$\begin{aligned} \max_{w_{1,m},\ldots,w_{n,m}\geq 0} \quad & \sum_{i=1}^{n} w_{i,m} \\ \text{s.t.} \quad & \left.\begin{aligned} & w_{1,m} \leq \tilde{W} \\ & w_{2,m} \leq \tilde{W} - (1-\theta_1)w_{1,m} \\ & \vdots \\ & w_{n,m} \leq \tilde{W} - \textstyle\sum_{i=1}^{n-1}(1-\theta_i)w_{i,m} \end{aligned}\right\} \text{Physical constraints} \\ & \left.\begin{aligned} & w_{i,m} \leq R_{i,m} \\ & w_{i,m} = \delta_{i,m}\left(\textstyle\sum_{j=1}^{n} w_j\right) \end{aligned}\right\} \text{Institutional constraints} \end{aligned} \tag{21}$$

   If $\tilde{W} > \grave{w}_1$, store the allocations from the minimization problem in step 1 in vector $\hat{\boldsymbol{w}}_{\mathbf{1}} = (\hat{w}_{1,m}, \hat{w}_{2,m}, ..., \hat{w}_{n,m})$ and proceed to step 3.

3. Subtract $\hat{\boldsymbol{w}}_{\mathbf{1}}$ from the vector of maximum allowable quantities $\boldsymbol{R}$ to get the remaining amount of water each user requires to satisfy their water right, $\hat{\boldsymbol{R}}_{\mathbf{1}} = (R_{1,m} - \hat{w}_{1,m}, R_{2,m} - \hat{w}_{2,m}, ..., R_{n,m} - \hat{w}_{n,m})$. Subtract $\grave{w}_1$ from $\tilde{W}$ to get the remaining amount of water in the stream, $\tilde{W}_1$. Set $\delta_{i,1} = 0$ since $R_{i,1} - \hat{w}_{i,1} = 0$, and generate new shares and proportions using Equation 19 for $i-1$ users. Calculate the minimum amount of water necessary, $\grave{w}_2$, to satisfy

the maximum allowable quantity for the second highest priority user ($R_{i,2} - \hat{w}_{i,2}$), who now holds the largest share, using the same minimization problem from step 1 with the updated parameters.

4. If $\tilde{W}_1 \leq \grave{w}_2$, use the maximization problem from step 2 with updated parameters to determine second stage allocations and add them to $\hat{\boldsymbol{w}}_{\mathbf{1}}$ to get final allocations. If $\tilde{W}_1 > \grave{w}_2$, repeat step 3 with updated parameters.
5. Repeat the process until there is no water remaining or all water rights are satisfied.

The iterative approach to PS allocation provides a final vector of water delivered to each user. This vector is the sum of all allocations across each iteration, typically culminating in the maximization problem in step 2 when there is not enough water left to satisfy the next highest-priority user. In essence, the PS allocation process always diverts the next available unit of water in proportion to each user's number of shares while abiding by the physical constraints of the waterway.

3.1.4 Equal sharing

Equal sharing (ES) represents the limiting case of a share-based allocation system in which all users receive the same amount of water, regardless of stream position or priority. This mechanism can be interpreted as an edge case of the PS framework where all users hold equal shares, i.e., $\gamma_i = 1$ for all i, implying $\delta_i = 1/n$. Like the other institutions, ES can be represented as a constrained linear optimization problem. The ES allocation function is defined as:

$$
\begin{aligned}
\max_{w_1,\ldots,w_n \geq 0} \quad & \sum_{i=1}^{n} v_i(w_i) & & \\
\text{s.t.} \quad & \left.\begin{aligned} & w_1 \leq \tilde{W} \\ & w_2 \leq \tilde{W} - (1-\theta_1)w_1 \\ & \vdots \\ & w_n \leq \tilde{W} - \textstyle\sum_{i=1}^{n-1}(1-\theta_i)w_i \end{aligned}\right\} & & \text{Physical constraints} \\
& \left.\begin{aligned} & w_i \leq R_i \ \ \forall i \in \{1,\ldots,n\} \\ & w_i = w_j \ \ \forall i \neq j. \end{aligned}\right\} & & \text{Institutional constraints}
\end{aligned}
\tag{22}
$$

The institutional constraint $w_i = w_j$ for all $i \neq j$ enforces equal allocations across users, while the physical constraints ensure hydrologic feasibility. This system removes any notion of priority and allows for a direct comparison of efficiency losses when hierarchy is entirely abandoned.

3.1.5 Leasing markets

The mechanisms described above span a spectrum of priority-based allocation rules: from full priority (PA) to partial priority (PS) to no priority (ES). They each assign water according to institutional rules rather than by equalizing marginal values across users. This introduced inefficiency can be remedied with a competitive leasing market by allowing users to trade water. Those with lower-valued uses can sell to those who generate higher value, leading to a more efficient allocation.

The following model emulates a leasing market in which users are allowed to trade water *after* allocations are determined from the PA, PS, or ES allocation functions. In practice, annual water supplies in the WUS within a basin are often estimated based on winter snowpack, so the initial allocation can be interpreted as an expected allocation for the given year. To proxy the outcome of a perfectly competitive market with no transaction costs, the model is structured as a Social Planner's problem with additional institutional constraints. Let $\hat{w}_i$ be the initial allocation for user i, who may choose to buy or sell an amount t_i:

$$
\begin{aligned}
\max_{t_1,\ldots,t_n \geq 0} \quad & \sum_{i=1}^{n} v_i(\hat{w}_i + t_i) & & \\
\text{s.t.} \quad & \left.\begin{aligned}
& \hat{w}_1 + t_1 \leq \tilde{W} \\
& \hat{w}_2 + t_2 \leq \tilde{W} - (1-\theta_1)(\hat{w}_1 + t_1) \\
& \vdots \\
& \hat{w}_n + t_n \leq \tilde{W} - \textstyle\sum_{i=1}^{n-1}(1-\theta_i)(\hat{w}_i + t_i)
\end{aligned}\right\} && \text{Physical constraints} \\
& \left.\begin{aligned}
& \hat{w}_i + t_i \geq 0 \;\; \forall i \in \{1,\ldots,n\}
\end{aligned}\right\} && \text{Non-negative allocations} \\
& \left.\begin{aligned}
& \textstyle\sum_{i=1}^{n} t_i = 0
\end{aligned}\right\} && \text{Market clearing condition.}
\end{aligned}
\tag{23}
$$

If $t_i > 0$, user i can be considered a net buyer and if $t_i < 0$, a net seller. The first set of constraints ensure that user i can only buy water up to an amount such that $\hat{w}_i + t_i$ is less than or equal to the total amount of water in the river at their point of diversion. The second set of constraints ensures

that no user can sell more water than they were initially allocated. The last constraint is a market clearing condition that ensures that the sum of all water bought is equal to the sum of all water sold. The solution to the water market optimization problem, $\hat{\boldsymbol{t}}$, can then be added to the initial vector of allocations $\hat{\boldsymbol{w}}$ to get the final allocations.

Importantly, the leasing market in Equation 23 emulates a system in which only privately profitable trades occur, as described in subsection 2.4. Trades take place only when they increase the joint value of the trading parties, without compensating or accounting for spillover effects on non-participants. Because users can only exchange their existing allocations, no new return flows are created through trading; the physical-feasibility and market-clearing constraints ensure that all transactions respect the initial allocation and return-flow patterns determined by the underlying allocation rule (PA, PS, or ES). As a result, the outcomes directly reflect the same wedges between private profitability and social efficiency as benchmarked by the social planner.

3.2 Simulations

To evaluate institutional performance under varied stream compositions, we implement a series of randomized simulations that emulate stylized river basins of increasing size. The number of users (n) ranges from 3 to 36, representing increasing population growth or competing demands.[7] For each basin size, we generate 10,000 unique simulations of basin characteristics and water availability, holding these fixed across institutional regimes to allow for direct comparison. The purpose of this design is to assess not only the average efficiency of each institution but also its robustness to diverse and sometimes adverse hydrologic configurations. Randomization ensures that institutional rules are tested within a broad spectrum of possible stream compositions—upstream versus downstream priority alignments, heterogeneous return flows and value generation, and varying scarcity levels—so that we can identify the conditions under which each institution performs well or performs poorly.

Each simulation involves drawing a total water supply $\tilde{W}$ from a log-normal distribution that scales with the number of users. Users are randomly assigned stream positions and priority numbers without replacement, and each is endowed with a maximum diversion right of 500 units. Return

[7] The upper bound of 36 is chosen somewhat arbitrarily; it reflects the point at which numerical solvers in `R` begin to experience convergence issues and fail to reliably identify solutions.

flow coefficients (θ_i) are drawn from the set $\{0.1, 0.2, ..., 0.9\}$ with replacement, while productivity is defined by a random value coefficient (v_i) drawn from a normal distribution and embedded in the concave production function $v_i(w_i) = v_i \cdot w_i^{1/2}$, which forces diminishing marginal returns to water use. From each user's assigned priority, share weights are derived for the PS mechanism. These elements (see Table 1) fully define the physical and economic structure of the basin for a given simulation. Table 2 offers an example parameter set for $n = 10$.

Table 1: Detailed parameters for water supply simulations

Parameter	**Sample Set**	**Replacement**
Water Supply ($\tilde{W}$)	Log-normal distribution	With
User Endowments (R)	500	N/A
Number of Users (n)	3 to 36 in increments of 1	N/A
Stream Position (i)	1 to n	Without
Priority Number (m)	1 to n	Without
Priority Coefficient (β)	20^{-m}	N/A
Return Flow Coefficient (θ)	0.1, 0.2, 0.3, ..., 0.9	With
Value Coefficient (v)	Normal distribution	With
Value Function	$v \cdot w^{\frac{1}{2}}$	N/A

Critically, institutional rules are the only element that varies within each iteration of the simulation. That is, for each of the 10,000 generated basins at each n, water is allocated under all six institutional regimes—PA, PS, ES, and their respective leasing market variants: PAM, PSM, and ESM, respectively. This structure ensures that differences in outcomes across mechanisms are attributable solely to institutional constraints, not to underlying differences in basin conditions. All simulations were conducted in `R` software, and the optimization problems were solved using the 'lpSolve' (Csárdi and Berkelaar, 2024) and 'nloptr' (Ypma et al., 2025) packages.

Table 2: Parameters for a sample simulation with ten users and $\tilde{W} = 2,252$

Stream position (i)	1	2	3	4	5	6	7	8	9	10
Priority no. (m)	5	6	4	10	7	8	2	1	3	9
Share coef. (δ)	0.085	0.065	0.11	0.01	0.048	0.034	0.193	0.293	0.143	0.021
Value coef. (v)	68	738	435	225	950	454	121	592	321	534
Return flow coef. (θ)	0.4	0.1	0.8	0.6	0.7	0.8	0.2	0.3	0.7	0.3
Endowment (R)	500	500	500	500	500	500	500	500	500	500

4 Results

Proposition 1 (Scale with n). Figure 1 displays the average total value generated (panel A) and average share of water reallocated in a leasing market (panel B) across all simulations by institution and number of users.[8] Panel A presents a clear trend: in the absence of a leasing market, share-based mechanisms (PS and ES) outperform PA in terms of average total value created, with the disparity widening as the user base expands. This makes sense intuitively (and mechanistically) under an assumption of diminishing marginal returns to water, since the share-based mechanisms ensure all users receive at least some allocation. With a leasing market in place, as indicated by the overlapping red, yellow, and orange curves, total average value is effectively equivalent across institutions, which aligns with the principle that the efficiency of resource allocation should not depend on the initial allocation of resources. Nonetheless, panel B provides a nuanced perspective on this narrative. Despite achieving similar levels of total value, the proportion of water that needs reallocation under PAM is substantially higher—almost double—compared to ESM. The need for significant reallocation in the PAM system to attain efficiency relative to the ESM system highlights an often-overlooked dimension of bargaining: the initial distribution of rights has directly drives the volume of resources that must be reallocated to achieve efficiency. A more efficient initial allocation, as evidenced in the share-based systems, can reduce the need for reallocations, thereby minimizing transaction costs and potential barriers to achieving optimal resource use.

[8] The share of water reallocated is calculated as the total volume traded, $\sum \hat{t}$, divided by the total water diverted, $\sum \hat{w}$, in each simulation.

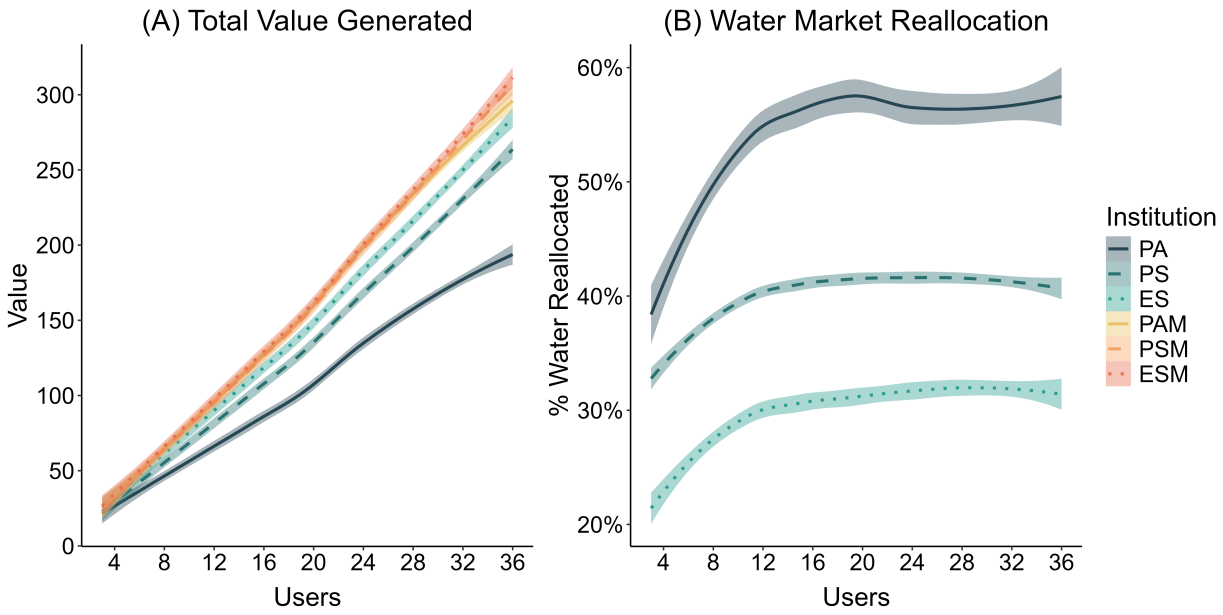


Figure 1: **Impacts of water allocation institutions on value generation and market activity.** Panel (A) shows a smoothed conditional mean curve of total value generated across all simulations for each institutional mechanism, as the number of users increases from 3 to 36. Panel (B) displays the same curve for the percentage of initial water allocations that is reallocated through leasing markets.

Proposition 2 (Priority—location misalignment). To further scrutinize the effect of priority locations along a waterway, we examine how stream composition influences both total water diverted and efficiency relative to the SP. Due to return flows, the SP typically allocates more water upstream to equate the marginal benefits of the last units of water diverted to each user (see subsection A.2). However, as demonstrated in the analytical models, a senior user downstream of a junior user can result in less overall diversions when basin supplies are insufficient to satisfy all users, because junior users cannot divert water if doing so would cause injury to senior users.[9] So, for each randomly generated basin composition, we calculate a Pearson correlation coefficient ρ to examine the strength and direction of the relationship between stream position and user priority under PA. For example, consider a case with $n = 3$ users where the priorities assigned (in order of stream position) are $1, 2, 3$. Here, the stream position correlates perfectly with the user priorities, yielding a correlation coefficient of $\rho = 1$. In contrast, if the user priorities are assigned as $3, 2, 1$, the stream position correlates perfectly negatively with the user priorities, resulting in a correlation coefficient

[9]This is also simulated for the two-user case in subsection A.1 and Table A1.

of $\rho = -1$. In essence, a positive correlation would suggest that users' priorities are aligned with the natural flow of the stream, i.e., upstream users have higher priority. Conversely, a negative correlation implies that high-priority rights are downstream.

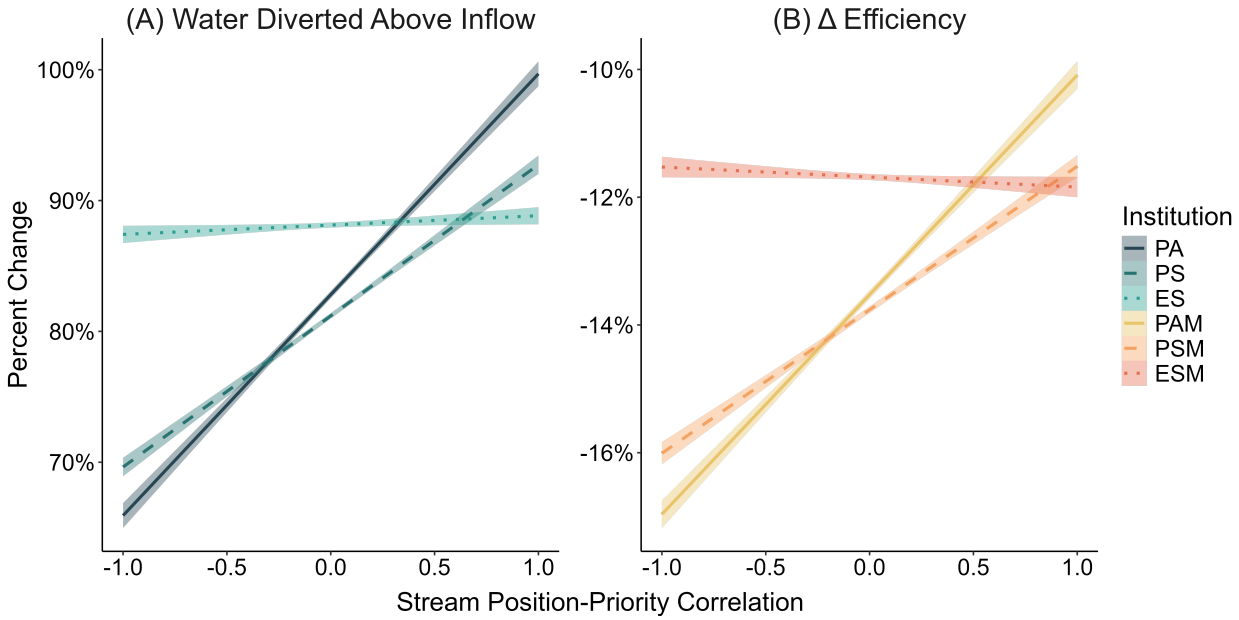


Figure 2: **Effect of stream–priority alignment on total diversions and efficiency.** Panel (A) shows the relationship between the correlation coefficient ρ—which measures alignment between user priority and stream position—and the total volume of water diverted relative to inflow ($\tilde{W}$). Panel (B) shows the corresponding change in efficiency relative to the Social Planner benchmark.

Figure 2 displays the relationship between stream–priority alignment and two key outcomes: the total volume of water diverted to all users relative to inflows (panel A) and the change in efficiency of the leasing markets relative to the Social Planner (panel B).[10] The plotted relationships are obtained by regressing ρ on these percentage changes across all simulations. For the ES mechanism, stream-priority alignment has no impact on the amount diverted, as the institutional constraint in Equation 22 forces all users to divert the same volume. The PA mechanism exhibits the steepest slope, such that with perfect alignment ($\rho = 1$), average diversions are nearly twice the natural inflow. When alignment is low, PA's downstream senior users can constrain upstream junior diversions, sharply reducing total water use. The PS mechanism maintains some degree of priority hierarchy, placing it between ES and PA in both magnitude and slope. Panel B mirrors the

[10]The percent change in total diversions is calculated as $(\sum \hat{w}/\tilde{W} - 1) \times 100$, where $\sum \hat{w}$ is the total volume diverted and $\tilde{W}$ is the natural inflow. The percent change in efficiency is calculated as $(V_{\text{inst}} - V_{\text{SP}})/V_{\text{SP}} \times 100$, where V_{inst} is the total value under each institution and V_{SP} is the total value under the Social Planner.

pattern in panel A almost exactly: efficiency losses relative to the Social Planner track the change in total diversions, indicating that these losses are driven almost entirely by institutional constraints that limit total diversions and thus reduce the water available for trade. Together, Figure 1 and Figure 2 show that the spatial distribution of priority rights not only shapes the market activity required to reach the social optimum, but also directly determines the total water that the system can harness.

Proposition 3 (Shortage amplifies PA losses). Figure 1 demonstrates that each of the market mechanisms generates the same total value on average, but this equivalence does not hold for all stream compositions. As shown in Figure 2, the total amount of water diverted depends on the spatial distribution of priorities. Accordingly, Figure 3 presents the distribution of total value generated (relative to the SP) across four categories defined by stream-priority alignment—high priority downstream ($\rho \leq -0.5$) versus high priority upstream ($\rho \geq 0.5$)—and shortage severity, with high and low shortage defined as the top and bottom 25^{th} percentiles of shortage per user.[11] The top four panels (A) show results without markets. Here, PA is particularly sensitive to both priority location and shortage level: distributions shift left (larger efficiency losses) when high-priority users are downstream, and flatten as shortage severity increases. Share-based mechanisms display less sensitivity in these dimensions. The bottom four panels (B) show results with markets. All distributions shift right relative to the no-market case (indicating increased efficiency), yet PAM remains sensitive to stream–priority location and shortage severity. Notably, PAM performs best under conditions of high priority upstream and low shortage, where it achieves the smallest efficiency losses among all market mechanisms. Because these simulated markets have zero transaction costs, differences in efficiency must stem from differences in total water traded (despite the same inflow). This reinforces the point that the location of priority rights can directly influence the total volume of water harnessed by the system and that such effects may be amplified during extreme shortages. Indeed, ES *without* a market still outperforms PAM in 17.5% of simulations, with these cases averaging a stream–priority correlation of -0.2 and shortage per user of 180 units (the 27^{th} percentile).

[11] Shortage per user is calculated as the minimum water required to satisfy all endowments minus the natural inflow $\tilde{W}$, divided by the number of users.

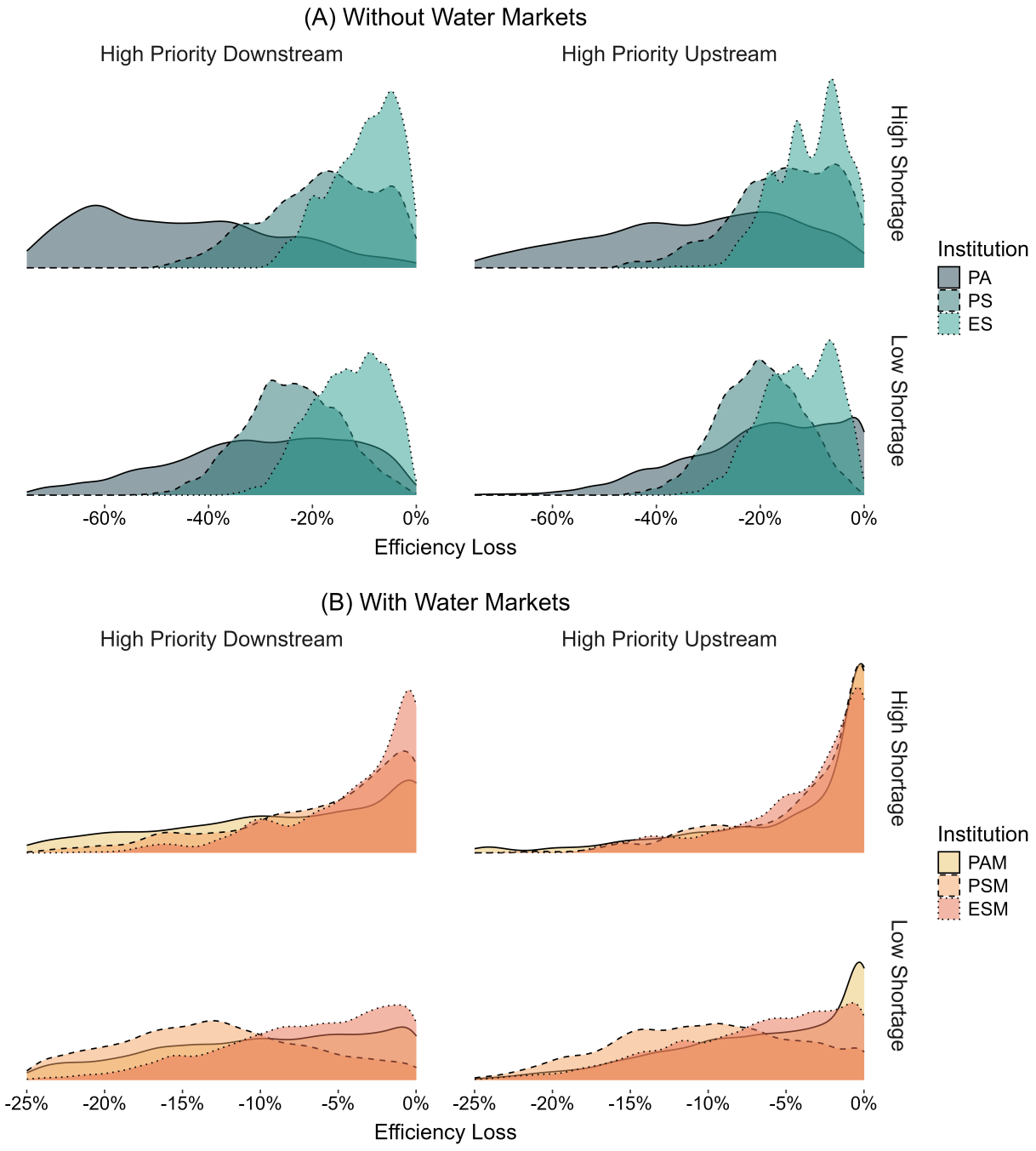


Figure 3: **Distributions of efficiency losses under different institutions by stream–priority alignment and shortage severity.** Each panel shows the distribution of efficiency loss (total value relative to the Social Planner) for a subset of simulated basins, categorized by stream–priority correlation and average shortage. Panel (A) shows results for mechanisms without markets, and Panel (B) shows results with markets. Columns distinguish whether users with high-priority rights are located downstream ($\rho \leq -0.5$) or upstream ($\rho \geq 0.5$). Rows separate basins by shortage severity: high shortage (top 25%) and low shortage (bottom 25%), based on inflow shortfall per user.

5 Discussion

While water markets can dramatically improve allocative efficiency, the results above show that due to institutional rules, some inefficiencies persist even under ideal market conditions. The lower performance of PAM, despite zero transaction costs and full participation, suggests that the problem is not the absence of mutually beneficial trades, but rather that private trades do not internalize the positive externalities of return flows to non-participating users, which can distort the sequence of reallocations and lead to outcomes below the social optimum. The leasing market formulation in Equation 23 represents the Coasean ideal to a degree: clearly defined property rights, no transaction costs, and a competitive market that reallocates water to its highest-valued uses. In such a setting, the initial allocation of rights should not affect efficiency. However, the cases and simulations above reveal several reasons why this prediction fails for priority based allocations.

First, the physical nature of water and the structure of priority-based allocation mechanisms imply that the total quantity of water available for reallocation may be endogenous to the initial distribution of rights. When high-priority rights are held downstream, upstream junior users may be unable to divert and generate return flows, reducing the total volume of water in the system available for use. Second, markets in this setting may equalize marginal values among trading parties but are not necessarily Pareto optimal. Because return flows are not owned or tradable, part of the marginal benefit or cost of use accrues to non-trading intermediate users. This leads to wedges between the social-efficiency and private-profitability conditions: positive externalities may be forgone because they are not internalized, and negative externalities may persist because they are not prohibited unless they injure a senior user. Third, this incomplete specification of property rights over return flows means that the relevant externalities fall outside the bargaining set. Finally, the above considerations still assume negligible transaction costs. In reality, leasing markets in the WUS face substantial costs from search, negotiation, legal review, and conveyance. Womble and Hanemann (2020) estimate that average transaction costs in Colorado amount to 35.7% of the market price plus transaction costs, and it has been suggested that less than 1% of withdrawn freshwater is traded worldwide (Rafey, 2023). The results presented in Figures 1-3 show that PA allocations can be highly inefficient *ex ante*, particularly during drought, with some junior users receiving no water and some senior users using water at very low marginal benefit. Because

so much water must change hands in PA to approach the social optimum, even modest frictions sharply limit the scope for efficiency gains.

6 Conclusion

This paper examines how water allocation institutions and markets perform under the hydrologic realities of river systems, with particular attention to how user location, shortages, and return flows influence allocative efficiency. We first use an analytical framework to show how institutional rules and user placement interact to shape efficiency outcomes and derive conditions under which privately profitable trades diverge from socially efficient ones in contemporary leasing markets. We then use numerical optimization methods that replicate the rules and constraints of prior appropriation and share-based allocation mechanisms, and use Monte Carlo simulations to assess performance across a wide range of basin compositions and scarcity conditions.

The analysis makes several simplifying assumptions that define its scope. While parameter choices in the simulations are necessarily stylized, the findings reflect general relationships rather than reliance on any particular calibration. Several important dynamics are not modeled: the role of storage infrastructure, more complex and realistic value functions, and dynamic trading where users update expectations based on observed return flows from prior trades. These extensions could further refine the results and explore additional drivers of efficiency, but they are beyond the scope of this study. The objective here is to characterize the sources of market inefficiency that arise from the interaction between allocation rules and the movement of water, rather than to produce basin-specific policy prescriptions.

The results highlight that the efficiency of any allocation institution depends on the hydrologic connectivity of users. Prior appropriation with an active leasing market can perform well—sometimes achieving near-socially optimal outcomes—when high-seniority water rights are held by upstream users whose return flows benefit downstream appropriators. In many other configurations, however, share-based mechanisms tend to yield both more equitable and efficient allocations by mechanistically internalizing the value of return flows. For policymakers, these findings underscore that institutional design cannot be divorced from basin hydrology: the same market rules can produce different outcomes depending on user location, connectivity, and return flows. Accounting

for these factors in market design and governance is essential for aligning private incentives with socially efficient water use.

Acknowledgments

This material is based upon work supported by the National Institute for Food and Agriculture under Award No. 2018-69011-28369 and the National Science Foundation under Grant No. 1828902.

References

Blumberg, J. and Warziniack, T. (2025), Drivers of us water use and implications for water scarcity and climate change, Preprint SSRN 5123252.
URL: *https://ssrn.com/abstract=5123252*

Brewer, J., Glennon, R., Ker, A. and Libecap, G. (2006), 'Transferring water in the american west: 1987-2005', *U. Mich. JL Reform* **40**, 1021.

Brewer, J., Glennon, R., Ker, A. and Libecap, G. (2008), '2006 presidential address water markets in the west: prices, trading, and contractual forms', *Economic Inquiry* **46**(2), 91–112.

Brown, T. C., Mahat, V. and Ramirez, J. A. (2019), 'Adaptation to future water shortages in the united states caused by population growth and climate change', *Earth's Future* **7**(3), 219–234.

Bruno, E. M. and Jessoe, K. (2024), 'Designing water markets for climate change adaptation', *Nature Climate Change* **14**(4), 331–339.

Burness, H. S. and Quirk, J. P. (1979), 'Appropriative water rights and the efficient allocation of resources', *The American Economic Review* **69**(1), 25–37.

Burness, H. S. and Quirk, J. P. (1980), 'Water law, water transfers, and economic efficiency: The colorado river', *The Journal of Law and Economics* **23**(1), 111–134.

Chong, H. and Sunding, D. (2006), 'Water markets and trading', *Annu. Rev. Environ. Resour.* **31**(1), 239–264.

Coase, R. (1960), 'The problem of social cost', *Journal of Law and Economics* pp. 1–44.

Csárdi, G. and Berkelaar, M. (2024), *lpSolve: Interface to 'Lp_solve' v. 5.5 to Solve Linear/Integer Programs*. R package version 5.6.23.
URL: *https://CRAN.R-project.org/package=lpSolve*

Debaere, P., Richter, B. D., Davis, K. F., Duvall, M. S., Gephart, J. A., O'Bannon, C. E., Pelnik, C., Powell, E. M. and Smith, T. W. (2014), 'Water markets as a response to scarcity', *Water Policy* **16**(4), 625–649.

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L. and Linsey, K. S. (2018), *Estimated use of water in the United States in 2015*, number 1441, US Geological Survey.

Donnelly, K. and Cooley, H. (2015), 'Water use trends', *Pacific Institute, Oakland, California* **851**.

Erfani, T., Binions, O. and Harou, J. J. (2014), 'Simulating water markets with transaction costs', *Water Resources Research* **50**(6), 4726–4745.

Garrick, D., Whitten, S. M. and Coggan, A. (2013), 'Understanding the evolution and performance of water markets and allocation policy: A transaction costs analysis framework', *Ecological Economics* **88**, 195–205.

Grafton, R. Q., Libecap, G., McGlennon, S., Landry, C. and O'Brien, B. (2011), 'An integrated assessment of water markets: a cross-country comparison', *Review of Environmental Economics and Policy* .

Gutzler, D. S. and Robbins, T. O. (2011), 'Climate variability and projected change in the Western United States: regional downscaling and drought statistics', *Climate Dynamics* **37**(5), 835–849.

He, L., Horbulyk, T. M., Ali, M. K., Le Roy, D. G. and Klein, K. (2012), 'Proportional water sharing vs. seniority-based allocation in the bow river basin of southern alberta', *Agricultural Water Management* **104**, 21–31.

Henry III, D. D., Muller, N. Z. and Mendelsohn, R. O. (2011), 'The social cost of trading: Measuring the increased damages from sulfur dioxide trading in the united states', *Journal of Policy Analysis and Management* **30**(3), 598–612.

Huffaker, R., Whittlesey, N. and Hamilton, J. R. (2000), 'The role of prior appropriation in allocating water resources into the 21st century', *International Journal of water resources development* **16**(2), 265–273.

Ji, X. and Cobourn, K. M. (2018), 'The economic benefits of irrigation districts under prior appropriation doctrine: An econometric analysis of agricultural land-allocation decisions', *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* **66**(3), 441–467.

Jones, P. A. and Cech, T. (2009), *Colorado water law for non-lawyers*, University Press of Colorado.

Leonard, B., Costello, C. and Libecap, G. D. (2019), 'Expanding water markets in the western united states: Barriers and lessons from other natural resource markets', *Review of Environmental Economics and Policy* .

Leonard, B. and Libecap, G. D. (2019), 'Collective action by contract: prior appropriation and the development of irrigation in the western united states', *The Journal of Law and Economics* **62**(1), 67–115.

Libecap, G. D. (2007), 'The assignment of property rights on the western frontier: lessons for contemporary environmental and resource policy', *The Journal of Economic History* **67**(2), 257–291.

Libecap, G. D. (2011), 'Institutional path dependence in climate adaptation: Coman's "some unsettled problems of irrigation"', *American Economic Review* **101**(1), 64–80.

Libecap, G. D. and Lueck, D. (2011), 'The demarcation of land and the role of coordinating property institutions', *Journal of Political Economy* **119**(3), 426–467.

Ostrom, E. and Gardner, R. (1993), 'Coping with asymmetries in the commons: self-governing irrigation systems can work', *Journal of economic perspectives* **7**(4), 93–112.

Pérez-Blanco, C. D., Essenfelder, A. H. and Gutiérrez-Martín, C. (2020), 'A tale of two rivers: Integrated hydro-economic modeling for the evaluation of trading opportunities and return flow externalities in inter-basin agricultural water markets', *Journal of Hydrology* **584**, 124676.

Rafey, W. (2023), 'Droughts, deluges, and (river) diversions: Valuing market-based water reallocation', *American Economic Review* **113**(2), 430–471.

Schwabe, K., Nemati, M., Landry, C. and Zimmerman, G. (2020), 'Water markets in the western united states: Trends and opportunities', *Water* **12**(1), 233.

Smith, S. M. (2021), 'The relative economic merits of alternative water right systems', *Journal of Environmental Economics and Management* **105**, 102389.

Spivak, D. S. (2021), 'The colorado river drought contingency plan', *Natural Resources Journal* **61**(2), 173–204.

Tingey-Holyoak, J. L. (2014), 'Water sharing risk in agriculture: Perceptions of farm dam management accountability in australia', *Agricultural Water Management* **145**, 123–133.

UNM (2023), 'Sharing means caring with the colorado river'. Accessed: 2025-11-10.
URL: *https://news.unm.edu/news/sharing-means-caring-with-the-colorado-river*

USU (2023), 'Combating drought: Equal water shares for colorado river users?'. Accessed: 2025-11-10.
URL: *https://artsci.usu.edu/social-sciences/political-science/international-studies/aggies-go/news/colorado-river-rights*

Wobus, C., Small, E., Carbone, J. C., Modi, P., Kamen, H., Szafranski, W. and Livneh, B. (2022), 'Water allocation, return flows, and economic value in water-scarce environments: Results from a coupled natural-human system model', *Water* **14**(20), 3280.

Womble, P. and Hanemann, W. M. (2020), 'Water markets, water courts, and transaction costs in colorado', *Water Resources Research* **56**(4), e2019WR025507.

Ypma, J., Johnson, S. G., Stamm, A., Borchers, H. W., Eddelbuettel, D., Ripley, B., Hornik, K., Chiquet, J., Adler, A., Dai, X., Ooms, J., Kalibera, T. and Jagan, M. (2025), *nloptr: R Interface to NLopt*. R package version 2.2.1.
URL: *https://CRAN.R-project.org/package=nloptr*

Appendix A Appendix

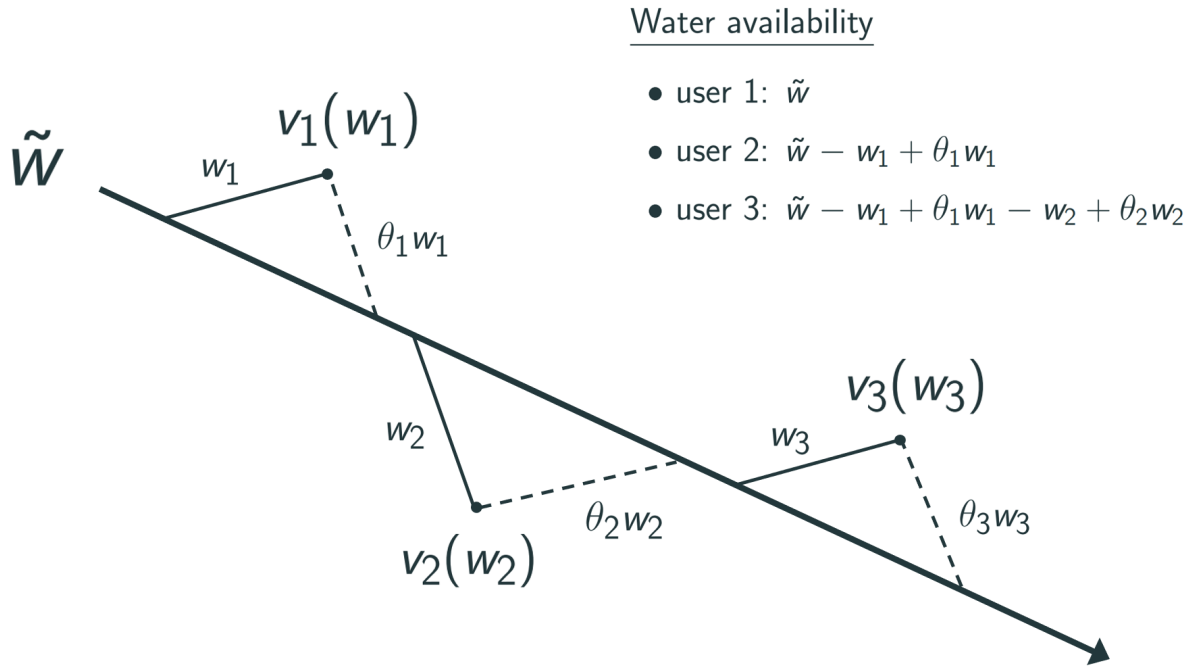


Figure A1: Visual representation of a hypothetical waterway with three users in which return flows from an upstream user are available to all downstream users

A.1 Water allocation mechanisms with two users

A.1.1 Social Planner

Before examining real-world water allocation institutions, consider the social planner (SP), whose objective is to maximize the total value of production within a river basin by allocating water w_i to two users with value functions $v_i(w_i)$, where i = upstream (u) or downstream (d). Available water in a given year is a random draw $\tilde{W} \in \boldsymbol{W}$, where $\boldsymbol{W}$ is the distribution of possible water supply conditions. Some proportion of water $\theta \in (0, 1)$ diverted to the upstream user returns to the stream and becomes available for the downstream user as a return flow. The SP's deterministic maximization problem with two water users in a given year is

$$\begin{aligned} \max_{w_u, w_d \geq 0} \quad & v_u(w_u) + v_d(w_d) \\ \text{s.t.} \quad & w_u \leq \tilde{W} \\ & w_d \leq \tilde{W} - (1-\theta) w_u. \end{aligned} \tag{A1}$$

If the two users are homogeneous, i.e., $v_u = v_d = v$, it can be easily shown that $w_u^* > w_d^*$ and $\partial v / \partial w_d^* > \partial v / \partial w_u^*$ at the optimum. Due to return flows, more water is diverted to the upstream user, resulting in a lower marginal benefit of water compared to the downstream user.[12] When $v_u \neq v_d$, optimal allocations will depend on the relative value functions.

[12]The role of return flows in determining upstream versus downstream diversions is further demonstrated in the **??**.

A.1.2 Prior Appropriation

Now consider the same basin layout, but water allocation is determined by PA. Each user owns a water right R_i^m that determines a maximum amount of water allowable for diversion, where $m =$ senior (s) or junior (j). Under PA, R_i^s must be met before the junior user can divert any water. In practice, any water use by a junior user that prevents any senior user from receiving their full allotment is prohibited. Water allocation is therefore no longer determined by a maximization problem, but by the location of the junior and senior user, their respective water rights, and the total water supply. There are two possible cases for a given level of water supply: (i.) the senior user is upstream or (ii.) the senior user is downstream.

i. When the upstream user is senior and the downstream user is junior, the resulting water allocations are:

$$w_u^s = \begin{cases} \tilde{W} & \text{if } \tilde{W} < R_u^s \\ R_u^s & \text{if } \tilde{W} \geq R_u^s \end{cases}$$
$$w_d^j = \begin{cases} \theta\tilde{W} & \text{if } \tilde{W} \leq R_u^s \\ \tilde{W} - (1-\theta)R_u^s & \text{if } R_u^s < \tilde{W} < (1-\theta)R_u^s + R_d^j \\ R_d^j & \text{if } \tilde{W} \geq (1-\theta)R_u^s + R_d^j. \end{cases} \tag{A2}$$

ii. When the upstream user is junior and the downstream user is senior, the resulting water allocations are:

$$w_d^s = \begin{cases} \tilde{W} & \text{if } \tilde{W} < R_d^s \\ R_d^s & \text{if } \tilde{W} \geq R_d^s \end{cases}$$
$$w_u^j = \begin{cases} 0 & \text{if } \tilde{W} \leq R_d^s \\ \frac{\tilde{W} - R_d^s}{1-\theta} & \text{if } R_d^s < \tilde{W} < (1-\theta)R_u^j + R_d^s \\ R_u^j & \text{if } \tilde{W} \geq (1-\theta)R_u^j + R_d^s. \end{cases} \tag{A3}$$

In each case, $(1-\theta)R_u^m + R_d^m$ characterizes the total amount of water necessary to fulfill the maximum allowable diversion for both users. For the senior user, location is irrelevant and $w_u^s = w_d^s$ in any water supply scenario. However, the junior user faces different outcomes in times of scarcity, i.e., $\tilde{W} < (1-\theta)R_u^m + R_d^m$, depending on location. If downstream, the junior user will always receive at least the return flows from the upstream user, $\theta\tilde{W}$. If upstream, the junior user faces the possibility of receiving nothing if the basin water supply is insufficient to fulfill even the senior right, i.e., $\tilde{W} < R_d^s$. Although simple, the two user case clearly conveys how risk is disproportionately assumed by the junior user. Table A1 provides a stylized example of how water allocations differ for the junior user under different stream location and supply scenarios. If the users have the same value functions in both cases, case (ii.) unambiguously results in lower basin-wide value during shortages due to the upstream, junior user receiving less water.

Table A1: Water allocation outcomes under different water supply scenarios; $R_i^s = 250$, $R_i^j = 200$, $\theta = 0.25$

	(i.)			(ii.)		
$\tilde{W}$	w_u^s	w_d^j	Total	w_d^s	w_u^j	Total
100	100	**25**	**125**	100	**0**	**100**
300	250	**112.5**	**362.5**	250	**66.7**	**316.7**
500	250	200	450	250	200	450

One criticism of the doctrine of prior appropriation is the economic inefficiency introduced through an uneven distribution of risk. During times of shortage, some users who could generate significant value per unit of water receive nothing, while others with higher priority rights may not put all of the resource to its most productive use. This uneven risk can lead to a suboptimal allocation of resources. Alternatively, share-based institutions can partly remedy this inefficiency by ensuring that all users receive a non-zero allocation of water, even during times of scarcity. Under a PS system, water is allocated proportionally according to shares until all endowments are satisfied in a descending order from highest to lowest number of shares. If a basin were to transition from PA to PS, shares may be determined by users' previous priorities or endowments or both. Share-based institutions aim to redistribute risk more equitably among water users, hopefully promoting a more efficient use of resources.

A.1.3 Proportional Sharing

Before addressing PS, here we characterize how water rights may have formed along a stream. This intermediate step has the benefit of ensuring return flows are accounted for when developing a stylized PS system. The senior user first appropriates some portion $\alpha_i^m \in (0,1)$ of average water supply from a completely unappropriated basin, and then the junior user appropriates some portion of the remaining available water. The characterization again depends on the location of each user, where $\overline{w}$ is the mean of $\boldsymbol{W}$.

i. The upstream user is senior, the downstream user is junior, and the resulting water rights are:

$$\begin{aligned} R_u^s &= \alpha_u^s \overline{w} \\ R_d^j &= \alpha_d^j \overline{w}[1-(1-\theta)\alpha_u^s]. \end{aligned} \tag{A4}$$

ii. The upstream user is junior, the downstream user is senior, and the resulting water rights are:

$$\begin{aligned} R_d^s &= \alpha_d^s \overline{w} \\ R_u^j &= \alpha_u^j \overline{w} \left(\frac{1-\alpha_d^s}{1-\theta} \right). \end{aligned} \tag{A5}$$

Again consider the same basin layout as in the SP and PA examples, but water allocation is now determined by PS. Assume this institution has replaced PA and the existing water users trade in their water rights for a share

$$\delta_i^m = \frac{R_i^m}{\sum_i R_i^m} \in (0,1). \tag{A6}$$

A more realistic formulation of shares that is driven by seniority is introduced in the n-user framework, however this simple version is sufficient for examining how allocations change under a share-

based mechanism. Under PS, if $\tilde{W}$ is not sufficient to fulfill the sum of all previous water rights, the shortage is distributed equally according to shares. Again, there are two cases depending on the location of the users.

i. The upstream user was senior, the downstream user was junior, and the resulting shares are:

$$\begin{aligned} \delta_u^s &= \frac{\alpha_u^s}{\alpha_u^s + \alpha_d^j[1-(1-\theta)\alpha_u^s]} \\ \delta_d^j &= \frac{\alpha_d^j}{\alpha_u^s + \alpha_d^j[1-(1-\theta)\alpha_u^s].} \end{aligned} \tag{A7}$$

ii. The upstream user was junior, the downstream user was senior, and the resulting shares are:

$$\begin{aligned} \delta_d^s &= \frac{\alpha_d^s}{\alpha_d^s + \alpha_u^j\left(\frac{1-\alpha_d^s}{1-\theta}\right)} \\ \delta_u^j &= \frac{\alpha_u^j}{\alpha_d^s + \alpha_u^j\left(\frac{1-\alpha_d^s}{1-\theta}\right)}. \end{aligned} \tag{A8}$$

In each case, the junior user is now guaranteed a nonzero allocation of water in a given period.

A.1.4 Two Users

To compare the efficiency of each institution, we allow users to be heterogeneous in terms of their value functions and use a Monte Carlo simulation to randomly sample 5,000 draws of $\tilde{W}$ from a log-normal distribution. Four cases are compared across the three institutions (SP, PA, PS): (i.) the senior user is upstream and high value, (ii.) the senior user is downstream and high value, (iii.) the senior user is upstream and low value, and (iv.) the senior user is downstream and low value. In each scenario, the junior user assumes the opposite characteristics. The outcomes of PA and PS would be identical when there is enough water to fulfill the endowments of both users, i.e., $\tilde{W} \geq (1-\theta)R_u^m + R_d^m$, so the following results examine only the iterations where there was a water *shortage*. The models are parameterized such that: $v_{high}(w) = 20w^{\frac{1}{2}}$, $v_{low}(w) = 10w^{\frac{1}{2}}$, $\theta = 0.25$, $\alpha_i^s = 0.6$, $\alpha_i^j = 0.75$, and $\overline{w} = 500$.

The results of the simulations are presented in Figure A1, which displays distributions of total value generated for each institution across all random water supply draws. Distributions with significant area on the right side of each panel indicate a larger proportion of simulations where high value is generated. In each scenario, the SP's distribution characterizes the upper bound on efficiency. In the scenario where the senior user is upstream and high value, PA generates more basin-wide value from water than PS on average, as evidenced by the higher peak on the right hand side of the panel. Under PA, water is first allocated to the higher value user upstream, and return flows are then available for the less valuable use downstream. Moving from PA to PS reallocates some water away from a high value use. In the scenarios where the senior user is low value, PS unambiguously generates more value than PA. When the senior user is downstream and low value (bottom right panel), there is potential for PA to generate very little relative value during extreme shortages, as shown in the peak in the left tail. An ambiguous finding results from when the senior user is downstream and high value. For minor shortages, PA may be preferred to PS, but during extreme shortages where the junior user receives no water, the senior uses all available water to a point of low marginal benefit. In general, Figure A1 conveys how PA is more sensitive to basin composition than PS in terms of risk. Moving from a desirable to less desirable composition causes

the distribution to flatten, increasing the risk of low-value outcomes in the left tail.

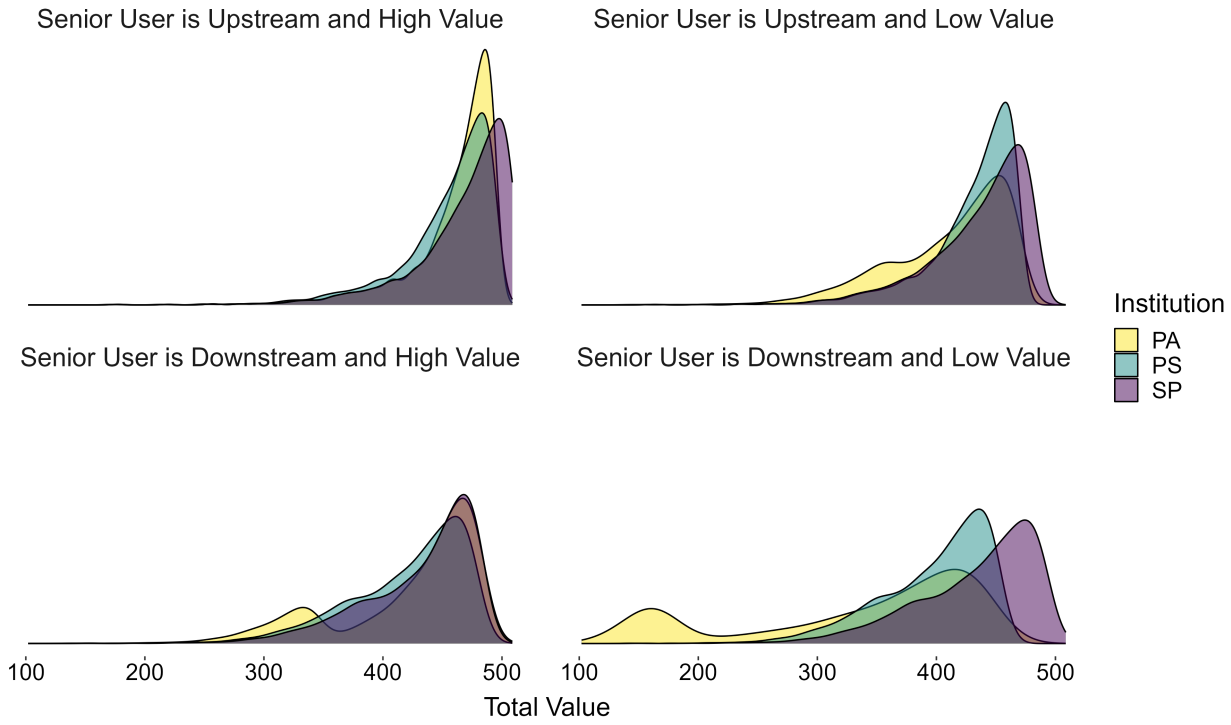


Figure A2: Monte Carlo generated total value distributions under different water allocation mechanisms with two water users

A.2 The role of return flows in water allocation

Consider n users along a river, with a uniform return rate θ for water used by each user. Each user i has a concave, differentiable value function $v(w_i)$, representing the total value generated from water diverted w_i. The objective of the Social Planner is to maximize the total value from water allocations subject to the total available water $\tilde{W}$. The simplified problem is formulated as:

$$\max_{w_1,\ldots,w_n \geq 0} \quad \sum_{i=1}^{n} v(w_i) \quad \text{s.t.} \quad w_1 + (1-\theta)w_2 + \ldots + (1-\theta)^{n-1} w_n \leq \tilde{W} \tag{A9}$$

The Lagrangian for this optimization problem is:

$$\mathcal{L} = \sum_{i=1}^{n} v(w_i) - \lambda \left(\sum_{i=1}^{n} (1-\theta)^{i-1} w_i - \tilde{W} \right), \tag{A10}$$

where λ is the Lagrange multiplier for the water availability constraint. Differentiating $\mathcal{L}$ with respect to w_i yields the first-order conditions:

$$\frac{dv}{dw_i} = \lambda (1-\theta)^{i-1}. \tag{A11}$$

Since $(1-\theta) \in (0,1)$ and i indexes stream position from upstream to downstream, this condition shows that the marginal value of water allocation decreases as we move downstream, adjusting for the return flow effect. Thus, the allocation w_i must decrease for increasing i to balance marginal value across users, leading to $w_1 > w_2 > \ldots > w_n$. This pattern indicates that, due to return flows, more water is allocated upstream *ceteris paribus*.

A.3 Consumptive use

Let c denote the allocated consumptive use to a representative user. Recall that, in practice, return flows are generated from both conveyance and application losses, so a user must divert some amount of water $x^c \geq c$ to account for those losses. Here, let $x^c = c/\eta$, where $\eta \in (0,1]$ is their conveyance/application efficiency, meaning $(1-\eta)x^c$ reenters the waterway as return flows. Furthermore, if the user wishes to trade some amount of their consumptive use, t^c, it thus requires a change in diversion t^c/η.

Consider a market where consumptive use allocations are fully transferable. We continue to focus on the three user cases under extreme scarcity, and each user has the same $\eta \in (0,1]$. Let inflow equal $\tilde{W} = c/\eta$, so there is only enough water in the stream to satisfy the most senior right, who needs to divert all available water in order to consumptively use quantity c.

Case 1: Downstream seniority (trade from $i = 3 \to i = 1$). Consider a transfer of $t^c > 0$ units of consumptive use from the most downstream user ($i = 3$) to the most upstream user ($i = 1$) when initial allocations are $\hat{c}_1 = \hat{c}_2 = 0$ and $\hat{c}_3 = c$. Because consumptive use must be supported by diversions, this implies initial diversions of $\hat{w}_1 = \hat{w}_2 = 0$ and $\hat{w}_3 = c/\eta$. A socially efficient trade occurs if the total value generated after the trade exceeds the value without the trade, or:

$$v_1(t^c) + v_2\left((1-\eta)\frac{t^c}{\eta}\right) + v_3\left(c - t^c + (1-\eta)^2\frac{t^c}{\eta}\right) > v_3(c). \tag{A12}$$

The left-hand side includes: (i) the value to user 1 from consumptively using t^c units, (ii) the value to user 2 from the return flows generated by user 1's diversion, and (iii) the value to user 3 from their reduced consumptive allocation plus return flows cascading back downstream.

If neither trader accounts for return flows, the trade occurs only if

$$v_1(t^c) + v_3(c - t^c) > v_3(c), \tag{A13}$$

which compares the buyer's gain in consumptive use with the seller's direct loss. If user 3 internalizes their own return flows from user 1's diversion, but not the spillover to user 2, the trade occurs if

$$v_1(t^c) + v_3\left(c - t^c + (1-\eta)\frac{t^c}{\eta}\right) > v_3(c). \tag{A14}$$

As before, the set of trades satisfying the socially efficient condition (Equation A12) is always weakly larger than those satisfying the private conditions, because return-flow benefits to user 2 are unpriced.

Case 2: Upstream seniority (trade from $i = 1 \to i = 3$). Suppose the most upstream user ($i = 1$) is most senior and initially consumes $\hat{c}_1 = c$, while downstream users receive only return–flow–supported consumptive use: $\hat{c}_2 = (1-\eta)c$ and $\hat{c}_3 = (1-\eta)^2 c$, with associated diversions $\hat{w}_1 = \frac{c}{\eta}$, $\hat{w}_2 = \frac{(1-\eta)c}{\eta}$, and $\hat{w}_3 = \frac{(1-\eta)^2 c}{\eta}$. Consider a trade of $t^c > 0$ units of consumptive use from

user 1 to user 3. A trade is socially efficient if total value after the trade exceeds the value without the trade:

$$v_1(c-t^c)+v_2\big((1-\eta)(c-t^c)\big)+v_3\big(t^c+(1-\eta)^2(c-t^c)\big) \;>\; v_1(c)+v_2\big((1-\eta)c\big)+v_3\big((1-\eta)^2c\big). \quad \text{(A15)}$$

The left-hand side comprises: (i) the reduced value to user 1 from keeping $c-t^c$ units of consumptive use, (ii) the value to user 2 from the smaller return flows generated by user 1's lower diversion, and (iii) the value to user 3 from the purchased consumptive use t^c *plus* the remaining return-flow–supported consumptive use $(1-\eta)^2(c-t^c)$.

If the trade is evaluated privately between users 1 and 3, user 2's loss is ignored. The private profitability condition is:

$$v_1(c-t^c)+v_3\big(t^c+(1-\eta)^2(c-t^c)\big) \;>\; v_1(c)+v_3\big((1-\eta)^2c\big). \quad \text{(A16)}$$

Let $\gamma := v_2\big((1-\eta)(c-t^c)\big)-v_2\big((1-\eta)c\big)<0$ for $t^c>0$. Then (A15) can be written as

$$v_1(c-t^c)+v_3\big(t^c+(1-\eta)^2(c-t^c)\big)+\gamma \;>\; v_1(c)+v_3\big((1-\eta)^2c\big), \quad \text{(A17)}$$

which makes clear that the set of privately profitable trades (A16) is (weakly) larger than the set of socially efficient trades (A15), because the negative externality to the non-trading intermediate user (captured by $\gamma<0$) is not internalized by the buyer and seller.

Observation 8. Both buyer and seller must divert more (less) than the traded consumptive volume to account for conveyance losses or other inefficiencies (e.g., irrigation technology), and these diversions still generate return flows that affect non-participants. While defining rights in terms of consumptive use does not solve the externality problem stemming from return flows, it can remedy the problem in some instances.

A.4 Regression exercises

In the above analysis, institutional performance is compared across hundreds of thousands of simulations under randomly generated basin compositions and different supply scenarios. While the mechanisms that consistently generate a large amount of value relative to others show robustness to varying basin characteristics, it is perhaps still not entirely clear how the institutional and physical constraints may mitigate or exacerbate inefficiencies. As a final exercise to try and characterize why certain mechanisms perform better or worse under different circumstances, we employ a series of regressions to explain why a mechanism generates the most value for a given simulation. Specifically, we run separate linear regression models for each water allocation mechanism:

$$\frac{y_k^j-y_k^{SP}}{y_k^{SP}}=\alpha^j+\beta_1^jX_k^j+\beta_2^jC_k+\varepsilon_k^j, \quad j=1,\ldots,J, \quad \text{(A18)}$$

where $\frac{y_k^j-y_k^{SP}}{y_k^{SP}}$ is the proportional decrease in value generated for mechanism j relative to the social planner in simulation iteration k. Matrix X_k^j contains variables characterizing institution-specific constraints, such as the correlation between priority and stream position and the correlation between priority and value. Matrix C_k contains information characterizing the physical characteristics of the basin such as the number of users, the level of shortage, and the correlation between stream position and return flows. The goal of this exercise is to identify which mechanisms are most sensitive to

institutional constraints and basin compositions. Regression output is available in Table A2.

Table A2: Cross-sectional regression results assessing the impacts of basin composition and water shortage on institutional efficiency relative to the Social Planner

Model:	PA (1)	PAM (2)	PS (3)	PSM (4)	ES (5)	ESM (6)
Constant	-0.034***	-0.078***	-0.141***	-0.095***	-0.086***	-0.037***
	(0.0008)	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
# Users	-0.017***	-0.005***	-0.009***	-0.005***	-0.010***	-0.008***
	(7.6×10^{-5})	(5.02×10^{-5})	(4.15×10^{-5})	(3.96×10^{-5})	(3.46×10^{-5})	(3.27×10^{-5})
# Users Sq.	0.0003***	$5.35 \times 10^{-5***}$	0.0001***	$7.94 \times 10^{-5***}$	0.0002***	0.0001***
	(1.94×10^{-6})	(1.28×10^{-6})	(1.06×10^{-6})	(1.01×10^{-6})	(8.82×10^{-7})	(8.32×10^{-7})
Shortage	-0.001***	0.0003***	-0.0004***	0.0001***	0.0002***	0.0002***
	(6.86×10^{-6})	(4.53×10^{-6})	(3.75×10^{-6})	(3.58×10^{-6})	(3.13×10^{-6})	(2.95×10^{-6})
Shortage Sq.	$-9.13 \times 10^{-7***}$	$-8.67 \times 10^{-7***}$	$1.22 \times 10^{-6***}$	$2.34 \times 10^{-7***}$	$-4.02 \times 10^{-7***}$	$-3.71 \times 10^{-7***}$
	(2.53×10^{-8})	(1.67×10^{-8})	(1.38×10^{-8})	(1.32×10^{-8})	(1.15×10^{-8})	(1.09×10^{-8})
Prio. SP Corr.	0.080***	0.035***	0.025***	0.023***	-0.0005**	-0.0006***
	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Prio. Value Corr.	0.148***	0.004***	0.106***	0.002***	0.0004	0.0002
	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
RF SP Corr.	-0.044***	-0.033***	-0.023***	-0.025***	-0.032***	-0.033***
	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Fit statistics						
Observations	467,773	467,773	467,773	467,773	467,773	467,773
R^2	0.59827	0.18693	0.38230	0.21444	0.40189	0.37203

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.5 Other considerations

In addition to system-wide value, there are other outcomes from allocating water that warrant attention. For a 36 user system, Figure A3 displays distributions for total value, average value per unit of diverted water, total diverted water, and basin outflow for each institution.

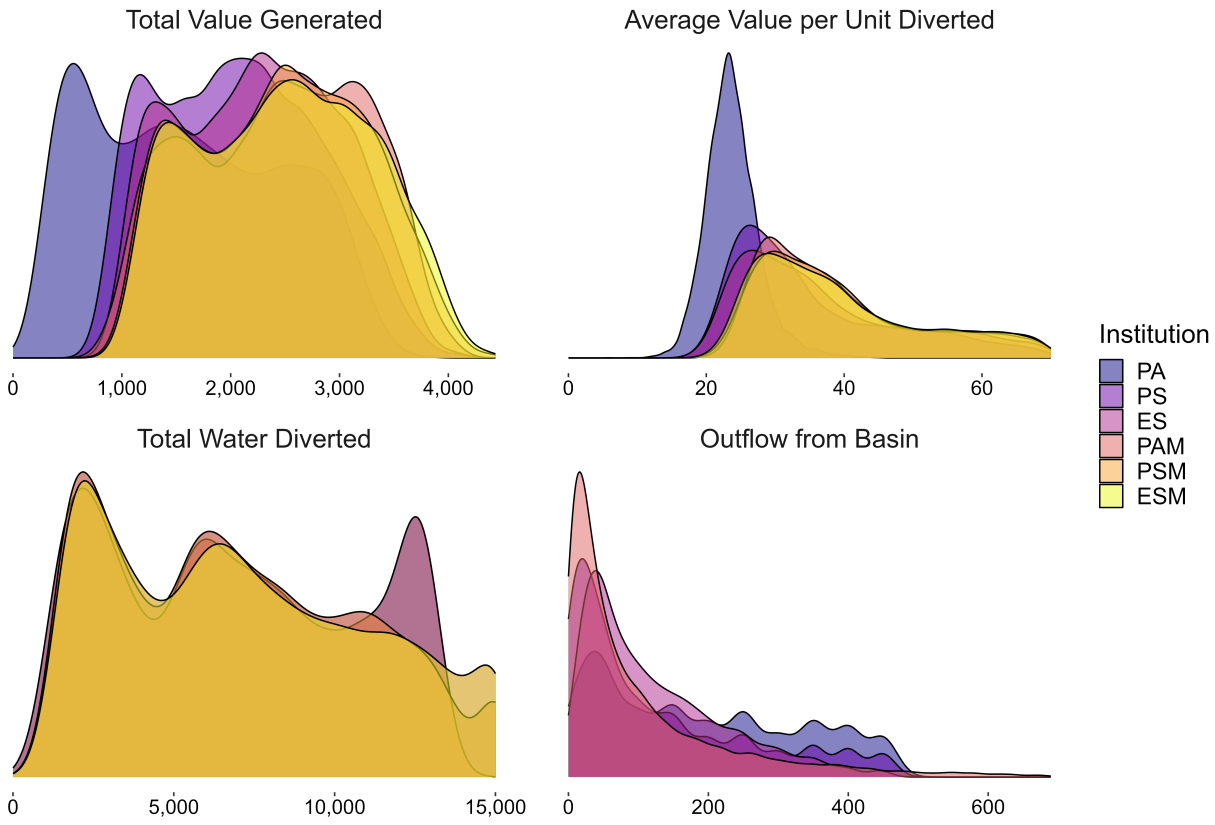


Figure A3: Monte Carlo generated total value, average value, total water, and basin outflow distributions with 36 water users