Pricing Urban Water: Rate Structures and Scarcity Rents Under Supply Uncertainty *

Connor Neff
University of Tennessee Knoxville
connor.neff@utk.edu

Charles Sims
University of Tennessee Knoxville
cbsims@utk.edu

November 2026

Abstract

Optimal pricing of scarce natural resources maximizes social welfare, yet public water utilities rarely price to reflect scarcity. We estimate residential and non-residential urban water demand using a new panel of Arizona utilities. We apply these elasticities to a dynamic Ramsey pricing model with uncertain supply and document how far prices fall below the scarcity-inclusive optimum. We find that residential rates fall well below the optimum at low-to-moderate consumption while non-residential rates tend to be above. Moving the utilities that buy Colorado River water to efficient prices cuts their draw on the river by 14 to 17 percent and raises welfare by roughly $21 million annually under normal conditions and $38 million under sustained shortage. The transition is financed almost entirely through bill increases for moderate-use residential customers while low-use and commercial bills change little. We argue that this asymmetry offers a structural explanation for why efficient water pricing remains politically difficult.

JEL codes: D12; L95; Q25
Key Words: water pricing, Ramsey pricing, natural monopolies, rate structures, scarcity rent

*We thank Casey Wichman, Ben Leard, and Scott Holladay for excellent comments and feedback.

1 Introduction

Standard results from dynamic resource economics imply that the efficient price charged by a water utility facing an uncertain supply source should embed a time-varying scarcity rent. Chronic underpricing of a resource encourages excess consumption, starves utilities of revenue for infrastructure investment, and generates welfare losses that compound as hydrological conditions deteriorate (McRae, 2015). Yet many U.S. municipal water utilities charge prices that fall below average operating costs, let alone the long-run marginal cost inclusive of the scarcity rent (Hanemann, 1998; Timmins, 2002b; Olmstead, 2010; Wichman, 2024). Water utilities are often constrained by political pressures to balance equity and affordability concerns with budget management and conservation signals. Utility managers together with regulators often employ nonlinear rate structures to balance these political, environmental, and budgetary pressures. The financial and distributional consequences of these pricing decisions are well-documented,[1] yet the natural resource management implications are much less understood. In water-abundant regions, where many studies of pricing focus, underpricing the scarcity rent element is much less of a concern than in water-scarce regions like the western United States.

Water supplies around the world are increasingly strained by intensifying droughts driven by climate change. Many urban water users in the southwestern United States rely on purchased water from a shared supply whose availability fluctuates with drought conditions. This region is more exposed to shortages than ever as it experiences historically exceptional drought severity (Williams et al., 2022). The U.S. Bureau of Reclamation's 2027-2028 Colorado River operating plan requires California, Arizona, and Nevada to reduce Colorado River water use by 1.25 million acre-feet (AF) in each of the following two years with 760,000 AF of reductions from Arizona alone (U.S. Bureau of Reclamation, 2026). Water utilities and regulators must adequately design and set rates to incorporate the scarcity and uncertainty of supplies to maximize social benefit.

In this paper, we ask whether the prices water utilities charge embed the scarcity value of water, and where along the consumption schedule they diverge from the dynamic optimum. We first develop a dynamic Ramsey pricing model for a two-sector (residential and non-residential) municipal water utility facing stochastic supply shocks. We extend Sağlam (2019) by centering

[1] As noted in Wichman (2024), these rate structures discount volumetric prices for low levels of consumption (Olmstead et al., 2007; Szabó, 2015; Wang and Wolak, 2022; Allaire and Dinar, 2022), may generate allocative inefficiencies (Borenstein, 2012, 2016; Borenstein and Davis, 2012), require subsidization from alternative municipal revenue sources (Timmins, 2002a; Renzetti, 1999), and ineffectively redistribute income (Levinson and Silva, 2022; Wietelman et al., 2025).

the manager's problem on the decision to purchase and bank external water supplies as insurance against future shortages, rather than relying solely on price adjustments to ration demand. The model yields a clean decomposition of the optimal price for each sector: marginal cost plus a Ramsey markup plus the scarcity rent. The scarcity rent evolves over time and is endogenous to current storage, carryover revenue, and the distribution of the supply shock. We apply the model to Arizona and estimate parameters using a panel of 194 water utilities over 2014, 2015, 2017, and 2019, assembled from utility-level rate schedules, supply and withdrawal records, financial filings, and purchased delivery data to provide evidence that observed prices deviate from the optimum. Finally, we simulate the optimal pricing policy for different utilities and characterize the divergence across the consumption distribution and by customer type.

The divergence between observed and efficient water prices is well-documented, but the existing literature benchmarks efficiency against average cost recovery rather than the socially optimal price that incorporates a forward-looking scarcity rent. McRae (2015) shows that politically constrained utilities are trapped in an equilibrium where rate increases are blocked by consumer-voters who do not internalize the resource cost. Wichman (2024) documents widespread cost-recovery failures across municipal utilities in the southeastern U.S. where water scarcity is driven more by population growth than resource limitation.

To our knowledge, no study asks whether increasing block rate (IBR) versus uniform rate structures differ in their capacity to embed the scarcity rent, or derives the efficient benchmark from a structural dynamic model. Timmins (2002b) estimates that municipal managers in the arid West price groundwater far below its dynamic marginal cost and recovers the in-situ scarcity value from a dynamic program, but characterizes each utility with a single marginal price and a single aggregate demand, and therefore cannot speak to how a utility's rates distribute scarcity along the consumption schedule nor to how the shortfall is spread across sectors. Baerenklau et al. (2014) find that one southern California utility's switch from uniform rates to increasing block rates cut residential demand by roughly 17 percent, and that consumption responds to the marginal rather than the average price[2], the channel through which a block schedule can transmit a scarcity signal. Sibly and Tooth (2014) instead show that under fixed supply and cost recovery a two-tier tariff must price its lower tier below, and its upper tier above, the efficient volumetric rate, distorting consumption for every household except the one at the threshold and creating deadweight loss relative to an optimally set two-part tariff. The tension is more apparent than real, because Sibly

[2] Ito (2014) finds evidence in electricity that consumers respond to average rather than marginal price, and Wichman (2014) documents similar inattention in residential water demand.

and Tooth (2014) compare a block tariff to a uniform rate that is itself set optimally in a static, full-information setting, whereas observed uniform rates are neither optimal nor adjustable as supply tightens between infrequent rate reviews. Analyzing drought surcharges added to individualized "budget-based" rates (BBRs) of two southern California utilities, Wietelman et al. (2025) find that budgets undercut the scarcity signal and simpler rate structures can dominate along equity dimensions. We show that uniform rates tend to underprice along the entire residential consumption schedule whereas IBR rates underprice at low consumption and overprice at high use. In contrast, we provide evidence that non-residential rates are priced at or above the scarcity optimum across the consumption schedule. Moving to optimal pricing would increase residential bills by 12–16% at moderate consumption while remaining nearly unchanged at low and high consumption.

We make four contributions. First, we characterize how far observed prices fall below the scarcity-inclusive optimum across the consumption distribution and by customer type, rather than as a single average. Most utilities price below even average cost, leaving the forward-looking scarcity rent almost entirely unpriced; across Central Arizona Project (CAP) utilities this rent has a median of $0.52 per kgal and reaches nearly $6 per kgal for the utilities closest to their resource constraint. The gap to the optimum is far from uniform along the schedule. At the lowest tier both rate structures price well below the optimum, with the median increasing-block utility roughly 43 percent below and the median uniform utility roughly 24 percent below at 4 kgal per month. The two diverge as use rises. The median IBR utility closes the residential gap entirely and prices the heaviest users 18 to 24 percent above the optimum, while the median uniform-rate utility stays between 5 and 24 percent below it across the schedule. Non-residential prices, by contrast, meet or exceed the optimum at almost all consumption levels. This is consistent with utilities cross-subsidizing volumetric rates by over-pricing non-residential customers relative to residential users, though it may also reflect sector differences in the cost of service or in the elasticity-driven Ramsey markup the model itself predicts.[3] Because ratemaking requires regulatory approval and rates adjust infrequently, a flat rate cannot differentiate scarcity across consumption levels between reviews, whereas an increasing block schedule embeds a price gradient that approximates the scarcity rent at upper tiers even when the overall rate level is fixed (Hanemann, 1998; McRae, 2015).

Second, we quantify the welfare and conservation payoff of moving to optimal prices and

[3]We include a Timmins-style hybrid weighted alternative objective function to describe how utilities may weight residential consumer surplus greater than non-residential. The pricing gaps shrink slightly, yet still demonstrate the same pattern; low-use residential rates are severely underpriced while non-residential rates sit above the scarcity inclusive optimum.

show that it depends on a utility's resource buffer. Simulating supply-shock paths over five- and hundred-year horizons, we measure welfare as the flow change in consumer surplus plus utility net revenue, benchmarking observed prices against a status quo that already purchases CAP water optimally so the comparison isolates the effect of pricing. Welfare gains are large and optimal pricing cuts demand substantially. Aggregated across the 21 CAP utilities over a hundred-year horizon, optimal pricing raises welfare by about $21 million per year under normal conditions and $38 million under sustained shortage, while cutting Central Arizona Project water purchases by 90,000 to 107,000 acre-feet, roughly 14 to 17 percent of the region's municipal and industrial entitlements. With only six uniform-rate utilities, the short-run gains are statistically indistinguishable across increasing block and uniform rates, while over the long run uniform utilities show larger point estimates, consistent with increasing block schedules already embedding part of the scarcity signal at their upper tiers. This conservation gain is financed by moderate residential bill increases concentrated at middling use. Once the over-recovering fixed charge is retired for an all-volumetric optimum, residential bills at ten thousand gallons per month rise by 12 to 16 percent, while low-use residential and non-residential bills change little, and the welfare changes are largest for utilities with thin resource buffers.

Third, we estimate price elasticities of demand for both the residential and non-residential sectors from the same panel, instrumenting the endogenous marginal price with the lagged average price of neighboring utilities in the same county, a cross-market instrument in the style of Hausman et al. (1994) and Nevo (2001).[4] The instruments are strong in both first stages. We estimate a residential elasticity of -0.53 and provide, to our knowledge, the first large-scale urban non-residential estimate, -0.87.[5] This is more elastic than recent residential water demand estimates but in line with recent evidence from Burlig et al. (2021) on agricultural water demand estimates. Because non-residential customers use roughly 30% of delivered water and are more price-elastic than households, restricting estimation to households would understate the quantity response, and hence the welfare gain, from pricing at the optimum.

Fourth, we assemble a panel of 194 Arizona utilities over 2014, 2015, 2017, and 2019 through

[4] Berry and Haile (2021) and MacKay and Miller (2025) detail the conditions under which this instrument identifies the price parameter; the key condition is that the demand shock in the focal market is mean-independent of the price in neighboring markets conditional on observables.

[5] Prior work estimates industrial water demand at the firm or sector level rather than at the utility scale. Renzetti (1992) and Renzetti (1993) estimate input demand for water among Canadian manufacturing plants using plant-level data. Reynaud (2003) estimates industrial water demand for French firms. These studies do not use price variation across public water utilities and focus on manufacturing rather than the full non-residential sector (commercial, institutional, and industrial customers) served by municipal systems.

public records requests and public sources that include quantities directly by supply source and by customer class, rather than inferred from revenue and rates. We also observe revenues, expenditures, CAP water deliveries, storage account balances, and bill amounts across consumption levels.

We build on two strands of literature, one on stochastic dynamic programming for water resource pricing and another studying demand effects and equity implications of water pricing. Moncur and Pollock (1988) develop an early valuation and pricing model for urban water scarcity in which the efficient price incorporates a rent reflecting the opportunity cost of current consumption. Sağlam (2019) formalizes this into a stochastic dynamic program for a water utility facing uncertain supply, characterizing optimal prices as a combination of the inverse-elasticity Ramsey markup, a scarcity rent of the water stock, and sector-specific marginal costs. Tsur (2020) derives optimal groundwater pricing when extraction costs depend on the water table, and Chu and Grafton (2021) incorporate supply uncertainty into a risk-adjusted user cost. We implement a Sağlam-style stochastic dynamic program using utility-level data, recovering this rent separately for each utility rather than as a single regional aggregate. Greater supply uncertainty changes the scarcity rent little, even though it leaves utilities less able to purchase and bank water for future periods.[6]

A large empirical literature studies demand effects and equity implications of nonlinear water pricing but takes the rate structure as given rather than asking how closely it tracks efficient prices. Griffin (2001) analyzes the conditions under which IBR schedules approximate efficient pricing when utilities face revenue constraints, but does so in a static framework that abstracts from dynamic supply uncertainty and the endogenous evolution of the scarcity rent. Wang and Wolak (2022) characterize the tradeoff between revenue recovery and conservation in nonlinear price schedule design, showing that optimal schedules depend critically on elasticity heterogeneity across usage tiers. With greater demand elasticity among high-use non-residential customers as estimated in this paper, it is unclear whether IBR or uniform rate structures better approximate the conservation signal. Mansur and Olmstead (2012) show that replacing quantity rationing with market-clearing prices raises welfare substantially. We depart from this literature by asking how well observed rate designs approximate the optimal price which incorporates the scarcity rent. This question has not been answered when looking at stochastic supply sources and rates for multiple consumer sectors.

[6] Abbott et al. (2026) show that stochasticity does not change scarcity rents substantially unless nonconvexities in the resource dynamics are present.

2 Model

In this model, a benevolent water manager chooses prices and the quantity of water purchased to maximize the net present value of customer consumer surplus. Customers fall into one of two user groups: residential and non-residential. The manager is constrained by a stochastic resource constraint and a budget constraint requiring cost recovery.

2.1 Water Management Problem

The water manager's optimization problem can be written recursively as follows:

$$V(w,b) = \max_{p_1,p_2,x(\theta)} \ \mathbb{E}_\theta\Big[\sum_{i=1}^{2} CS_i(p_i) + \beta V(w'(\theta), b'(\theta))\Big] \tag{1}$$

$$\text{s.t.} \quad w'(\theta) = w + w_f + x(\theta) - \sum_{i=1}^{2} q_i(p_i); \quad \forall\theta, \tag{2}$$

$$b'(\theta) = (p_1 - c_1)q_1(p_1) + (p_2 - c_2)q_2(p_2) + Rb - \tau - p_x\, x(\theta); \quad \forall\theta, \tag{3}$$

$$x(\theta) \le \bar{x}(\theta), \tag{4}$$

$$x(\theta) \ge 0, \tag{5}$$

$$w \ge 0, \tag{6}$$

where all prices and quantities are non-negative. The carryover flow $w'(\theta)$ may be negative, representing a net draw on accumulated storage; the beginning-of-period stock w is constrained to be non-negative by equation (6). The term w_f denotes fixed water entitlements (groundwater and surface rights), treated as exogenous. The unit cost of purchasing external water is p_x; constraint (4) limits purchases to $\bar{x}(\theta) = \bar{x} \cdot \theta$, where $\bar{x}$ is the contractual maximum and $\theta \in [0,1]$ is the supply shock realized after prices are set. The gross rate of return on bond savings is R. The discount factor is β, and $CS_i(p_i)$ is consumer surplus for user group $i \in \{1,2\}$.

The expectation operator $\mathbb{E}_\theta(\cdot)$ is over the supply shock θ, which is drawn independently each period from a known distribution. The control variables for the manager's stochastic dynamic programming (SDP) problem are the prices for each user sector $\{p_1, p_2\}$, the quantity of purchased water $\{x(\theta) \ge 0\}$, and the carryover stock and revenue $\{w'(\theta), b'(\theta)\}$. After solving for prices and purchased water, the carryover stock and revenue can be computed from the constraints.

Constraint (4) caps current-period purchases at $\bar{x}(\theta) = \bar{x} \cdot \theta$, where the shock θ is drawn independently each period from a Beta distribution on $[0, 1]$ and $\bar{x}$ is the contractual maximum. A lower realization of θ tightens the purchase limit, reducing the water manager's ability to supplement fixed supplies.

Let $\lambda(\theta)$ and $\mu(\theta)$ be the Lagrange multipliers on the resource constraint (2) and the revenue constraint (3), respectively. Let $\eta(\theta) \geq 0$ and $\psi(\theta) \geq 0$ denote the Lagrange multiplier on the purchase limit equation (4) and the non-negativity constraint (5), with complementary slackness $\eta(\theta)\,[\bar{x}(\theta) - x(\theta)] = 0$ and $\psi(\theta)\,x(\theta) = 0$. Equation (6) is the non-negativity constraint on the water stock.

2.2 Timing

We consider two sectors: residential (sector 1) and non-residential (sector 2). Let θ denote a supply-side stochastic shock to the model. We denote sectoral demand by sector i as $q_i(p_i)$.

At the beginning of the period, the water manager observes the carryover stock (w). Before the supply shock is realized, the manager sets prices (p_1, p_2). Sectoral prices are determined ex ante and do not depend on the current-period supply shock. Once the shock is observed the manager can decide to augment the existing stock by purchasing water (x) at unit cost (p_x). At the end of the period the remaining stock $(w'(\theta))$ is saved for the next period.

2.3 Resource Constraint

The dynamic resource constraint for the intertemporal resource allocation is:

$$w'(\theta) + q_1(p_1) + q_2(p_2) \leq w + w_f + x(\theta); \quad \forall\theta, \tag{7}$$

where the left-hand side (LHS) is the sum of savings and total withdrawals. Specifically, $w'(\theta)$ is the remaining stock at the end of the period that is saved for next period and $q_1(p_1)$ and $q_2(p_2)$ are sectoral demands. In the optimization problem this resource constraint generates Lagrange multiplier λ, which is the shadow value of the resource, and when divided by the revenue constraint Lagrange multiplier μ generates the scarcity rent. This scarcity rent is the main measure of natural resource value evaluated throughout the results section of this paper.

The right-hand side (RHS) is the total supply. This is the sum of savings from last period (w),

fixed water stock (w_f)[7], and the additional purchased water $(x(\theta))$. We assume that the water manager can only purchase this water and not sell it, therefore $x(\theta) \geq 0$.

A shortage occurs in a period when aggregate demand exceeds the amount of fixed water supply available plus available water for purchase. For a given supply shock θ:

$$q_1(p_1) + q_2(p_2) > w_f + x(\theta). \tag{8}$$

A shortage may occur due to high consumption or low supply resulting from an adverse supply shock. The water manager can respond to the shortage by drawing down its savings or by charging a higher price to ration demand.[8]

2.4 Budget Constraint

The water manager is bound by a dynamic revenue constraint which requires cost recovery. Let $[c_1, c_2]$ denote constant marginal costs for the two sectors and τ represent the fixed cost of production for all users. Following Sağlam (2019), we assume marginal costs may differ between sectors due to differences in production costs or accessibility[9]. The water manager's budget constraint is:

$$b'(\theta) + p_x\, x(\theta) \leq (p_1 - c_1)q_1(p_1) + (p_2 - c_2)q_2(p_2) + Rb - \tau, \tag{9}$$

where b and $b'(\theta)$ denote the carryover revenue (in the form of bonds) from last period and this period, respectively, and R is the fixed gross rate of return on bonds. The bond savings $b'(\theta)$ are non-negative, meaning the water manager cannot finance losses by issuing bonds.

Equation (9) is consistent with cost recovery, as revenue must cover costs throughout the periods.

[7] This is water that the water manager has a fixed right to. This includes groundwater rights and surface water rights, which we assume are fixed.

[8] A third option beyond the scope of the model presented here would be a command and control reduction in water use. Mansur and Olmstead (2012) evaluates the inefficiency of such non-price management options compared to price-based rationing.

[9] Production costs may include differing water quality required by user groups. Accessibility costs such as transfer costs may differ by user groups. Empirically, we use marginal costs that are the same across user-groups within a utility.

2.5 Optimal Policy Functions

The optimal policy functions follow from the first-order and envelope conditions (FOC and EC, respectively) of the dynamic program, equations (1)–(6). Because sectoral prices (p_1, p_2) are set ex ante before the supply shock θ is realized while the CAP purchase $x(\theta)$ is chosen ex post after observing θ, the analysis has a natural two-stage structure. We first characterize the intertemporal Euler conditions governing the carryover stock $w'(\theta)$ and revenue $b'(\theta)$, then derive the optimal pricing rule, which is the central result. Throughout, a prime denotes a next-period value and θ' the next-period supply shock.

2.5.1 Optimal Carryover Stock and Revenue

Taking the first-order and envelope conditions of the Bellman equation with respect to $w'(\theta)$ and w:

$$\text{FOC:} \quad \frac{\beta\, \partial V(w'(\theta),\, b'(\theta))}{\partial w'} = \lambda(\theta), \tag{10}$$

$$\text{EC:} \quad \frac{\partial V(w,\, b)}{\partial w} = \mathbb{E}_\theta[\lambda(\theta)]\,. \tag{11}$$

Iterating equation (11) forward one period and substituting equation (10) yields the Euler equation for the carryover water stock:

$$\lambda(\theta) = \beta\, \mathbb{E}_{\theta'}\left[\lambda'(\theta')\right]. \tag{12}$$

Equation (12) equates the current shadow value of one unit of water stock with the discounted expected shadow value tomorrow. When storage is low and $\lambda(\theta)$ is high, the value function is steeply increasing in w; the utility raises prices to reduce current withdrawals and preserve stock for future periods.

The analogous conditions for carryover revenue $b'(\theta)$ and b yield the Euler equation for revenue:

$$\text{FOC:} \quad \frac{\beta\, \partial V(w'(\theta),\, b'(\theta))}{\partial b'} = \mu(\theta), \tag{13}$$

$$\text{EC:} \quad \frac{\partial V(w,\, b)}{\partial b} = R\, \mathbb{E}_\theta[\mu(\theta)]\,, \tag{14}$$

which gives:

$$\mu(\theta) = \beta R \, \mathbb{E}_{\theta'} \left[\mu'(\theta') \right]. \tag{15}$$

The shadow value of a dollar of carryover revenue today equals the discounted expected shadow value next period multiplied by the gross bond return R. The utility saves budget surplus when high prices generate revenue that can finance CAP purchases or underwrite below-cost pricing in future periods of abundance, following Sağlam (2019).

Supply uncertainty induces precautionary management through the curvature of the resource problem. Because the value of stored water is concave in available supply, its marginal value, $\lambda(\theta)$, rises steeply and convexly as supply tightens toward the resource constraint. A mean-preserving increase in delivery risk raises the expected shadow value of water above its value at the mean shock, $\mathbb{E}_\theta[\lambda(\theta)] > \lambda(\mathbb{E}_\theta[\theta])$, by Jensen's inequality. This higher expected shadow value feeds directly into the optimal price through the scarcity-rent term $\bar{\lambda}/\bar{\mu}$ and into more aggressive banking of carryover stock. Stochasticity therefore matters for a reason distinct from risk aversion. A risk-neutral manager facing a concave resource value prices higher and holds larger buffers under uncertainty than under the mean shock, and the size of this precautionary wedge scales with the amount of supply risk.

2.5.2 Optimal Sectoral Prices

Sectoral prices are the primary policy instrument and the focus of the empirical analysis. Taking the derivative of the Bellman objective equation (1) with respect to p_i, and using the fact that $q_i(p_i)$ does not depend on θ since prices are set ex ante, the first-order condition for sector i is:

$$0 = \frac{\partial CS_i}{\partial p_i} + \mathbb{E}_\theta \left[-\lambda(\theta) \frac{\partial q_i}{\partial p_i} + \mu(\theta) \left((p_i - c_i) \frac{\partial q_i}{\partial p_i} + q_i(p_i) \right) \right]. \tag{16}$$

Applying Roy's identity, $\partial CS_i/\partial p_i = -q_i(p_i)$, and defining $\bar{\lambda} \equiv \mathbb{E}_\theta[\lambda(\theta)]$ and $\bar{\mu} \equiv \mathbb{E}_\theta[\mu(\theta)]$ as the expected shadow prices on the resource and revenue constraints, respectively, equation (16) simplifies to:

$$(\bar{\mu} - 1)\, q_i(p_i) + \left(\bar{\mu}(p_i - c_i) - \bar{\lambda} \right) \frac{\partial q_i}{\partial p_i} = 0. \tag{17}$$

Because demand $q_i(p_i)$ and its derivative $\partial q_i/\partial p_i$ are non-stochastic, the expectations in equation (16) reduce to the scalar multipliers $\bar{\lambda}$ and $\bar{\mu}$. Solving equation (17) for p_i yields the following result.

The optimal volumetric price for sector $i \in \{1, 2\}$ satisfies:

$$p_i = \underbrace{\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right)\left(\frac{-q_i(p_i)}{\partial q_i / \partial p_i}\right)}_{\text{Ramsey markup}} + \underbrace{\frac{\bar{\lambda}}{\bar{\mu}}}_{\text{scarcity rent}} + \underbrace{c_i}_{\text{marginal cost}}, \tag{18}$$

where $\bar{\lambda} = \mathbb{E}_\theta[\lambda(\theta)]$ and $\bar{\mu} = \mathbb{E}_\theta[\mu(\theta)]$ are the expected Lagrange multipliers on the resource and revenue constraints.

The optimal price decomposes into three additive components. The first term is the Ramsey markup: the inverse-elasticity rule scaled by $(\bar{\mu} - 1)/\bar{\mu}$, which measures the tightness of the revenue constraint. A sector with less elastic demand bears a proportionally larger markup, recovering fixed costs τ while minimizing aggregate welfare loss. The second term, $\bar{\lambda}/\bar{\mu}$, is the scarcity rent: the expected shadow value of the water stock normalized by the shadow value of revenue. This is the forward-looking opportunity cost of drawing down one unit of storage–the same object identified as the user cost of a scarce resource by Moncur and Pollock (1988) and as the natural capital value of a stock by Fenichel and Abbott (2014). The multiplier $\bar{\lambda}$ is the marginal continuation value of a banked unit of water, as equation (10) makes explicit, measured in units of welfare; dividing by $\bar{\mu}$, the marginal welfare value of a dollar of revenue, re-expresses it in dollars so that it enters the price alongside the marginal cost c_i. The rent is thus the dollar value of a unit of water held in storage. Because the carryover stock consists of banked CAP water and purchases $x(\theta)$ are the only margin on which the utility can add to that stock, $\bar{\lambda}/\bar{\mu}$ is the marginal value of a stored unit of CAP water: within any single period the three supply sources enter the resource constraint additively and are therefore fungible, so the rent prices stored water rather than one source over another. The scarcity rent is endogenous to the state (w, b) and the distribution of θ: it rises when storage is low relative to expected demand or when the supply shock tightens the CAP constraint, embedding a conservation signal into the optimal price. The third term is the sector-specific marginal cost of supply.

Two limiting cases clarify the structure of equation (18). When the revenue constraint is slack ($\bar{\mu} \to 1$), the Ramsey markup vanishes and the rule collapses to $p_i = \bar{\lambda} + c_i$: full marginal user cost pricing, the efficient benchmark from dynamic resource economics. When storage is abundant and no shortage threatens ($\bar{\lambda} \to 0$), the scarcity rent disappears and the rule reduces to the static Ramsey benchmark, recovering fixed costs through inverse-elasticity markups alone. The empirically relevant case for Arizona utilities combines both: revenue constraints are real and the CAP

shortage mechanism makes supply uncertainty salient, so both components should be present in observed prices to achieve efficiency.

Assuming marginal cost, c_i, does not differ by sector, the scarcity rent is the same for both sectors, reflecting the fact that a unit of water drawn from storage by any user group carries the same opportunity cost to the utility.

2.5.3 Demand for External Water

The dynamic program is linear in the CAP purchase $x(\theta)$ for each realization of θ. Taking the first-order condition with respect to $x(\theta)$:

$$\lambda(\theta) - p_x \, \mu(\theta) = \eta(\theta) \geq 0, \tag{19}$$

with complementary slackness $\eta(\theta)[\bar{x}(\theta) - x(\theta)] = 0$. Equation (19) states that the utility purchases CAP water when the shadow value of an additional unit, $\lambda(\theta)$, is at least as large as its effective cost $p_x \, \mu(\theta)$. When $\lambda(\theta) < p_x \, \mu(\theta)$, the external source is too costly relative to the value of additional storage and $x(\theta) = 0$; when the CAP constraint binds, $\eta(\theta) > 0$ and $x(\theta) = \bar{x}(\theta) = \bar{x} \cdot \theta$. Dividing equation (19) by $\mu(\theta)$ ties the rent to the CAP price: $\lambda(\theta)/\mu(\theta)$ pivots around p_x, falling below it in abundant states when the utility buys nothing, equaling it when the utility replenishes at an interior quantity, and rising above it by $\eta(\theta)/\mu(\theta)$ when drought binds the allotment.

3 Institutional Background

Arizona municipal water utilities draw from three primary supply sources: surface water rights governed by prior appropriation doctrine, groundwater extracted from local aquifers, and purchased Colorado River water delivered through the Central Arizona Project (CAP). Groundwater and surface water rights are held long-term and typically not subject to year-to-year variation. Unused water from either source does not roll over to the following year. Excess water from any source may be stored through the Arizona Water Banking Authority (AWBA). The relative weight of each source varies considerably across the state, with CAP-dependent utilities concentrated in the Phoenix and Tucson metropolitan corridors and groundwater-reliant utilities predominating in rural areas.

The regulatory environment divides the state into Active Management Areas (AMAs) and non-AMA regions. The five AMAs (Phoenix, Tucson, Pinal, Prescott, and Santa Cruz) were established under Arizona's 1980 Groundwater Management Act in response to severe aquifer overdraft, and utilities operating within them are subject to mandatory conservation programs and a 100-year assured water supply requirement administered by the Arizona Department of Water Resources (ADWR). Non-AMA utilities face no equivalent supply adequacy standard and are regulated primarily through the Arizona Corporation Commission (ACC) on financial and service-quality grounds, with groundwater extraction constrained only by well permitting and local basin conditions.

The Central Arizona Project (CAP) is a 336-mile aqueduct system that diverts Colorado River water from Lake Havasu eastward and southward to serve the Phoenix and Tucson metropolitan areas (see Figure A5 in the appendix), with first water deliveries in 1985 and the backbone aqueduct substantially complete by 1993 (Hanemann, 2002). The Central Arizona Water Conservation District (CAWCD) holds Arizona's Colorado River allotment and delivers water to end-users through long-term subcontracts; the 52 current municipal and industrial (M&I) subcontractors collectively hold entitlements totaling approximately 620,000 acre-feet per year. In a normal year, CAP delivers approximately 1.5 million acre-feet across all user classes, accounting for around 35 percent of Arizona's total water use (Central Arizona Project, 2024). CAP water is distributed by a strict priority hierarchy: tribal water holds the highest priority, followed by M&I subcontract water, then non-Indian agricultural (NIA) water, and residual excess water at the lowest tier. CAP's allotment is determined by Lake Mead's pool elevation. The lake's elevation has declined on net over the last 25 years, reaching historically low levels near the end of 2025 as shown in Figure 1.

The shortage mechanism translates Lake Mead reservoir elevations directly into reductions in CAP deliveries to Arizona. Under the Bureau of Reclamation's 2007 Interim Guidelines, a Tier 1 shortage is triggered when Lake Mead is projected to fall below 1,075 feet, reducing Arizona's CAP allocation by 512,000 acre-feet annually; Tier 2 and Tier 3 conditions cut deliveries by up to 720,000 acre-feet from the full entitlement (U.S. Bureau of Reclamation, 2007). Because cuts fall first on NIA and excess water pools, M&I subcontract holders retain most of their deliveries under moderate shortage conditions, though severe shortages affect all user classes. In August 2021, the Bureau of Reclamation declared the first Tier 1 shortage for calendar year 2022, and conditions worsened to Tier 2a in 2023. Figure 2 shows that Lake Mead elevations approach Tier 3 in late 2027 under both most probable and minimum probable scenarios. Climate-change-driven drought projections

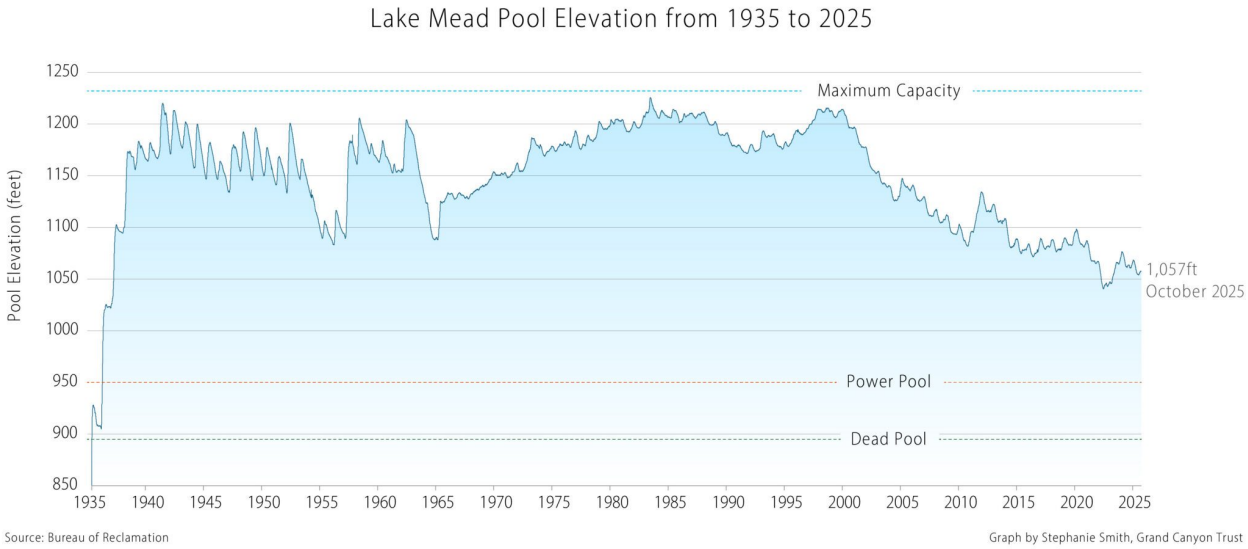


Figure 1: Lake Mead Pool Elevation, 1935–2025

Notes: Monthly pool elevation in feet above sea level, January 1935 through October 2025. The reservoir stood at 1,057 feet in October 2025, its lowest sustained level in the historical record. Dashed reference lines mark maximum storage capacity (≈1,229 ft), power pool (≈950 ft), and dead pool (≈895 ft). Source: U.S. Bureau of Reclamation; figure by Stephanie Smith, Grand Canyon Trust (reproduced with permission).

suggest that Colorado River rationing may become more frequent and intense. This uncertainty in supply is captured by the supply shock in the model.

To buffer against shortage-driven delivery reductions, Arizona utilities and the Arizona Water Banking Authority (AWBA) bank surplus CAP water underground during years of abundant supply. When deliveries exceed immediate demand, utilities recharge water into permitted underground storage facilities and ADWR tracks the resulting Long-Term Storage Account (LTSA), which can be withdrawn in future periods. Through 2024, the AWBA had accumulated approximately 4.4 million acre-feet in its LTSA, with roughly 2.3 million acre-feet designated specifically to firm M&I subcontract supplies during shortage events (Arizona Water Banking Authority, 2024). Utilities can also hold storage credits directly and draw them down when CAP deliveries fall short, converting a stochastic flow entitlement into a more reliable supply through active inventory management. This arrangement is the empirical analog to the water stock in the model, generating the intertemporal tradeoff at the center of the Ramsey pricing problem.

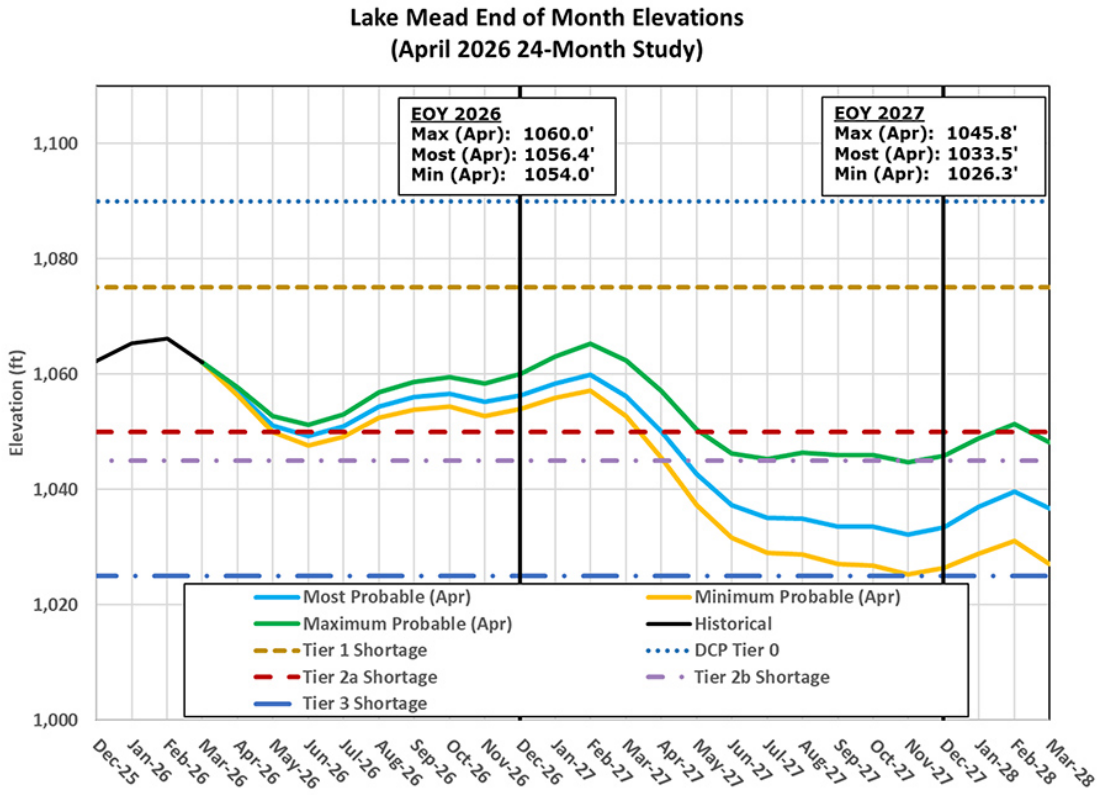


Figure 2: Lake Mead End-of-Month Elevation Projections, 2026–2028

Notes: Bureau of Reclamation April 2026 24-Month Study projections of Lake Mead end-of-month pool elevation through March 2028. Lines show most probable, maximum probable, and minimum probable outcomes. Horizontal dashed lines mark shortage tier thresholds: Tier 1 (1,075 ft), Tier 2a (1,050 ft), Tier 2b (1,045 ft), and Tier 3 (1,025 ft). Under the most probable scenario, end-of-year elevations are projected at 1,056 ft in 2026 and 1,034 ft in 2027. At these elevations, 2026 conditions correspond to Tier 1 (below 1,075 ft but above the Tier 2a threshold of 1,050 ft) and 2027 conditions to Tier 2b (below 1,045 ft). Source: U.S. Bureau of Reclamation, Colorado River Basin 24-Month Study (April 2026), `https://www.usbr.gov/lc/region/g4000/riverops/24mo.html`.

4 Data

We combine data on rates, supply quantities, withdrawal quantities, revenues, and expenditures from disparate public data sources and public records requests to assemble a four-year panel for 194 utilities in Arizona covering 5.5 million of the 7.6 million population.

4.1 Water Rates Data

We retrieve Arizona water rates data from UNC-Chapel Hill's Environmental Finance Center (EFC) for years 2014, 2015, 2017, and 2019 covering around 400 utilities. In these data we observe utility name, county, approximate population served, number of connections, type of rate structure, date of last rate change, and monthly bill amounts at different consumption levels for both residential and non-residential customers. Later in this section we detail how these bill amounts are used to compute average marginal prices.

4.2 Revenue and Expenditure

We compile revenue and expenditure data from several sources. We make public records requests for community water service (CWS) providers data.[10] For non-CWS utilities, we use operating ratio[11] if available from EFC data[12] to approximate revenue and expenditure. Otherwise, we manually collect PDF files of city and county annual financial reports through Google searches as well as the Arizona Financial Transparency Portal.[13]

4.3 Arizona Supply and Withdrawal Quantities

We submit public records requests to the Arizona Department of Water Resources (ADWR) to obtain annual utility-level supply and withdrawal quantities in acre-feet (AF). These data provide annual utility-level quantity of water supplied to single-family, multi-family, non-residential, and other customers. We also observe the quantity of water withdrawn by source type. Source types are

[10] A CWS is a public water system that serves at least 15 connections used by year-round residents or regularly serves 25 or more year-round residents. A non-CWS is a public water provider that has at least 15 service connections or regularly serves an average of at least 25 individuals daily, but does not serve year-round residential populations.

[11] Operating Ratio is operating revenue divided by expenditure including depreciation and amortization.

[12] For years 2014, 2015, and 2017 we observe the utility's operating ratio. For later years we compile revenue and expenditure data by hand from annual reports provided by the Arizona Corporation Commission.

[13] Arizona Financial Transparency Portal: https://www.azcc.gov/utilities/industry-types/water/annual-reports

groundwater (GW), surface water (SW), Colorado River, water from the Central Arizona Project (CAP), and other.[14] we match 194 utilities with rates and supply data for all four years.

To normalize non-residential water use, we obtain counts of business establishments from the Census Bureau's ZIP Business Patterns (ZBP) and allocate them to each utility's service area by area-proportional intersection of ZIP Code Tabulation Areas with the service-area boundary. Establishment counts, together with average payroll per establishment, construct the per-establishment non-residential quantity used in the non-residential demand estimation.

4.4 Central Arizona Project

We collect annual data on quantity of water purchased from CAP by entity and contract type, CAP water rates, and quantity of water stored in underground storage facilities (USF) and groundwater storage facilities (GSF) by entity through filing public records requests.

We observe the quantity of water (in AF) purchased by year, utility, and contract type. Entities holding CAP contracts are permitted to lease water to other entities. For example, the City of Phoenix which holds a Municipal & Industrial (M&I) user contract may lease some or all of its contract allocation to another entity. More commonly observed, however, are large municipalities entering into leases to purchase additional CAP water from other entities with CAP contracts. We observe the quantity of water purchased under an entity's own contract and the quantity purchased through any leases along with the leasing entity.

The contract user type determines the delivery priority. In shortage years, total CAP water will see reductions; CAP then passes these reductions along to its users in order of user priority. Tribal users and M&I users are higher in priority than the non-Indian agriculture (NIA) user group as well as the excess group. Individual tribal use contracts may have higher priority than some M&I use contracts, which illustrates why some cities choose to lease water from tribes to hedge against reductions in their own contract entitlement.

The contract user type also determines the rate paid for the water ordered. Water that falls under a contract for municipal and industrial customers pays a CAP base rate ($/AF) plus a capital charge ($/AF) that is used to pay the federal government back for the costs incurred in the construction of the canal and infrastructure used to transport water to Central Arizona.

[14] Other includes reclaimed water.

4.5 Variable Construction

From the rates data, we construct average marginal price (AMP) between each consumption threshold as in Wichman (2024):

$$AMP_{4kgal} = [Bill_{4kgal} - Bill_{0kgal}]/4, \qquad AMP_{5kgal} = [Bill_{5kgal} - Bill_{4kgal}]/1,$$
$$AMP_{10kgal} = [Bill_{10kgal} - Bill_{5kgal}]/5, \qquad AMP_{15kgal} = [Bill_{15kgal} - Bill_{10kgal}]/5,$$
$$AMP_{25kgal} = [Bill_{25kgal} - Bill_{15kgal}]/10, \qquad AMP_{50kgal} = [Bill_{50kgal} - Bill_{25kgal}]/25,$$
$$AMP_{100kgal} = [Bill_{100kgal} - Bill_{50kgal}]/50.$$

These variables allow us to capture approximate marginal price changes across different levels of consumption. We include larger AMP thresholds than Wichman (2024) because several utilities in our data set have an average monthly demand per connection greater than 25 kgal and we want to capture potential marginal price increases for large-volume non-residential customers.

4.6 Summary Statistics

Table 1 displays summary statistics for the core data needed for the empirical exercise detailed in the next section. One advantage of this setting over previous studies is that we observe utility-level annual quantity supplied. As shown in the table, about 86 percent of a utility's water supply comes from groundwater, 9 percent from CAP, 2 percent from other surface water sources, and the remaining 3 percent from reclaimed water on average. There is wide variation across utilities in average monthly supply per connection. However, the median falls in the range of average family monthly water consumption.[15] Utilities in our sample have revenues approximately 12 percent above operating expenditures on average.

The sample is primarily made up of small, groundwater-dependent utilities. The median utility serves 764 connections drawing 100 percent of its supply from groundwater, and both the 50th and 90th percentile of the groundwater share equal one. The mean connection count of 9,282 is an order of magnitude larger than the median, reflecting the presence of a small number of large metropolitan systems, including Phoenix, Tucson, and Mesa, that account for the bulk of Arizona's total water consumption despite representing a fraction of the utility count. A connection is an active

[15] The Environmental Protection Agency (EPA) estimates the average family uses more than 300 gallons of water per day or about 9,000 gallons per month (EPA, 2024).

billed customer on the utility's system. The CAP share distribution is similarly bimodal: the median utility holds no CAP contract at all, while the 90th percentile reaches 43 percent of supply. We use the full sample of 194 utilities for demand elasticity estimation and marginal cost estimation, exploiting the price and quantity variation that spans this heterogeneous population. The structural dynamic model and welfare simulations, however, are calibrated to and applied exclusively to the 21 CAP-dependent utilities. Results from the full-sample estimation inform the model's demand parameters; the model itself speaks to a narrower but policy-relevant set of utilities that face the intertemporal tradeoff at the center of the Ramsey pricing problem.

Table 1: Summary Statistics: Arizona Water Utility Panel, 2014–2019

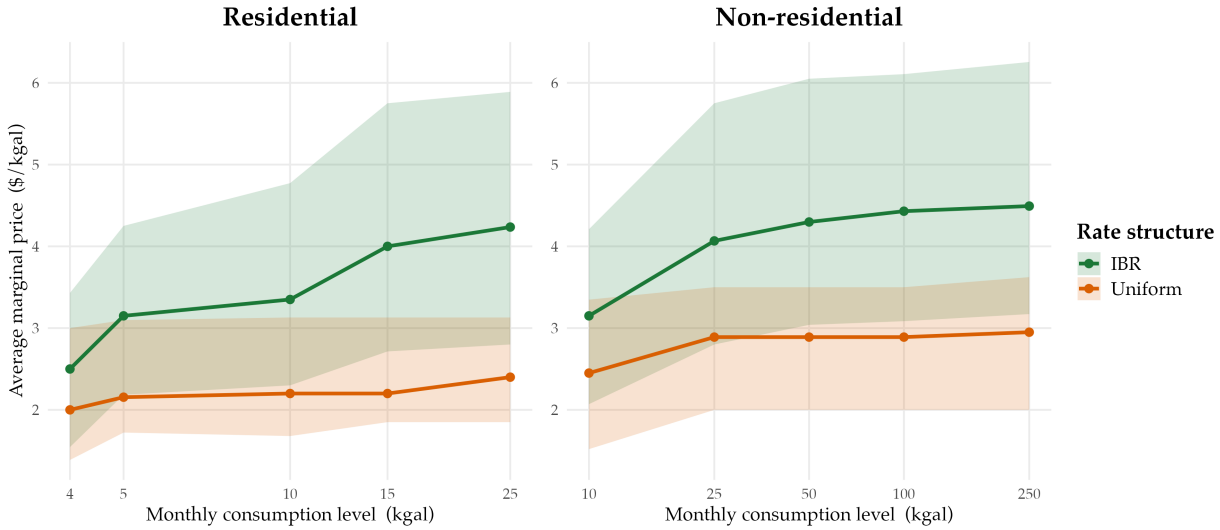
	Mean	SD	10th %ile	Median	90th %ile	N
Utility characteristics						
Service Population	30,883	145,536	198	1,922	32,740	589
# of Connections	9,282	40,253	96	764	13,185	589
Total Annual Supply (AF)	5,149	27,447	12	158	3,979	749
LTSA Balance (AF)	10,263	54,981	0	0	0	749
Supply per Connection (kgal/month)	12.45	13.05	3.65	8.62	22.44	589
Operating Ratio (Rev./Exp.)	1.12	0.47	0.73	1.07	1.52	749
Supply shares						
Groundwater Share	0.86	0.33	0.08	1.00	1.00	749
Surface Water Share	0.02	0.09	0.00	0.00	0.00	749
CAP Share	0.09	0.26	0.00	0.00	0.43	749
Reclaimed Water Share	0.03	0.13	0.00	0.00	0.07	749
Residential demand (estimation sample)						
Quantity, q_1 (AF)	3,615	16,567	16	194	3,974	526
AMP_{4kgal}	2.52	1.73	0.79	2.27	4.00	526
AMP_{5kgal}	3.30	2.04	1.20	3.00	5.50	526
AMP_{10kgal} (p_1)	3.56	2.16	1.40	3.08	5.86	526
AMP_{15kgal}	4.37	3.32	1.65	3.68	7.50	526
AMP_{25kgal}	4.46	3.34	1.68	3.82	7.56	526
AMP_{50kgal}	4.70	3.65	1.85	3.98	7.78	526
$AMP_{100kgal}$	4.84	3.84	1.88	4.14	8.75	526
Non-residential demand (estimation sample)						
Quantity, q_2 (AF)	2,122	9,254	4	136	2,982	392
Establishments (count)	685	3,550	0	5	560	392
AMP_{50kgal} (p_2)	4.57	3.03	2.00	3.98	7.42	392

Notes: Observations are utility-years (survey years 2014, 2015, 2017, 2019); N varies with missing records and because quantities and prices use the demand estimation samples (2015, 2017, 2019). *AF* denotes acre-feet (1 AF = 325.851 kgal). *LTSA balance* is banked CAP water; supply shares are of total annual supply and sum to one. q_1, q_2 are residential and non-residential water delivered; AMP_X is the average marginal price (\$/kgal) at X kgal/month, with estimation prices p_1 (10 kgal) and p_2 (25–50 kgal tier). Connections from EPA SDWIS; establishments from Census ZBP.

4.7 Price Gradients by Rate Structure

The average marginal price variable captures how utilities escalate prices with consumption, and the escalation differs sharply by rate structure. Figure 3 plots the median AMP at each consumption level for increasing block and uniform-rate utilities, separately for residential and non-residential customers. Increasing block schedules step upward as consumption rises: the median residential IBR marginal price climbs from about \$2.50 per kgal in the first block to roughly \$4.25 per kgal at 25 kgal per month, and the non-residential IBR schedule follows the same upward path. Uniform-rate utilities, by construction, hold a nearly constant marginal price across the consumption range. The modest early increase for uniform-rate utilities reflects the free initial allocation, often around 3 kgal, that many utilities bundle with the fixed access charge; the marginal price over the first few kgal therefore understates the volumetric rate that applies once the free block is exhausted. This cross-consumption variation in the marginal price is the identifying variation that both the demand estimation and the tier-level gap analysis exploit.

Figure 3: Average Marginal Price by Consumption Level and Rate Structure

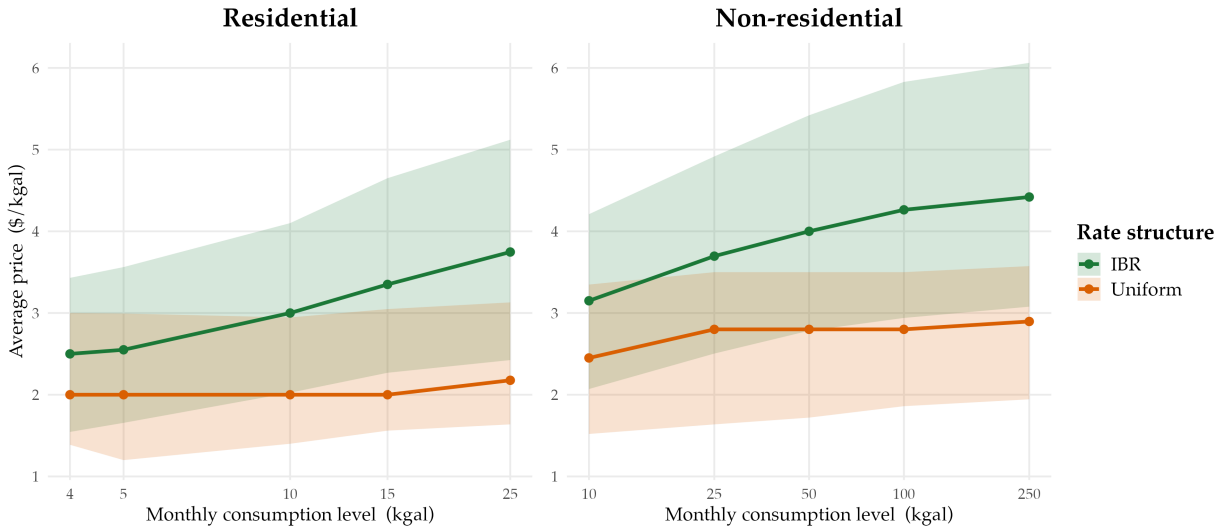


Notes: Each line plots the median average marginal price (AMP) across utility-years within a rate structure at each monthly consumption level; shaded bands are the interquartile range. Residential tiers are evaluated at 4, 5, 10, 15, and 25 kgal per month and non-residential tiers at 10, 25, 50, 100, and 250 kgal. The horizontal axis is on a log scale. IBR denotes increasing block rate and Uniform denotes uniform volumetric utilities. The sample matches the demand estimation (2015, 2017, and 2019).

The volumetric average price traces the same contrast in smoother form. Figure 4 plots the median average price, the total volumetric charge divided by consumption and net of the fixed access fee, at each consumption level. Because the average price integrates over the marginal schedule,

the IBR blocks appear as a gradual upward slope rather than discrete steps, rising from about \$2.50 to \$3.70 per kgal for residential customers, while the uniform-rate line stays close to flat near \$2 per kgal. Both figures make clear that the realized price per unit depends jointly on the rate structure and the consumption level.

Figure 4: Average Volumetric Price by Consumption Level and Rate Structure



Notes: Each line plots the median volumetric average price across utility-years within a rate structure at each monthly consumption level; shaded bands are the interquartile range. The average price at consumption L is $(\text{Bill}_L - \text{Bill}_0)/L$, the charge per kgal net of the fixed access fee. Residential average prices are reconstructed from the marginal-price tiers; non-residential average prices use the observed bills at each consumption level. The horizontal axis is on a log scale. The sample matches the demand estimation (2015, 2017, and 2019).

5 Demand Estimation

We estimate residential and non-residential demand separately by two-stage least squares, instrumenting for the endogenous marginal price with the average marginal price charged by neighboring utilities in the same county, lagged one period. The identification challenge is standard: utilities with higher demand set higher prices, therefore OLS estimates of the price coefficient are attenuated toward zero. The instrument follows the cross-market price strategy introduced by Hausman et al. (1994) and applied prominently by Nevo (2001). The identifying assumption is that utilities within the same county share common cost conditions; similar aquifer depth, comparable construction and regulatory costs, and exposure to the same county-level input price environment but set prices in response to their own local demand. A change in a neighboring utility's price therefore signals a common cost shift rather than a correlated demand shock. The leave-one-out construc-

tion, which excludes utility i from the county average, prevents any mechanical correlation between the instrument and the endogenous price. Lagging by one period eliminates contemporaneous simultaneity: a utility's own current-period demand shock cannot have caused neighboring utilities' rate decisions in the prior year, as water rate changes require formal regulatory approval and are set in advance (Hanemann, 1998). We apply the same strategy to the fixed fee, constructing a second instrument as the lagged leave-one-out county average of fixed charges. MacKay and Miller (2025) formalize the conditions under which cross-market price instruments of this type identify the price parameter, showing that identification requires the unobserved demand shock in the focal market to be mean-independent of prices in neighboring markets conditional on observables, a condition satisfied here by the combination of county-by-year fixed effects and the one-period lag. The chief remaining threat is spatially or serially correlated demand. If local demand shocks persist and spill across neighboring utilities within a county, the lagged neighbor price is not fully purged of them. We explore the potential for this identification threat later in this section and provide evidence that spatially or serially correlated demand shocks are not a concern in our setting. The general form of the instrument for sector s in utility i, county c, and year t is:

$$\bar{p}_{s,c,t-1} = \frac{\sum_{j \neq i} p_{s,j,c,t-1}}{N_{c,t-1} - 1}$$

where $N_{c,t-1}$ is the number of utilities in county c in year $t-1$. Because the instrument is lagged one period, the earliest panel year (2014) has no prior year price and is dropped from estimation; the demand samples therefore span 2015, 2017, and 2019.

The general form of the log-log (in quantity and price) demand estimation regression is:

$$\ln(q_{sict}) = \alpha_1 \widehat{\ln(p_{sict})} + \alpha_2 \widehat{\ln(F_{sict})} + X'_{ict}\gamma_1 + \delta_{ct} + \epsilon_{ict} \tag{20}$$

where q_{sict} is the equilibrium water quantity (in AF) sold in sector s of utility i in county c in year t, $\widehat{\ln(p_{sict})}$ is the fitted value of price from the first stage regression for sector s of utility i in county c in year t, $\widehat{\ln(F_{sict})}$ is the analogously fitted fixed monthly fee (both the volumetric price and the fixed fee are treated as endogenous), α_1 is the price elasticity of demand $\hat{\varepsilon}_s$ for sector s, X_{ict} are controls, δ_{ct} are county-by-year fixed effects, and ϵ_{ict} is the error term. The controls X_{ict} include median household income, which is reported for each utility's service area in the rate survey and therefore measured at the utility-year level, not as a county aggregate. Because income varies across utilities within a county-year, it is identified separately from the county-by-year fixed

effects δ_{ct}, which absorb only shifters common to all utilities in a county-year; the non-residential specification additionally controls for average payroll per establishment.

Because both the volumetric price and the fixed fee are treated as endogenous, there are two first-stage regressions, each projecting one endogenous variable on both excluded instruments–the lagged leave-one-out county price $\bar{p}_{s,c,t-1}$ and the lagged leave-one-out county fixed fee $\bar{F}_{s,c,t-1}$:

$$\ln(p_{sict}) = \pi_1 \ln(\bar{p}_{s,c,t-1}) + \pi_2 \ln(\bar{F}_{s,c,t-1}) + X'_{ict}\gamma_2 + \psi_{ct} + v_{sict}, \tag{21}$$

$$\ln(F_{sict}) = \phi_1 \ln(\bar{p}_{s,c,t-1}) + \phi_2 \ln(\bar{F}_{s,c,t-1}) + X'_{ict}\gamma_3 + \psi_{ct} + \omega_{sict}, \tag{22}$$

where all prices and fees are in logs, ψ_{ct} are county-by-year fixed effects, and v_{sict} and ω_{sict} are error terms. The fitted values $\widehat{\ln(p_{sict})}$ and $\widehat{\ln(F_{sict})}$ from these regressions enter the second stage above.

Before turning to the estimates, we illustrate the price and quantity variation that identifies the demand elasticities. Appendix Figures A1 and A2 plot water use against the average marginal price at each of several monthly consumption levels, separately for residential and non-residential customers and distinguishing IBR from uniform rate utilities. The clouds show substantial dispersion in marginal prices across utilities serving comparable quantities, the raw variation that the estimation exploits. Appendix Figures A3 and A4 localize this variation within the counties with the most utility-years in the sample, evaluated at 10 kgal per month for residential demand and the 25–50 kgal tier for non-residential demand. The price is measured at a single consumption benchmark: for an IBR utility this is one marginal price on its tiered schedule, rather than the full block structure. Therefore, demand estimates identify the response to the marginal price at a fixed benchmark rather than to the shape of the schedule. Within a given county, utilities charge markedly different marginal prices at the same consumption level. This cross-utility, within-county dispersion, purged of common cost conditions by the county-by-year fixed effects and instrumented by neighboring utilities' lagged rates, is the source of identification.

The principal threat to the exclusion restriction is a demand shock correlated across neighboring utilities, and in this setting the most salient such shock is drought. A dry year increases outdoor water demand throughout a county while also pushing rates upward, meaning the lagged neighbor price could proxy for common weather-driven demand rather than common cost. We test this channel directly by adding the utility's own service-area precipitation and the lagged county-peer precipitation to the demand equation. The residential price elasticity is nearly identical, moving from -0.53 to -0.51, while the non-residential estimate becomes slightly smaller in magnitude

from -0.87 to -0.72 and the coefficients on the precipitation terms are small and insignificant (Appendix Tables A6 and A7). That the elasticity does not respond to controlling for the drought channel indicates the instrument reflects common cost rather than correlated demand. Because persistent demand heterogeneity leaves the residual serially correlated within utilities, I cluster all standard errors at the utility level and defend the exclusion restriction through the precipitation controls directly rather than through the one-period lag alone.

5.1 Residential Water Demand

Table 2 presents estimates of the residential demand equation. We treat both volumetric price (AMP_{ict} and AVP_{ict}) and fixed monthly fee as endogenous, instrumenting for each with the corresponding lagged county-mean price among all other utilities in the same county. Both the average marginal price (AMP_{ict}) and average volumetric price (AVP_{ict}), each evaluated at a single consumption level of 10 kgal per month, yield similar elasticity estimates of -0.53 and -0.46, respectively. The instruments are strong in both first stages. We report that the Cragg-Donald minimum eigenvalue statistic is more than an order of magnitude above the Stock and Yogo (2005) 10 percent maximal-bias critical value. We also report the Angrist and Pischke (2009) F-statistics, which test each excluded instrument in its own first stage holding the other constant. First stage regression results appear in Appendix Table A2.[16] The fixed fee elasticity estimates are statistically insignificant and we cannot conclude that the monthly fixed fee alters consumption levels. The income elasticity is 0.93, consistent with water being a normal good with moderate income sensitivity. We report the fixed fee-inclusive total bill average price (AP_{ict}) demand elasticity at 10 kgal in Column (4). Consistent with previous studies, we find that households are more responsive to the average price than marginal price of water. We use $\hat{\varepsilon}_1 = -0.53$ throughout the model calibration for residential water demand elasticity.

The direction of OLS bias is as expected. The OLS estimate of the price coefficient is -0.51 in Column (1), less elastic than the IV estimate, reflecting the standard attenuation bias from endogenous rate-setting: utilities facing high demand set higher prices, attenuating the measured price response toward zero. The IV estimate corrects for this simultaneity by isolating variation in marginal prices driven by changes in neighboring utilities' rates rather than by demand conditions within the utility itself.

[16] The Cragg-Donald statistic is the minimum eigenvalue statistic of Cragg and Donald (1993); the Angrist-Pischke F is cluster-robust.

Table 2: Residential Water Demand (Two Endogenous Variables)

	OLS	IV		
	(1)	(2)	(3)	(4)
$ln(AMP_{ict})$	-0.51***			
	(0.12)			
$ln(\widehat{AMP_{ict}})$		-0.53***		
		(0.19)		
$ln(\widehat{AVP_{ict}})$			-0.46**	
			(0.19)	
$ln(\widehat{AP_{ict}})$				-0.67***
				(0.21)
$ln(\widehat{FixedFee_{ict}})$	0.10	-0.05	-0.14	
	(0.10)	(0.16)	(0.15)	
ln(Median Income)	0.93***	0.92***	0.93***	0.91***
	(0.16)	(0.15)	(0.15)	(0.16)
Observations	526	526	526	526
County×Year FE	✓	✓	✓	✓
Cragg-Donald F		126.6	132.7	260.1
AP F: Price		20.6	22.3	26.3
AP F: Fixed Fee		67.3	63.4	

Notes: Dependent variable is log residential water delivered per capita; Arizona utilities, 2014–2019. AMP and AVP are instrumented jointly with the fixed fee; AP enters as a single endogenous regressor. Standard errors clustered by utility in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.10.

5.2 Non-Residential Water Demand

Credible non-residential water demand estimates are scarce. Establishment-level water use is rarely observed at scale, so most of the pricing literature estimates residential demand alone and assigns the non-residential sector a borrowed or assumed elasticity. The rate-survey panel, combined with establishment counts, lets us estimate non-residential demand directly across the Arizona utilities, and that estimate proves consequential for both the aggregate conservation gain and the incidence of the pricing correction. We estimate non-residential demand following the same IV strategy, treating the non-residential volumetric price and fixed fee as endogenous and instrumenting with their lagged county-mean counterparts. The dependent variable is the log of non-residential water delivered per non-residential establishment, where establishment counts are from the Census ZIP Business Patterns survey allocated to utility service areas by area-proportional intersection. This quantity-per-establishment measure standardizes for differences in the non-residential density of utility service areas and isolates the intensive margin response of individual establishments to price.

Table 3 presents results for three alternative price measures. Column (2) uses the average marginal price in the 25–50 kgal per month consumption tier, the range most representative of moderate-scale non-residential users; Column (3) uses the 10–25 kgal tier; Column (4) uses the volumetric average price at 50 kgal; Column (5) uses the total bill average price at 50 kgal. Column (1) presents the OLS baseline. All four IV specifications yield price elasticities in a tight range, from -0.70 to -0.87. The Angrist and Pischke (2009) first-stage F-statistics are above conventional thresholds for both the price and the fixed fee and we report the first stages in Appendix Table A3. In contrast to the residential estimates, the average price elasticity is smaller than the marginal price. One possible explanation for this is that non-residential consumers use more water and therefore the fixed fee makes up a smaller share of their total bill. We use $\hat{\varepsilon}_2 = -0.87$ from the preferred Column (2) specification in the model calibration for non-residential water demand elasticity.

Non-residential demand is more elastic than residential demand at -0.87 versus -0.53. One interpretation is that non-residential establishments treat water as an input cost subject to profit maximization rather than a consumption good, making them more attentive to marginal prices. A business that monitors operating costs through accounting has stronger incentives to reduce water-intensive processes, adjust scheduling, or adopt water-saving equipment when the price of a variable input rises. Residential users, by contrast, may be less price-responsive at the margin

Table 3: Commercial Water Demand (Per Establishment)

	OLS	IV			
	(1)	(2)	(3)	(4)	(5)
$ln(AMP_{ict}^{com})$ (25–50 tier)	-0.72**				
	(0.33)				
$ln(\widehat{AMP_{ict}^{com}})$ (25–50 tier)		-0.87**			
		(0.36)			
$ln(\widehat{AMP_{ict}^{com}})$ (10–25 tier)			-0.76**		
			(0.32)		
$ln(\widehat{AVP_{ict}^{com}})$ (vol, 50 kgal)				-0.84**	
				(0.35)	
$ln(\widehat{AP_{ict}^{com}})$ (fee-incl, 50 kgal)					-0.70*
					(0.36)
$ln(\widehat{FixedFee_{ict}^{com}})$	0.32	0.64*	0.41	0.55	
	(0.34)	(0.37)	(0.37)	(0.36)	
ln(Payroll per Establishment)	-0.29	-0.26	-0.27	-0.27	-0.35
	(0.28)	(0.28)	(0.29)	(0.29)	(0.28)
ln(Median HH Income)	-2.1***	-2.0***	-2.1***	-2.1***	-2.1***
	(0.58)	(0.57)	(0.58)	(0.58)	(0.60)
Observations	392	392	392	392	392
County×Year FE	✓	✓	✓	✓	✓
Cragg-Donald F		72.9	62.9	69.7	142.7
AP F: Price		28.6	29.3	31.4	33.1
AP F: Fixed Fee		32.5	30.5	31.8	

Notes: Dependent variable is log non-residential water per establishment (kgal/yr); Arizona utilities, 2014–2019. Column (1) is OLS; IV columns treat the price and fixed fee as endogenous, instrumented by lagged county-year leave-one-out prices. Columns (2)–(3) use AMP (25–50, 10–25 kgal/month); (4) AVP; (5) fee-inclusive $AP = \text{Bill}_{50}/50$. SEs clustered by utility. Significance: ***p<0.01, **p<0.05, *p<0.10.

because water expenditure is a small share of household income, and there is evidence that the perceived price signal is muted by the complexity of block-rate billing structures (Ito, 2014; Wichman, 2014). If residential IBR customers respond to average rather than marginal prices, the conservation advantage of IBR over uniform pricing would be smaller than the tier-structure analysis implies, a caveat that applies equally to the estimated elasticity, which is identified from marginal price variation, and to the welfare comparisons in Section 7.5 that use it. The fixed fee elasticity of 0.64 in Column (2) is imprecisely estimated, reflecting the limited variation in non-residential fixed charges and consistent with statistically insignificant estimates in the residential demand specification. The income coefficient ranges from -2.0 to -2.1 across specifications, indicating that utilities serving higher-income areas have substantially lower non-residential water use per establishment. This is better interpreted as a compositional effect than a true income response: wealthier Arizona service areas, concentrated in Maricopa and Pima counties, are dominated by less water intensive businesses such as offices, healthcare, and retail while lower-income rural areas contain more food processing, light manufacturing, and construction supply operations. Service-area median household income, measured at the utility-year level, serves here as a reduced form proxy for industrial composition rather than an income shifter for non-residential demand per se. We construct service-area-level manufacturing and food service establishment shares from NAICS-level ZIP Business Patterns data, treating utilities with no establishments in a sector as having a share of zero. Adding these controls to the preferred specification leaves the estimation sample unchanged and moves the price elasticity only slightly, to -0.94, well within the range of sampling variability across the three IV price-tier specifications, while the income coefficient is essentially unchanged. The stability of the income coefficient conditional on industry controls reflects the fact that the county-by-year fixed effects already absorb most county-level industrial composition variation; the remaining within-county utility-level variation in sector shares provides little additional information. The price elasticity is therefore robust to direct controls for service area industrial composition. The income coefficient is not used in model calibration; only $\hat{\varepsilon}_2 = -0.87$ enters the SDP. Non-residential customers consume nearly one third of delivered water in our sample and their more elastic response materially shapes the aggregate. A residential-only elasticity applied to the model would understate the quantity response and total welfare gain from pricing at the optimum, and it would misstate the incidence of the correction.

As a robustness check, Appendix Tables A4 and A5 present results restricting to county-years in which at least five utilities report valid prices. The leave-one-out instrument for each utility is then

constructed from at least four peers; this restriction excludes 17 residential and 28 non-residential utility-years. The elasticity estimates for both residential and non-residential customers stay within a close range of the full sample estimates.

6 Marginal Cost Estimation

The optimal pricing formula in equation (18) requires utility-year estimates of marginal cost c_i. We recover these from a log-linear cost function estimated on the panel of utility-year operating expenditures, following the approach in Wichman (2024). Estimating this elasticity directly from cost and quantity data therefore yields a utility-specific marginal cost without requiring observation of input prices.

6.1 Specification

We estimate separate cost regressions for AMA and non-AMA utilities: AMA utilities operate under the 100-year assured water supply requirement and tend to be larger, while non-AMA utilities are usually smaller and rely predominantly on groundwater extraction. Pooling the two groups would require strong homogeneity assumptions across very different cost structures. Within the AMA sample, we interact the log-quantity term with a CAP-user indicator to allow the cost elasticity to differ for utilities with active CAP contracts, which face the additional variable cost of M&I contract water purchases.

The estimating equation is:

$$\ln(E_{it}) = \beta_1 \ln(Q_{it}) + \beta_2 \left[\ln(Q_{it}) \times \mathbf{1}(\text{CAP}_i)\right] + \gamma \, \text{Precip}_{it} + \eta_{ct} + \kappa_{g(i)} + \varepsilon_{it}, \tag{23}$$

where E_{it} is annual operating expenditure for utility i in year t (dollars), Q_{it} is total annual supply (kgal), $\mathbf{1}(\text{CAP}_i)$ is an indicator equal to one if the utility holds a CAP contract, Precip_{it} is service-area precipitation (inches), η_{ct} are county-by-year fixed effects that absorb common variation in input costs, drought severity, and regulatory conditions, and $\kappa_{g(i)}$ are customer size-group fixed effects for small (<1,000 connections), medium (1,000–5,000), and large (>5,000) utilities, following Wichman (2024). Standard errors are clustered by utility to account for serial dependence in costs within utilities over time. The non-AMA regression omits the CAP interaction.

From equation (23), the implied marginal cost is:

$$\widehat{MC}_{it} = \hat{\beta}_1 \cdot \frac{E_{it}}{Q_{it}} = \hat{\beta}_1 \cdot AC_{it}, \tag{24}$$

where $\hat{\beta}_1$ is the estimated cost elasticity from the regression for utility i's regulatory group (AMA or non-AMA) and $AC_{it} = E_{it}/Q_{it}$ is the observed average cost. Because the AMA CAP interaction is statistically insignificant (Table 4), we set $\hat{\beta}_2 = 0$, so marginal cost is the product of the group cost elasticity and average cost. The fixed cost component that must be recovered through the Ramsey markup is:

$$\hat{\tau}_{it} = (1 - \hat{\beta}_1) \cdot E_{it}, \tag{25}$$

which is the dollar amount of expenditure not attributable to variable production costs and which enters the budget constraint in equation (3) as the period fixed cost τ.

6.2 MC Estimates

Table 4 presents the cost function estimates. In the AMA sample (Column 2), the cost elasticity with respect to output is $\hat{\beta}_1 = 0.858$, implying that a 1 percent increase in total supply is associated with a 0.86 percent increase in operating expenditure. The CAP interaction is small and statistically insignificant ($\hat{\beta}_2 = 0.018$, $SE = 0.013$), so we set it to zero and apply the common AMA elasticity $\hat{\beta}_1 = 0.858$ to CAP and non-CAP utilities alike. Non-AMA utilities (Column 3) exhibit greater economies of scale, with a cost elasticity of $\hat{\beta}_1 = 0.544$, consistent with small groundwater-dependent utilities operating with high fixed infrastructure costs relative to their volume of supply. This pattern is consistent with the cost structure of groundwater-dependent systems, in which aquifer rights represent a one-time investment and variable costs are limited to pumping, treatment, and delivery.

The implied marginal cost estimates have a median of \$3.52 per kgal for AMA utility-years and \$5.93 per kgal for non-AMA utility-years. For AMA utilities, the small difference between marginal and average cost ($\hat{\beta}_1 = 0.858$ implies $MC \approx 0.86 \times AC$) means the fixed cost component is modest and the optimal Ramsey markup required for budget balance is small relative to the marginal cost. For non-AMA utilities, the larger gap between MC and AC ($\hat{\beta}_1 = 0.544$) means the Ramsey markup would account for roughly 46 percent of average cost.

Two features of this cost calibration bear on how the recovered scarcity rent should be read.

Table 4: Log-Linear Cost Function Estimates (Own-Water Cost)

	Pooled (1)	AMA (2)	Non-AMA (3)
$\ln(Q_{it})$	0.600***	0.858***	0.544***
	(0.062)	(0.053)	(0.080)
Precipitation (in.)	0.005	-0.007	0.004
	(0.016)	(0.024)	(0.019)
$\ln(Q_{it}) \times$ CAP user	0.045***	0.018	
	(0.012)	(0.013)	
Observations	543	152	391
R^2	0.91	0.96	0.82
County-Year fixed effects	✓	✓	✓
Size-group fixed effects	✓	✓	✓

Notes: Dependent variable: $\ln(E_{it}^{\text{own}})$, log annual expenditure net of CAP purchases (ACC income statements). County×Year and customer size-group fixed effects, where size groups are small (<1,000 connections), medium (1,000–5,000), and large (>5,000), following Wichman (2024). SE clustered by utility. $MC_{it} = \hat{\beta}_1(E_{it}/Q_{it})$ and $\tau_{it} = (1 - \hat{\beta}_1)E_{it}$; the CAP interaction $\hat{\beta}_2$ is insignificant and set to 0. Significance: ***p<0.01, **p<0.05, *p<0.10.

First, expenditure and output are jointly determined: utilities that plan for larger output volumes also authorize larger operating and capital budgets, so higher Q predicts higher E partly through budget-setting rather than through production technology alone, biasing $\hat{\beta}$ upward. Through the fixed-cost channel this makes the scarcity rent a conservative lower bound: an upward-biased $\hat{\beta}$ understates the fixed cost $\hat{\tau} = (1 - \hat{\beta})E$, relaxing the budget constraint in the dynamic program and lowering the model-optimal price and hence the recovered approximate scarcity rent $\bar{\lambda}/\bar{\mu} = p^* - AC$. An upward-biased $\hat{\beta}$ simultaneously raises implied marginal cost, which pushes the optimal price up, so we assess the net effect directly by re-solving the model across a range of $\hat{\beta}$ in the calibration-sensitivity analysis. Second, unlike Wichman (2024), who recovers marginal cost from operating expenditure alone and treats $\hat{\beta} \cdot AC$ as a short-run lower bound on long-run marginal cost, we measure expenditure inclusive of depreciation and amortization, so $\widehat{MC}$ already embeds capital costs and sits closer to long-run marginal cost.

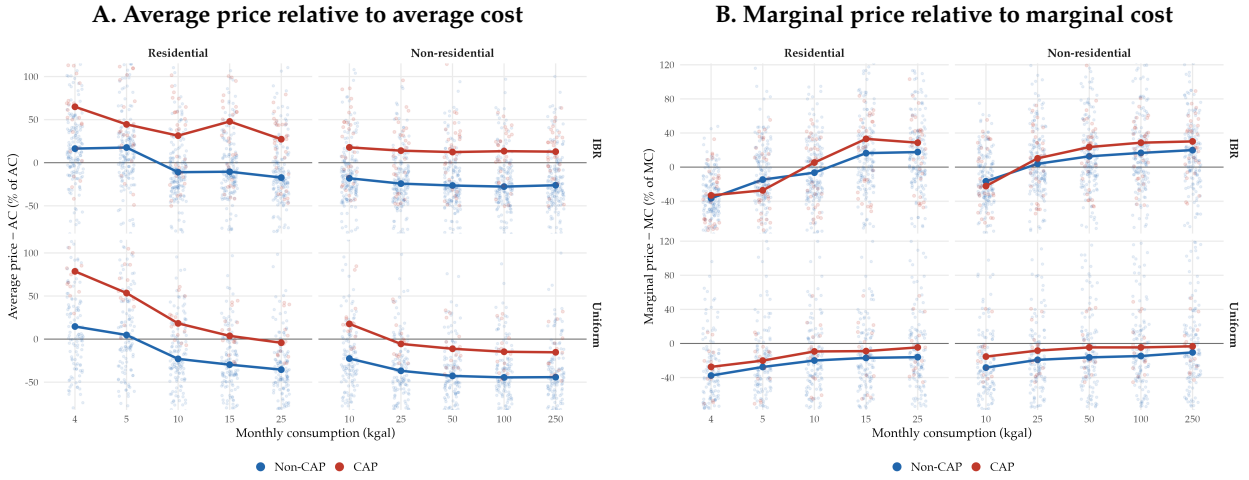
7 Results

7.1 Observed Prices versus Cost Recovery

Before looking at model-derived optimal pricing, we begin by comparing observed prices to the cost of providing water. Figure 5 plots two gaps across the consumption distribution. Panel A displays the difference in average price a customer pays relative to the utility's average cost in percentage terms. Panel B displays the difference in marginal price and marginal cost. Two patterns emerge. Looking at cost recovery in Panel A, the fixed fee makes low-use bills look cost-covering but recovery erodes as usage rises and is systematically worse for uniform and non-CAP utilities. For the median CAP IBR utility, non-residential average prices are almost exactly cost recovering across all consumption tiers.

Panel B evaluates whether prices follow marginal costs. Observed volumetric rates underprice the marginal unit almost everywhere. The caveat is that under IBR, low-usage customers are subsidized by the highest consumption tiers. Only high IBR consumption tiers send a conservation signal above marginal costs. Volumetric prices are persistently below MC for uniform rate utilities. Across all of Panel B, non-CAP and CAP pricing patterns are nearly indistinguishable from one another.

Figure 5: Observed Prices Relative to the Cost of Service



Notes: Gap between observed prices and the cost of service across monthly consumption levels, by rate structure and sector. Panel A: average price less average cost (percent); Panel B: average marginal price less estimated marginal cost (Section 6, percent). Points are utility-years; solid lines are group medians; red denotes CAP and blue non-CAP utilities. The zero line marks cost recovery (A) and marginal-cost pricing (B); points below it price below cost. Axes trimmed to the 5th–95th percentile. Sample: Arizona utility-years, 2014, 2015, 2017, and 2019.

7.2 Supply Shock Calibration

The supply shock θ is calibrated to the Central Arizona Project's shortage-sharing framework from the recent past as well as ambiguous future management scenarios. CAP delivers roughly 1.5 million acre-feet per year across all user classes, with its allocation tied to Lake Mead's elevation through the Bureau of Reclamation's tiered shortage mechanism. Under the 2019 Drought Contingency Plan (DCP), in place through 2026, a Tier 2a/2b shortage removes approximately 600,000 acre-feet and a Tier 3 shortage approximately 720,000 acre-feet from CAP's supply (Western Resource Advocates, 2023; Central Arizona Project, 2026). These supply reductions are not evenly allocated. The DCP reduces entitlements in strict priority order, where non-Indian agricultural (NIA) and excess pools absorb the first cuts, while M&I subcontracts are largely protected until Tier 3.

A utility's exposure therefore depends on its contract portfolio, not on the aggregate CAP cut. Large municipalities hold a mix of contract types. For example, 65% of the City of Phoenix's CAP water is delivered from M&I subcontracts and roughly 31% comes from NIA pools. A less severe shortage that does not impact M&I still reduces a utility's effective deliveries through its NIA and excess holdings. The stressed shock regime represents this portfolio-weighted exposure under a Tier 3 or more severe shortage scenario. The shock distribution $\theta \sim \text{Beta}(4, 1)$ has mean 0.80 and standard deviation 0.16, placing the most mass on moderate cuts while assigning roughly a six percent probability to deliveries below half of contract. A left tail draw is representative of an outcome a utility with substantial NIA holdings would face in a deep-shortage year. The mild regime, $\theta \sim \text{Beta}(50, 1)$ with mean 0.98, reflects the near-full M&I deliveries observed over the 2014–2019 sample, before the first shortage declaration.

The DCP expires in 2026 and the post-2026 Lower Basin operating guidelines remain unsettled. Facing record low Lake Mead elevation levels, initial federal guidelines require Arizona to cut total Colorado River use by roughly 760,000 acre-feet per year, comparable to a Tier 3 reduction. The stressed regime should thus be read as a stand-in for the uncertain severity and incidence of future shortages—anchored to the realized DCP framework but spanning the range of portfolio exposures utilities may face.

7.3 Optimal Prices

Table 5 reports the model-implied optimal prices for the customer sectors across the rate structure and CAP status groups. Residential optimal prices are evaluated at 10 kgal per connection per month and non-residential at 25 to 50 kgal per connection per month, matching the consumption benchmarks used throughout the demand estimation. For CAP utilities, p_1^* and p_2^* are the 2019 welfare-simulation optima under the mild supply-shock regime representative of the 2014–2019 sample period, the path-mean optimal price starting from each utility's observed 2019 reserves and water stock; for non-CAP utilities, the scarcity component is zero and the benchmark collapses to average cost. Optimal prices for non-CAP utilities are higher than those of CAP utilities, reflecting the higher average costs outside CAP service areas. These figures are economically plausible: against median observed prices of \$2.91 to \$4.75 per kgal, the model-implied optimum sits roughly 20 to 30 percent above prevailing residential rates for CAP utilities and from modestly below to about 20 percent above for non-residential rates, and more than double for the high-cost non-CAP utilities, whose benchmark is average cost rather than a scarcity-inclusive optimum. Non-CAP utilities are also substantially smaller in quantity delivered annually. Column (4) shows CAP utilities deliver orders of magnitude more water per year than non-CAP utilities. This scale difference allows utility managers to spread fixed costs over many users and charge much lower rates.

The optimal price decomposes into the three terms of the pricing rule in equation (18): marginal cost, a Ramsey markup that recovers fixed costs through the volumetric rate, and the scarcity rent $\bar{\lambda}/\bar{\mu}$ that prices the shadow value of stored water. The scarcity rent is the natural-capital component the model recovers and is absent from any static cost-of-service benchmark. Across the CAP utilities its distribution is right-skewed, with a median of \$0.52 per kgal and a mean of \$1.17: it is near zero for utilities holding ample banked water and rises to nearly \$6 per kgal for those closest to the resource constraint. This term is what a scarcity-inclusive optimum adds to price beyond cost recovery, and these magnitudes are conditional on the model's calibration.

The smaller utilities without access to CAP water exhibit higher average optimal prices and substantially greater variance. The optimal price standard deviation is roughly the magnitude of the average prices for non-CAP utilities across rate structure. This is because the unit cost of supplying water varies substantially across utilities and makes up the largest portion of the optimal price composition. Mean optimal prices are also greater than those of CAP utility peers across both sectors. This likely reflects the scale difference and the more limited set of customers over which

Table 5: Model-Implied Optimal Prices by Sector and Rate Structure

	Mean	Median	SD	n	Avg. Demand (AF)
Panel A: *Uniform rates, CAP utilities*					
Residential optimal, p_1^*	3.62	3.43	0.96	6	21,589
Non-residential optimal, p_2^*	3.52	3.51	0.88	6	
Panel B: *Uniform rates, non-CAP utilities*					
Residential optimal, p_1^*	7.88	5.97	6.03	50	573
Non-residential optimal, p_2^*	5.85	4.57	4.39	50	
Panel C: *IBR rates, CAP utilities*					
Residential optimal, p_1^*	5.00	4.06	2.59	15	44,787
Non-residential optimal, p_2^*	4.77	4.15	2.31	15	
Panel D: *IBR rates, non-CAP utilities*					
Residential optimal, p_1^*	8.62	7.04	5.73	65	870
Non-residential optimal, p_2^*	6.57	5.74	4.21	65	

Notes: Model-implied optimal prices in \$/kgal, 2019. For CAP utilities (Panels A, C), p_1^* and p_2^* are the welfare-simulation optima; for non-CAP (Panels B, D), the two-sector Ramsey cost-recovery solution ($p_1^* > p_2^*$). n counts utilities; *Avg. Demand* is mean annual water demanded (AF), shown once per panel. Sample: $AC_{it} \leq$ \$30/kgal.

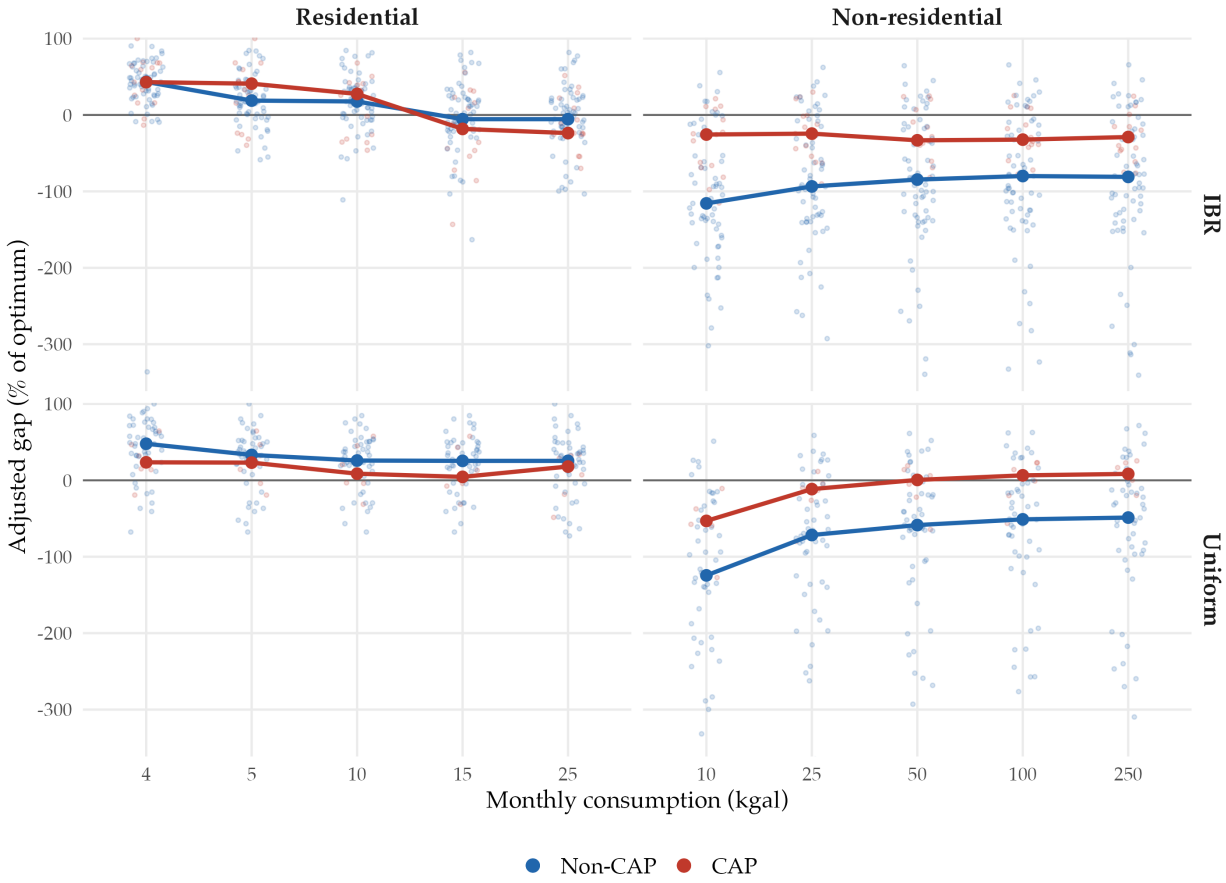
to spread fixed costs.

7.4 Observed Prices Relative to Optimal Benchmark

Figure 6 plots the median adjusted price gap, expressed as a percent of the scarcity-inclusive optimum, against consumption level for residential (left panel) and non-residential (right panel) sectors, separately by rate structure and CAP status. The adjusted gap uses $(AC - MC) = (1 - \hat{\beta}) \cdot AC$ as an observable proxy for the fixed-cost recovery component of p^*, equal to the Ramsey markup when observed prices approximate marginal cost. This fixed cost proxy is subtracted before comparing to the observed average marginal price, yielding a volumetrically comparable benchmark of approximately $MC +$ scarcity component.

The gap from observed to optimum is largely driven by pricing differences across sector groups not rate structure types. Figure 6 shows that observed non-residential prices vary from approximately optimal to nearly double the optimum, with uniform rate utilities having a smaller gap across consumption levels. Residential rates for IBR utilities are consistent with the well-documented pattern that low-use customers face prices well below efficient and the gap erodes at higher consumption tiers. Residential uniform utility rates are roughly 20% below the model-implied efficient

Figure 6: Median Adjusted Price Gap as a Percent of the Optimum, by Consumption Level, Rate Structure, and CAP Status



Notes: Each line plots the median adjusted gap as a percent of the scarcity-inclusive optimum, $100 \times [p^* - (AC - MC) - \text{Avg. MP}]/[p^* - (AC - MC)]$, where a positive value means the marginal price lies that many percent below the optimum. $(AC - MC)$ proxies the fixed-cost recovery component. Residential gaps (left) are evaluated at 4, 5, 10, 15, and 25 kgal/month; non-residential (right) at 10, 25, 50, 100, and 250 kgal/month. Dark solid lines are CAP utilities, light dashed non-CAP; red denotes IBR and blue uniform rates. For non-CAP utilities the scarcity component is zero and the benchmark reduces to MC. CAP optima are the 2019 welfare-simulation prices; sample is IBR and uniform-rate utilities with $AC \leq \$30$/kgal in 2019.

price, and the gap to optimum is consistent across consumption levels. Appendix Table A8 displays the price gap to optimum in relative percent terms across rate structures and CAP status.

7.5 Welfare Simulation: Status Quo versus Optimal Pricing

Adjusting observed prices to the optimum yields large welfare gains and substantial quantity conservation across both shock regimes and both horizons we consider. We simulate shock paths of five and one hundred years for each of the 21 CAP utilities in the sample. Appendix Table A9 confirms that IBR and uniform utilities in this group are broadly comparable on observable characteristics: connections, costs, quantities, revenue, and service-area income are all statistically indistinguishable across rate structure types. IBR utilities do charge higher non-residential prices (median \$3.92 versus \$3.45 per kgal), hold somewhat larger CAP entitlements, and changed residential rates more frequently over the sample period. Because each utility chooses its own rate structure, the IBR–uniform comparison throughout is descriptive: observable characteristics are balanced across the two groups, but we cannot rule out that unobserved determinants of rate-structure adoption also shape where prices fall relative to the optimum.

The simulation proceeds as follows. Starting from the observed 2019 long-term-storage balance as the initial water stock and the observed 2019 operating surplus as the initial financial reserve, we draw sequences of CAP supply shocks $\{\theta_t\}$ from a calibrated Beta distribution and propagate the state forward under each pricing scenario over horizons of five and one hundred years. Under the optimal scenario, we evaluate the value-function-iteration policy at each realized state (w_t, b_t) to obtain the period-t optimal prices and CAP purchase. The status-quo benchmark holds prices at their observed 2019 levels but purchases CAP water under the same optimal rule, carrying revenue forward identically, so the contrast isolates the effect of pricing rather than of procurement. We repeat each simulation for 1,000 shock sequences under two shock regimes: a mild regime representing recent normal conditions and a stressed regime calibrated to conditions more severe than Colorado River Tier 3 shortage declarations.[17] we measure the welfare change from optimal pricing as the total flow surplus it delivers, $\Delta W = \Delta CS + \Delta NR$: the change in discounted consumer surplus plus the change in discounted operational net revenue along a shock path, each the optimal-minus-status-quo difference. To place the two horizons on a common footing, we express ΔW as a constant per-year annuity per connection, the T-year discounted present value scaled by

[17] The mild regime draws $\theta \sim \text{Beta}(50, 1)$, implying mean CAP delivery of approximately 98 percent of contract. The stressed regime draws $\theta \sim \text{Beta}(4, 1)$, implying mean delivery of 80 percent of contract and approximately 6 percent probability of delivery below half of contract.

$(1-\beta)/(1-\beta^T)$.

The unit cost of purchased CAP water, p_x, is calibrated differently across the two regimes to reflect how shortage conditions affect delivery pricing. Under the mild regime p_x starts from the 2019 observed M&I rate of \$0.485 and grows at 2.5 percent per year with inflation, following the same deterministic path in every year regardless of the shock draw. Under the stressed regime, p_x is tied to each period's shock realization: $p_x(\theta_t) = 0.90 + (1 - \theta_t)$, so each reduction in contract delivery raises the per-unit cost proportionally. This functional form is motivated by the post-2022 CAP rate data shown in Figure A6 in the appendix. Delivery rates exhibited no statistically significant relationship with delivery volume during the pre-shortage period from 2014 to 2021, but from 2022 onward under Drought Contingency Plan conditions, delivery reductions and rate increases were tied to the same shortage tier, generating a strong negative correlation between deliveries and price per unit that was absent before. At the mean stressed delivery level ($\theta = 0.80$), this implies $p_x = \$1.10$ per kgal, roughly double the mild-regime rate. The stressed scenario therefore applies pressure through two simultaneous channels. The same shock θ that limits available purchase volume ($\bar{x}(\theta_t) = \bar{x} \cdot \theta_t$) also raises the per-unit cost, amplifying the budget constraint's bite and increasing the marginal value of demand reduction through higher prices.

Table 6: Welfare Change from Optimal Pricing, by Rate Structure (per Connection-Year)

Rate structure	Mild	Stressed
Panel A: 5-year horizon		
IBR	9.3	10.8
Uniform	11.1	18.5
All	9.6	12.0
Panel B: 100-year horizon		
IBR	14.3	21.9
Uniform	30.0	77.8
All	16.7	30.6

Notes: Connection-weighted mean flow welfare change $\Delta W = \Delta CS + \Delta NR$ from optimal pricing (\$/conn/yr), as a constant per-year annuity so the 5- and 100-year panels are comparable. Sample is the 21 CAP utilities. The status quo prices at observed tariffs, isolating the pricing effect.

Table 6 reports the connection-weighted mean flow welfare change per connection-year, ΔW, for each rate structure and supply regime at both horizons. Optimal pricing raises welfare by \$9.6

per connection-year in the mild regime and $12.0 in the stressed regime over five years, rising to $16.7 and $30.6 over one hundred years. The long-run gains exceed the short-run ones because the dynamic-Ramsey advantage of banking reserves to fund CAP and smooth scarcity materializes only over a long horizon. Uniform-rate utilities gain more than increasing-block utilities at both horizons, and the gap widens over the long run. Over one hundred years, uniform utilities gain roughly twice as much as increasing-block utilities under mild conditions and more than three times as much under stress. This ordering is consistent with the pricing-gap evidence above, where IBR upper tiers already track the scarcity-inclusive optimum more closely. With only six uniform-rate utilities and wide dispersion in the group, we cannot distinguish the two rate structures statistically, so we read the difference as suggestive rather than decisive and pool rate types for the remainder of the results.

We note three considerations on the scope of the results. First, the observed price for IBR utilities is constructed as the average marginal price at a fixed consumption level of 10 kgal per month, a single-point proxy for the full tier schedule; utilities with steep upper tiers will appear more underpriced than they truly are in this framework, because the AMP at 10 kgal understates the price paid by heavy users. Second, the model-optimal price is a uniform-equivalent rather than an optimal IBR tier schedule; a model that solved directly for the optimal block structure would quantify the within-schedule distributional effects. Third, the welfare and conservation magnitudes rely on extrapolating the constant-elasticity demand estimates beyond the range of price variation used to identify them, because the optimum sits well above observed prices for many utilities. Each of these qualifications bear on the magnitude or utility-level interpretation of the gaps, but do not threaten the aggregate results in Tables 6 and 7.

7.6 Aggregate Welfare and Conservation

The headline payoff of efficient pricing is a conservation effect delivered through prices alone. Scaling each utility's connection-weighted per-connection change by its connections and summing across the 21 CAP utilities, which together serve roughly 1.26 million connections, optimal pricing raises welfare by approximately $21.0 million per year in the mild regime and $38.5 million in the stressed regime over a hundred-year horizon, and by $12.0 million and $15.0 million over five years (Table 7). Over 100-year simulations, demand is cut by roughly 85,000 acre-feet per year under mild conditions and 146,000 under stress.

Table 7: Welfare and Quantity Change from Optimal Pricing

	5-year		100-year	
	Mild	Stressed	Mild	Stressed
ΔW (\$M/yr)	12.0	15.0	21.0	38.5
ΔQ (AF/yr)	-75,574	-95,929	-85,101	-146,455
ΔX CAP (AF/yr)	-16,300	-21,390	-89,655	-106,533

Notes: Sample total across all 21 CAP utilities. $\Delta W = \Delta CS + \Delta NR$ is the flow welfare change as a per-year annuity (\$M/yr); ΔQ (water demanded) and ΔX (CAP purchased) are annual quantities (AF/yr; negative = reduction). Benchmark as in Table 6: observed prices, optimal CAP.

The conservation dividend falls directly on Central Arizona Project water. Optimal pricing lowers CAP purchases by roughly 90,000 acre-feet per year under mild conditions and 107,000 under stress, 14 to 17 percent of the region's municipal and industrial entitlement[18] when prices efficiently signal scarcity. These figures should be read as back-of-the-envelope: they scale the simulated per-utility effects to the full connection base of the 21 modeled utilities, hold the estimated demand elasticities fixed, and assume the model-optimal price is implemented in full. They nonetheless establish that the pricing gaps documented above would free roughly 100,000 acre-feet of water each year from the region's most stressed supply source while generating tens of millions of dollars in annual welfare. Who bears the cost of this correction, and how it reaches household bills, depends on the supply regime.

7.7 Who Bears the Correction?

Table 8 decomposes the welfare change by sector, measuring the status quo inclusive of the fixed service charge that utilities levy on top of the volumetric rate and crediting its removal to customers when they move to an all-volumetric optimum. The incidence is sharply state-contingent. In abundant years, moving to optimal pricing raises consumer surplus for both sectors–by \$21.1 per connection-year for residential customers and \$24.1 for non-residential–while utility net revenue falls by \$28.5. The household gain is structural rather than a scarcity effect: Arizona utilities over-recover their fixed costs through the service charge, and the efficient volumetric price lets them retire part of that regressive levy, returning surplus to customers. Under sustained shortage the pattern reverses. The scarcity-inclusive price must rise well above the observed rates. Consumer surplus falls by \$33.0 per connection-year for residential customers and \$18.6 for non-residential,

[18] The 52 CAP municipal and industrial subcontractors hold entitlements totaling approximately 620,000 AF per year (Central Arizona Project, 2024).

and the $82.2 increase in net revenue becomes a genuine transfer from customers to the utility.

Table 8: Sector Decomposition of the Welfare Change from Optimal Pricing (fixed-fee-inclusive status quo, AMP marginal price)

	All CAP		IBR		Uniform	
	Mild	Stressed	Mild	Stressed	Mild	Stressed
ΔCS, residential	$20.9	$-33.2	$24.4	$-40.6	$1.8	$6.9
ΔCS, non-residential	$24.1	$-19.2	$31.8	$-18.1	$-17.8	$-25.0
ΔNR (utility)	$-28.4	$83.0	$-42.1	$80.6	$46.0	$96.1
ΔW (total)	$16.6	$30.7	$14.1	$21.9	$30.0	$78.1

Notes: Connection-weighted mean welfare change per connection-year ($/conn/yr) across the 21 CAP utilities at the 100-year horizon, reported as a per-year annuity. The status quo charges observed AMP marginal prices plus the observed fixed fee; the optimum recovers fixed costs through the volumetric price. The fixed fee is split across sectors by each utility's residential connection share, so ΔCS nets forgone fixed charges against the higher volumetric price; $\Delta W = \Delta CS + \Delta NR$ and total ΔCS are invariant to the split. IBR $n = 15$, uniform $n = 6$.

Table 9 translates these changes into monthly bills at representative consumption levels, replacing each utility's observed schedule with the optimal uniform price and dropping the fixed charge as the welfare measure assumes. Connection-weighted across the 21 utilities, the bill changes are modest and concentrated at moderate residential use. At 10 kgal per month household bills rise by 12 to 16 percent, roughly $55 to $75 per year, while at low consumption (5 kgal) and high consumption (15 kgal) the change is negligible, within a couple of dollars per month. Non-residential bills increase within five percent at every consumption level under the stressed regime, but see small decreases under normal conditions. The equity concern is therefore smaller than the price gaps alone suggest: once the regressive fixed charge is retired, efficient pricing raises bills only modestly for moderate-use households and leaves them essentially unchanged elsewhere. The bill changes exhibit stark contrasts by rate structure and the detailed bill changes are displayed in Appendix Table A14. Uniform-rate utilities account for the larger non-residential increases while their residential bills fall, whereas the increasing residential bills are concentrated among low-to-moderate consumption blocks under IBRs.

Although the aggregate welfare gains are large, even the modest bill increases on moderate-use households may face political and public opposition. This equity concern does not require abandoning efficient pricing. A separate literature argues that the volumetric rate should carry the scarcity signal while distributional goals are pursued through a distinct instrument – a fixed rebate or lifeline credit to low-use or low-income households – rather than through the infra-marginal volumetric subsidies embedded in current rates (Boland and Whittington, 2000; Borenstein, 2012;

Table 9: Monthly Bill Change from Moving to Optimal Pricing

	Mild		Stressed	
Consumption level	$/mo	%	$/mo	%
Residential				
5 kgal	$0.24	1.2%	$1.00	4.8%
10 kgal	$4.64	12.4%	$6.16	16.4%
15 kgal	$-1.22	-1.9%	$1.06	1.6%
Non-residential				
25 kgal	$-1.08	-1.1%	$4.21	4.1%
50 kgal	$-0.95	-0.5%	$9.63	4.7%
100 kgal	$-3.48	-0.8%	$17.69	4.3%

Notes: Change in the monthly bill at the stated consumption (q kgal) when the observed 2019 schedule is replaced by the utility's optimal all-volumetric price, pooled across all 21 CAP utilities. $/mo is the connection-weighted mean; % is the connection-weighted aggregate change. A positive value is a bill increase.

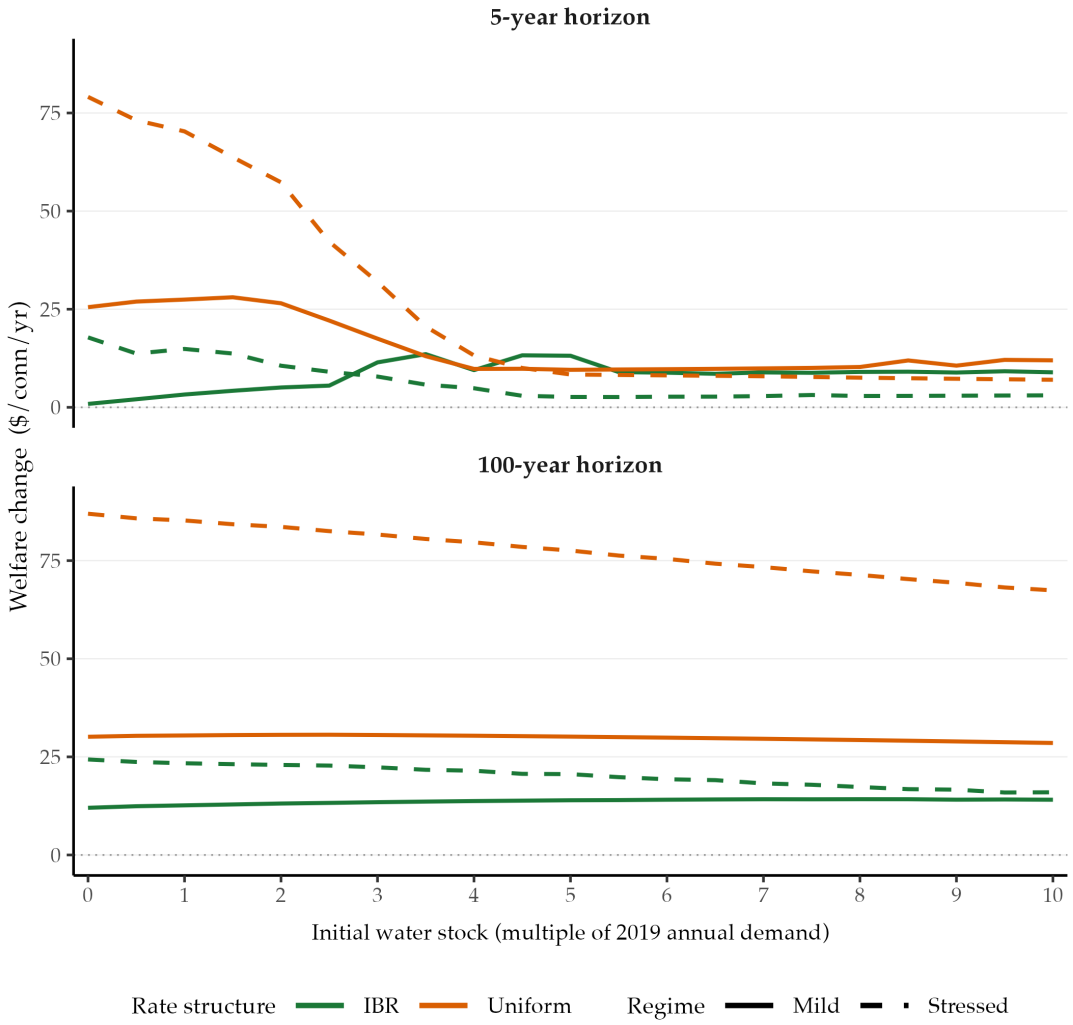
Wichman, 2024). The decomposition above shows that pairing the scarcity-inclusive marginal price with removal of the fixed charge leads to increased residential bills for IBR utilities in our sample and lower bills for uniform rate utilities.

7.8 Robustness

To trace the sensitivity of the welfare gain across a range of possible initial conditions, we vary the initial water stock from zero to ten times 2019 demand while holding initial financial reserves at zero (Figure 7). The two horizons behave differently. Over five years (top panel) the gains are small and similar across rate structures and regimes, on the order of $9 to $10 per connection-year once the utility holds any meaningful buffer, so the initial stock does little to separate the four cases. Over one hundred years (bottom panel) the ranking sharpens. Under mild conditions the gain is essentially insensitive to the initial stock, with uniform-rate utilities staying above increasing-block utilities throughout. Under stress a larger initial buffer pulls the gain down, as stored water substitutes for the scarcity price, but the gain remains much larger for uniform-rate utilities than for increasing-block utilities at every stock level. The uniform–IBR ordering is the same as in the headline results, reinforcing that the welfare gap by rate structure is a feature of the long run rather than of any particular initial condition.

We also assess four alternative specifications of the model, collected in Appendix Table A11;

Figure 7: Welfare Gain Sensitivity to Initial State Allocation



Notes: Each line is the connection-weighted mean flow welfare change $\Delta W = \Delta CS + \Delta NR$ ($/conn/yr) from optimal pricing, against the initial water stock (a multiple of 2019 annual demand; financial reserve fixed at zero). The status-quo benchmark prices at observed tariffs but purchases CAP optimally, as in Table 6. Top panel: 5-year horizon; bottom: 100-year.

across all four the headline conclusions hold, with positive welfare gains and substantial demand and CAP-purchase reductions throughout. Eliminating supply uncertainty leaves the long-run per-connection gain essentially unchanged, though the quantity reductions shrink somewhat when the optimal policy has no tail-shortfall risk to hedge. This difference is the precautionary wedge appearing in the solved policies. With supply risk removed the optimal price falls, demand and CAP purchases rise, the stored buffer is drawn down, and the residual shortages the stochastic policy guards against disappear. Removing the revenue reserve lowers the welfare gains to \$10.8 and \$27.6 per connection-year. The reserve matters most under mild conditions, where it funds the intertemporal smoothing of CAP purchases. We display results from applying the single residential elasticity estimate to both sectors in Appendix Table A12. Welfare gains drop substantially to \$15.7 and \$22.5. This reduction is concentrated in the stressed regime where the more inelastic non-residential demand of the headline generates larger scarcity-pricing gains, and it shifts the sectoral incidence, confirming that the two-sector structure carries real information for who bears the correction. The bill changes move in the same direction (Appendix Table A13): applying the residential elasticity to both sectors treats non-residential demand as less elastic than it is, so that sector absorbs more of the correction. Non-residential bills rise even in the mild regime, where the two-sector optimum left them roughly flat, while the residential increase at 10 kgal attenuates from 12–16 percent to 11–13 percent. Substituting the more elastic average price estimates (-0.67 residential and -0.70 non-residential) instead increases welfare gains to \$18.8 and \$41.9 from switching to optimal pricing.

We also propose an alternative group-weighted objective that places each utility's own revealed weights on the two sectors similar to Timmins (2002b). The aggregate welfare and conservation effects are slightly attenuated and the CAP-purchase reduction stays within a few thousand acre-feet of the headline. We provide complete results from this alternative objective function exercise in Appendix Table A15. The weighting alters the incidence of the move to efficient pricing. We recover weights favoring households for most utilities in our sample. For these utilities, the group-weighted optimum decreases the residential prices and increases the non-residential, shifting the surplus loss off residential and onto non-residential customers, so that residential bills fall or barely move while non-residential bills bear the correction. Appendix Table A16 displays the welfare decomposition. The equal-weight headline is therefore the conservative benchmark for residential incidence, and the conservation result is invariant to the welfare weights.

8 Conclusion

This paper asks whether water utility pricing embeds the scarcity value of water and where the pricing gaps fall on the consumption schedule. We address both questions with four complementary components. First, we assemble a uniquely detailed utility-level panel that directly observes quantities by source and customer class, linking billing schedules, financial records, and CAP delivery data for 194 Arizona utilities over four years. Second, we estimate residential and non-residential price elasticities of -0.53 and -0.87, using variation in neighboring utilities' prices as an instrument for average marginal price. Estimating sectoral demand elasticities is critical because non-residential use is a substantial fraction of total demand and the correction toward efficient pricing falls largely on households. Third, these estimates enter a two-sector stochastic dynamic program calibrated to each utility-year, yielding a scarcity-inclusive optimal price for each of the 21 CAP utilities in the sample. Finally, we use the optimal policy to simulate short-term and long-term water scarcity scenarios and recover changes in quantity, external water purchases, and welfare from moving to optimal pricing.

The results reveal systematic and substantial underpricing. CAP utilities' residential rates sit roughly 20 to 30 percent below the model-derived optimum while their non-residential rates are close to or above it. Non-CAP utilities price further below, in some instances needing to double prevailing prices to reach the optimum. Median price gaps are roughly equal between CAP utilities of differing rate structures; however, this obscures where in the consumption distribution these differences accumulate. Rate structure is associated with how this shortfall is distributed across consumption tiers. IBR utilities price moderate and heavy users substantially closer to the scarcity-inclusive optimum than their uniform-rate counterparts. In relative terms, the residential shortfall under IBR falls from roughly 40 percent at the lowest usage tier to below zero for the heaviest users, who are priced above the optimum, whereas under uniform rates it stays in the 5 to 24 percent range across levels. The welfare gains from moving to optimal pricing are large, and their headline payoff is a conservation gain delivered through prices alone. The distributional incidence is state-contingent: in normal years the efficient volumetric price lets utilities retire fixed charges and raises consumer surplus for both sectors, while under sustained shortage the scarcity-inclusive price transfers surplus from customers, mostly residential, to utility net revenue. Scaled to the roughly 1.26 million connections represented in our sample, moving to optimal prices over a hundred-year horizon would raise welfare by approximately $21 million per year under normal supply conditions and $38 million under severe shortage. Efficient pricing also cuts Central Ari-

zona Project water purchases by 14 to 17 percent of the municipal and industrial entitlements. The CAP purchase reductions represent 11 to 14 percent of the mandated 760,000 AF of Colorado River water Arizona is required to cut in 2027 and 2028.

The pricing and welfare results illustrate the efficiency, conservation, and equity tradeoff at the heart of water rate design. The largest pricing shortfalls sit on residential volumetric rates, but once the over-recovering fixed charge is retired for an all-volumetric optimum the bill consequences are modest and concentrated at middling use: residential bills at ten thousand gallons per month rise by 12 to 16 percent, while low-use residential and non-residential bills change little, with many non-residential customers seeing bills fall. Decomposing the bill changes by rate structure illustrates how the IBR utilities price residential customers well below optimal and non-residential customers very close to efficient rates. Uniform rate utilities moving to optimal prices would decrease bills for low-to-moderate use residential customers substantially while bills would increase for high use non-residential.

The analysis proceeds at the utility level, and three limitations follow from this. The welfare figures measure the aggregate gain from transitioning from observed marginal prices to optimal uniform pricing. The results do not account for how those gains and losses would be distributed across consumers within a utility. Under a transition from an IBR schedule to optimal uniform pricing, low-use residential customers currently paying below the scarcity-inclusive first-block price would incur a consumer surplus loss, while high-use customers on elevated upper tiers would gain. The burden to residential customers can be offset by pairing the efficient volumetric price with a fixed rebate or lifeline credit rather than inframarginal volumetric subsidies. The extent of this and the precise conservation versus equity effects depend on the distribution of customers across consumption tiers, information not available in the data. The model-optimal price is also a uniform-equivalent rather than an optimally designed IBR schedule; the second-best block structure that jointly satisfies efficiency, cost recovery, and distributional objectives is a natural extension left for future work. Finally, we proxy each utility's marginal price with the average price at 10 kgal of monthly use rather than the full block schedule, so the tier-level gaps should be read as representative rather than exact.

The findings carry a direct message for water regulators and utility managers. The welfare and conservation stakes are large; in the short run they do not appear to hinge on the choice of rate structure, while over the long run uniform-rate utilities show larger gains, though with only a handful of such utilities we cannot rule out that the two are the same. Prevailing rates largely

omit the scarcity rent. This gap between observed rates and the scarcity-inclusive optimum is concentrated in residential volumetric rates. Correcting it raises bills modestly for moderate-use households while leaving low-use and non-residential bills little changed, and where the residual increases raise equity concerns a fixed rebate or lifeline credit is a well-documented remedy. As Colorado River shortage declarations deepen and CAP delivery uncertainty grows, efficient pricing substantially reduces demand and positions utilities to be financially better equipped for future shocks.

References

Abbott, Joshua K., Eli P. Fenichel, and Seong D. Yun, "Risky (Natural) Assets: Stochasticity, Nonconvexity, and the Value of Natural Capital," *Journal of the Association of Environmental and Resource Economists*, 2026, *13* (5), 1269–1309.

Allaire, Maura and Ariel Dinar, "What Drives Water Utility Selection of Pricing Methods? Evidence from California," *Water Resources Management*, 2022, *36*, 1–17.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press, 2009.

Arizona Water Banking Authority, "Arizona Water Banking Authority Annual Report," Arizona Water Banking Authority, Phoenix, AZ 2024.

Baerenklau, Kenneth A., Kurt A. Schwabe, and Ariel Dinar, "The Residential Water Demand Effect of Increasing Block Rate Water Budgets," *Land Economics*, 2014, *90* (4), 683–699.

Berry, Steven T. and Philip A. Haile, "Foundations of Demand Estimation," in "Handbook of Industrial Organization," Vol. 4, Elsevier, 2021, pp. 1–62.

Boland, John J. and Dale Whittington, "The Political Economy of Water Tariff Design in Developing Countries: Increasing Block Tariffs versus Uniform Price with Rebate," in Ariel Dinar, ed., *The Political Economy of Water Pricing Reforms*, New York: Oxford University Press, 2000.

Borenstein, Severin, "The Redistributional Impact of Nonlinear Electricity Pricing," *American Economic Journal: Economic Policy*, 2012, *4* (3), 56–90.

_ , "The Economics of Fixed Cost Recovery by Utilities," *The Electricity Journal*, 2016, *29* (7), 5–12.

_ **and Lucas W Davis**, "The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets," *The Journal of Law and Economics*, 2012, *55* (1), 75–128.

Burlig, Fiona, Louis Preonas, and Matt Woerman, "Groundwater and Crop Choice in the Short and Long Run," Working Paper 28706, National Bureau of Economic Research April 2021.

Central Arizona Project, "Colorado River Operations," Central Arizona Project 2024.

_ , "Economic Impact to CAP," Technical Report, Central Arizona Water Conservation District 2026.

Chu, Long and R. Quentin Grafton, "Dynamic Water Pricing and the Risk Adjusted User Cost (RAUC)," *Water Resources and Economics*, 2021, *35*, 100181.

EPA, U.S., "How We Use Water," Sep 2024.

Fenichel, Eli P. and Joshua K. Abbott, "Natural Capital: From Metaphor to Measurement," *Journal of the Association of Environmental and Resource Economists*, 2014, *1* (1/2), 1–27.

Griffin, Ronald C., "Effective Water Pricing," *JAWRA Journal of the American Water Resources Association*, 2001, *37* (5), 1335–1347.

Hanemann, W. M., "Price and Rate Structures," in D. D. Baumann, J. J. Boland, and W. M. Hanemann, eds., *Urban Water Demand Management and Planning*, New York: McGraw-Hill, 1998, chapter 5, pp. 95–134.

Hanemann, W. Michael, "The Central Arizona Project," CUDARE Working Paper 937, University of California, Berkeley 2002.

Hausman, Jerry A., Gregory K. Leonard, and J. Douglas Zona, "Competitive Analysis with Differentiated Products," *Annales d'Économie et de Statistique*, 1994, *34*, 159–180.

Ito, Koichiro, "How Do Consumers Respond to Nonlinear Pricing? Evidence from a Natural Experiment in the US Electricity Market," *American Economic Review*, 2014, *104* (2), 537–563.

Levinson, Arik and Emilson Silva, "The Electric Gini: Income Redistribution through Energy Prices," *American Economic Journal: Economic Policy*, 2022, *14* (2), 341–65.

MacKay, Alexander and Nathan H Miller, "Estimating Models of Supply and Demand: Instruments and Covariance Restrictions," *American Economic Journal: Microeconomics*, 2025, *17* (1), 238–281.

Mansur, Erin T and Sheila M Olmstead, "The Value of Scarce Water: Measuring the Inefficiency of Municipal Regulations," *Journal of Urban Economics*, 2012, *71* (3), 332–346.

McRae, Shaun, "Infrastructure Quality and the Subsidy Trap," *American Economic Review*, January 2015, *105* (1), 35–66.

Moncur, James E. T. and Richard L. Pollock, "Scarcity Rents for Water: A Valuation and Pricing Model," *Land Economics*, 1988, *64* (1), 62–72.

Nevo, Aviv, "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 2001, *69* (2), 307–342.

Olmstead, Sheila M., "The Economics of Managing Scarce Water Resources," *Review of Environmental Economics and Policy*, 2010, *4* (2), 179–198.

_ **, W. Michael Hanemann, and Robert N. Stavins**, "Water Demand under Alternative Price Structures," *Journal of Environmental Economics and Management*, 2007, *54* (2), 181–198.

Renzetti, Steven, "Estimating the Structure of Industrial Water Demands: The Case of Canadian Manufacturing," *Land Economics*, 1992, *68* (4), 396–404.

_ , "Examining the Differences in Self- and Publicly Supplied Firms' Water Demands," *Land Economics*, 1993, *69* (2), 181–188.

_ , "Municipal Water Supply and Sewage Treatment: Costs, Prices, and Distortions," *Canadian Journal of Economics*, 1999, pp. 688–704.

Reynaud, Arnaud, "An Econometric Estimation of Industrial Water Demand in France," *Environmental and Resource Economics*, 2003, *25* (2), 213–232.

Sağlam, Yiğit, "Welfare Implications of Water Scarcity: Higher Prices of Desalination," *Environmental and Resource Economics*, 2019, *73* (4), 995–1022.

Sibly, Hugh and Richard Tooth, "The Consequences of Using Increasing Block Tariffs to Price Urban Water," *Australian Journal of Agricultural and Resource Economics*, 2014, *58* (2), 223–243.

Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," in Donald W. K. Andrews and James H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, New York: Cambridge University Press, 2005, pp. 80–108.

Szabó, Andrea, "The Value of Free Water: Analyzing South Africa's Free Basic Water Policy," *Econometrica*, 2015, *83* (5), 1913–1961.

Timmins, Christopher, "Does the Median Voter Consume Too Much Water? Analyzing the Redistributive Role of Residential Water Bills," *National Tax Journal*, 2002, *55* (4), 687–702.

_ , "Measuring the Dynamic Efficiency Costs of Regulators' Preferences: Municipal Water Utilities in the Arid West," *Econometrica*, 2002, *70* (2), 603–629.

Tsur, Yacov, "Optimal Water Pricing: Accounting for Environmental Externalities," *Ecological Economics*, 2020, *170*, 106429.

U.S. Bureau of Reclamation, "Record of Decision: Colorado River Interim Guidelines for Lower Basin Shortages and Coordinated Operations for Lake Powell and Lake Mead," Technical Report, U.S. Department of the Interior, Bureau of Reclamation 2007.

_ , "Interior Department Finalizes Plans for 2027–2028 Colorado River Operations," News release, U.S. Department of the Interior, Bureau of Reclamation August 2026.

Wang, Zeyu and Frank A. Wolak, "Designing Nonlinear Price Schedules for Urban Water Utilities to Balance Revenue and Conservation Goals," *Working Paper*, 2022.

Western Resource Advocates, "Arizona's Water Future: Strengthening Water Security in a Drier Future," Technical Report, Western Resource Advocates 2023.

Wichman, Casey J., "Perceived Price in Residential Water Demand: Evidence from a Natural Experiment," *Water Resources Research*, 2014, *50* (12), 9470–9479.

Wichman, Casey J, "Efficiency, Equity, and Cost-Recovery Trade-Offs in Municipal Water Pricing," Technical Report, Resources for the Future 2024.

Wietelman, Derek C., Casey J. Wichman, and Daniel A. Brent, "Conservation and Distributional Consequences of Pricing Scarce Water During Droughts," *Working Paper*, November 2025.

Williams, A. Park, Benjamin I. Cook, and Jason E. Smerdon, "Rapid Intensification of the Emerging Southwestern North American Megadrought in 2020–2021," *Nature Climate Change*, 2022, *12* (3), 232–234.

Appendix

A Model Solution and Calibration

A.1 Demand Functional Form and Consumer Surplus

The manager's problem requires a demand curve $q_i(p_i)$ and the associated consumer surplus $CS_i(p_i)$ for each sector. We adopt the constant-elasticity (log-log) form implied by the demand regressions in Section 5.1, in which log quantity is linear in log price. The fixed fee and household income enter those regressions as shifters but are not choice variables in the manager's problem, so we hold them at their sample means and sectoral demand collapses to

$$q_i(p_i) = A_i \, p_i^{\varepsilon_i}, \qquad i \in \{1, 2\}, \tag{26}$$

where $\varepsilon_i < 0$ is the estimated price elasticity and $A_i > 0$ is a scale factor absorbing the regression intercept together with the mean fixed fee, income, and service population. The estimated elasticities satisfy $\varepsilon_i > -1$ ($\hat{\varepsilon}_1 = -0.53$, $\hat{\varepsilon}_2 = -0.87$), so demand is inelastic and the surplus integral below converges.

We calibrate the scale factor so that demand passes through the observed price–quantity pair. For each utility-year we set

$$A_i = \bar{q}_i \, \bar{p}_i^{-\varepsilon_i}, \tag{27}$$

where $\bar{p}_i$ is the observed sector price and $\bar{q}_i$ is observed sector deliveries. This anchoring guarantees that at observed prices the model reproduces observed quantities; the elasticity governs only the curvature of demand away from the calibration point, and A_i is therefore utility-year specific.

Consumer surplus is the area under inverse demand from p_i up to a choke price $\bar{p}_i^{\max}$:

$$CS_i(p_i) = \int_{p_i}^{\bar{p}_i^{\max}} q_i(s) \, ds = \frac{A_i}{\varepsilon_i + 1} \left[\left(\bar{p}_i^{\max} \right)^{\varepsilon_i + 1} - p_i^{\varepsilon_i + 1} \right]. \tag{28}$$

We set the choke price well above any observed rate ($\bar{p}_i^{\max} = \$30$/kgal). Because it enters CS_i only additively, its level shifts the value function uniformly and does not affect optimal prices or the scarcity rent. Differentiating equation (28) recovers Roy's identity, $\partial CS_i / \partial p_i = -q_i(p_i)$, the property used to derive the optimal-price condition in equation (17). The constant-elasticity form

also gives the inverse semielasticity term in the Ramsey markup a closed form,

$$\frac{-q_i(p_i)}{\partial q_i/\partial p_i} = \frac{p_i}{|\varepsilon_i|}, \tag{29}$$

so the markup in equation (18) is $\frac{\bar{\mu}-1}{\bar{\mu}} \cdot \frac{p_i}{|\varepsilon_i|}$: the less elastic sector (residential) carries the larger markup.

A.2 Numerical Solution and Calibration

We solve the manager's problem in equations (1)–(6) by value function iteration, once for each CAP utility-year. We approximate the value function $V(w, b)$ with a two-dimensional Chebyshev polynomial of degree six on a grid of $N = 16$ collocation nodes in each state dimension, the water stock w and the reserve balance b. The stock grid runs from zero to an upper bound set to at least 2.5 times observed storage (and to multiples of the utility's entitlements and contract capacity), so that utilities holding little or no carryover storage still have a well-posed grid; the reserve grid is scaled to the utility's own fixed-cost magnitude. We discretize the post-price supply shock θ into $N_\theta = 7$ nodes placed at the inverse Beta cumulative distribution function of evenly spaced probabilities, each carrying equal weight.

At every collocation node we solve the stage problem –choosing prices (p_1, p_2), purchases $x(\theta)$, and next-period states $(w'(\theta), b'(\theta))$ subject to the resource, budget, and purchase-limit constraints— with a sequential quadratic programming routine, providing analytic gradients of the objective and constraints. The stage objective is not globally concave in price because the continuation value can generate several local optima, so a single-start solution is unreliable. We use a two-step procedure, first sweeping a set of pinned candidate prices, the cost-recovery price and fixed multiples of it, optimizing only over quantities at each candidate, and then polish the best pinned solution with a free-price solve. Because the cost recovery price is always among the candidates, the Bellman operator never returns a policy worse than cost recovery, which eliminates the too-high-price local optima that trap a naive solver.

We iterate the Bellman operator to convergence of the price policy rather than the value-function norm. The value-function change is dominated by the large constant welfare level $CS/(1-\beta)$, so the relative Chebyshev-coefficient norm clears a tight tolerance after a single Bellman step while prices are still moving and the continuation-value curvature—and hence the scarcity rent $\bar{\lambda}/\bar{\mu}$— has not yet developed. We therefore require a minimum of twelve iterations and stop when the

relative change in the price policy falls below 5×10^{-3}, capping the loop at sixty iterations. Forty-two of the forty-six CAP utility-years satisfy this strict policy-convergence criterion; the remainder settle into a shallow limit cycle in which the policy plateaus near 10^{-2}, and for these we retain the last iterate, which remains a valid policy.

We calibrate the model at the utility-year level. The annual discount factor is $\beta = 0.97$ and the gross return on reserves is $R = 1.02$. We solve and simulate the model under two supply-shock regimes, each policy is solved under that regime's own shock distribution. The mild regime takes $\theta \sim \text{Beta}(50, 1)$ (mean 0.98, standard deviation 0.02), reflecting the low-shortage conditions of the 2014–2019 sample, while the stressed regime takes $\theta \sim \text{Beta}(4, 1)$ (mean 0.80, standard deviation 0.16), approximating a Tier-3 CAP shortage. Demand uses the estimated elasticities $\hat{\varepsilon}_1 = -0.53$ and $\hat{\varepsilon}_2 = -0.87$ with the scale factors of equation (27). Marginal cost is $c_i = \hat{\beta}\, AC_{it}$ from the cost regression in Section 6, and the fixed-cost term is $\tau = (1 - \hat{\beta})\, E_{it}$, where E_{it} is operating expenditure and AC_{it} average cost. The purchase price p_x is the utility's CAP municipal-and-industrial delivery rate; fixed entitlements w_f are groundwater and surface water rights; the contract cap $\bar{x}$ is the M&I entitlement; and the observed stock w is the long-term-storage-account balance. For non-CAP utilities, we do not model storage scarcity and take the optimal-price benchmark to be the two-sector Ramsey cost-recovery price that recovers AC_{it} through inverse-elasticity markups over c_i; only CAP utilities are solved by the full stochastic dynamic program above. Table A1 collects every calibrated and estimated parameter, its value, and its source.

Table A1: Simulation Parameters, Values, and Sources

Parameter	Symbol	Value	Source
Preferences and finance			
Annual discount factor	β	0.97	Assumed (standard annual rate)
Return on financial reserves	R	1.02	Assumed
Demand			
Residential price elasticity	ε_1	-0.53	Estimated, Table 2
Non-residential price elasticity	ε_2	-0.87	Estimated, Table 3
Residential demand scale	A_1	Utility-specific	Calibrated to observed $(\bar{p}_1, \bar{q}_1)$
Non-residential demand scale	A_2	Utility-specific	Calibrated to observed $(\bar{p}_2, \bar{q}_2)$; q_2 per establishment (ZBP)
Max water price (CS integration)	$\bar{p}^{\max}$	\$30/kgal	Assumed upper bound
Water connections (per-conn scaling)	n_i	Utility-specific	Data: EPA SDWIS; service population $\div 2.5$ where missing
Cost			
Average cost of service	AC_{it}	Utility-year	Data: $E_{it}/Q_{it}^{\text{sold}}$; E from ACC income statements, ADWR PRRs, or UNC EFC rate surveys (by year/utility)
Cost (output) elasticity	$\hat{\beta}$	AMA 0.86; non-AMA 0.54	Estimated, Table 4 (CAP interaction insignificant, set to 0; following Wichman, 2024)
Marginal cost	c_i	$\hat{\beta} \cdot AC_{it}$	Estimated, Table 4
Fixed cost	τ_i	$(1-\hat{\beta}) \cdot E_{it}$	Estimated, Table 4
CAP supply and shock			
CAP entitlement (purchase cap)	$\bar{x}$	Utility-specific (kgal)	Data: CAP (CAWCD) public records request
CAP price, mild regime	$p_x(t)$	0.485×1.025^t	Observed CAP M&I rates (CAP PRR), grown 2.5%/yr (Fig. A6)
CAP price, stressed regime	$p_x(\theta_s)$	$0.90 + (1-\theta_s)$	Observed post-2022 DCP shortage rates (CAP PRR)
Shock, mild regime (VFI & simulation)	θ_m	Beta$(50,1)$, mean 0.98	Calibrated (low-shortage era)
Shock, stressed regime (VFI & simulation)	θ_s	Beta$(4,1)$, mean 0.80	Calibrated ($\sim$Tier-3 shortage)
Initial state (headline: 2019 observed)			
Ground and Surface Water	w_f	Utility-specific (kgal)	Data: ADWR public records request (PRR)
Initial water stock	w_0	2019 LTSA balance (kgal)	Data: ADWR LTSA (public records request)
Initial financial reserve	b_0	$\max(0, \text{rev} - \text{exp})_{2019}$	Data: ACC income statements, ADWR PRRs, or UNC EFC (by year/utility)
Numerical			
Chebyshev nodes per state	N	16	—
Chebyshev polynomial degree	—	6	—
Shock quadrature nodes	—	7	Equal-probability Beta bins
Simulation horizon	T	5 and 100 years	—
Monte Carlo paths	—	1,000 per utility-regime	—

Notes: Utility-specific parameters vary across the 21 CAP utilities (Section 4); referenced tables report the estimated parameters. A separate optimal policy is solved for each regime under that regime's own shock distribution—$\theta_m \sim$ Beta$(50,1)$ (mild) and $\theta_s \sim$ Beta$(4,1)$ (stressed)—and each regime is simulated forward on its own policy. PRR = public records request; ZBP = ZIP Business Patterns; SDWIS = EPA Safe Drinking Water Information System.

B Additional Tables

B.1 First-Stage Regressions

Table A2: First-Stage Regressions: Residential Demand IV

	$\ln(AMP_{ict})$ (1)	ln(Fixed Fee) (2)
$ln(\overline{AMP}_{t-1})$	-4.26***	0.764*
	(0.938)	(0.415)
$ln(\overline{F}_{t-1})$	-0.749	-5.91***
	(0.618)	(0.720)
ln(Median Income)	0.030	-0.029
	(0.136)	(0.125)
Observations	526	526
R^2	0.45	0.43
County-Year fixed effects	✓	✓
Angrist–Pischke F	20.6	67.3

Notes: First stages for the two-endogenous residential demand IV (Table 2, Column 2); each endogenous regressor projected on both excluded instruments ($\overline{AMP}_{t-1}$, $\overline{F}_{t-1}$), with County×Year fixed effects and SEs clustered by utility. Significance: ***p<0.01, **p<0.05, *p<0.10.

Table A3: First-Stage Regressions: Non-Residential Demand IV

	$\ln(AMP^{com}_{ict})$ (1)	ln(Fixed Fee) (2)
$ln(\overline{AMP}_{t-1})$	-2.68***	0.204
	(0.501)	(0.338)
$ln(\overline{F}_{t-1})$	-0.393	-4.37***
	(0.610)	(0.766)
ln(Payroll per Estab.)	-0.098*	-0.128**
	(0.054)	(0.060)
ln(Median Income)	0.009	-0.103
	(0.128)	(0.153)
Observations	392	392
R^2	0.47	0.47
County-Year fixed effects	✓	✓
Angrist–Pischke F	28.6	32.5

Notes: First stages for the two-endogenous non-residential demand IV (Table 3, Column 2); each endogenous regressor projected on both excluded instruments ($\overline{AMP}_{t-1}$, $\overline{F}_{t-1}$), with County×Year fixed effects and SEs clustered by utility. Significance: ***p<0.01, **p<0.05, *p<0.10.

B.2 Demand Estimation Robustness

Table A4: Residential Water Demand: Thin-County Robustness (Counties with ≥ 5 Utilities)

	(1) AMP	(2) AVP	(3) AP
$ln(\widehat{AMP_{ict}})$	-0.58***		
	(0.22)		
$ln(\widehat{AVP_{ict}})$		-0.52**	
		(0.22)	
$ln(\widehat{AP_{ict}})$			-0.71***
			(0.22)
$ln(\widehat{\text{Fixed Fee}_{ict}})$	-0.02	-0.10	
	(0.18)	(0.17)	
ln(Median Income)	0.94***	0.95***	0.93***
	(0.15)	(0.15)	(0.16)
Observations	509	509	509
County×Year FE	✓	✓	✓
Cragg-Donald F	135.5	136.1	319.5
AP F: Price	45.7	56.6	54.6
AP F: Fixed Fee	65.1	55.4	

Notes: Arizona utilities, 2014–2019, restricted to county-years (period $t-1$) with $\geq$ 5 utilities reporting valid prices. Specifications mirror Table 2: AMP (marginal at 10 kgal) and AVP ($(\text{Bill}_{10}-\text{fee})/10$) are instrumented jointly with the fixed fee; AP ($\text{Bill}_{10}/10$, fee-inclusive) enters as a single endogenous regressor. Dependent variable is ln(residential water per capita, kgal/yr). Standard errors clustered by utility in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.10.

Table A5: Non-Residential Water Demand: Thin-County Robustness (Counties with ≥ 5 Utilities)

	(1) AMP (25–50)	(2) AMP (10–25)	(3) AVP	(4) AP
$ln(\widehat{AMP^{nr}_{ict}})$ (25–50 tier)	-0.90*			
	(0.51)			
$ln(\widehat{AMP^{nr}_{ict}})$ (10–25 tier)		-0.96**		
		(0.45)		
$ln(\widehat{AVP^{nr}_{ict}})$ (vol, 50 kgal)			-0.96*	
			(0.49)	
$ln(\widehat{AP^{nr}_{ict}})$ (fee-incl, 50 kgal)				-0.78
				(0.50)
$ln(\widehat{\text{Fixed Fee}^{nr}_{ict}})$	0.70*	0.44	0.61	
	(0.39)	(0.42)	(0.39)	
ln(Payroll/Establishment)	-0.24	-0.30	-0.27	-0.36
	(0.31)	(0.32)	(0.32)	(0.31)
ln(Median HH Income)	-2.03***	-2.17***	-2.08***	-2.13***
	(0.58)	(0.61)	(0.59)	(0.62)
Observations	364	364	364	364
County×Year FE	✓	✓	✓	✓
Cragg-Donald F	48.2	56.9	54.6	112.3
AP F: Price	16.7	13.8	15.0	15.9
AP F: Fixed Fee	63.9	61.3	62.4	

Notes: Arizona utilities, 2014–2019, restricted to county-years (period $t-1$) with ≥ 5 utilities reporting valid non-residential prices. Specifications mirror Table 3: AMP (25–50 and 10–25 kgal/month tiers) and AVP ($(\text{Bill}_{50}-\text{fee})/50$) are instrumented jointly with the fixed fee; AP ($\text{Bill}_{50}/50$, fee-inclusive) enters as a single endogenous regressor. Dependent variable is ln(non-residential water per establishment, kgal/yr). Standard errors clustered by utility in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.10.

B.3 Exclusion-Restriction Robustness to Drought Controls

Table A6: Residential Demand: Exclusion-Restriction Robustness to Drought Controls

	Base	+ Own precip.	+ Own & neighbor
	(1)	(2)	(3)
$\ln(\widehat{AMP_{ict}})$	-0.53***	-0.53***	-0.51**
	(0.19)	(0.20)	(0.21)
$\ln(\widehat{\text{Fixed Fee}})$	-0.05	-0.05	-0.07
	(0.16)	(0.16)	(0.16)
ln(Median Income)	0.92***	0.91***	0.91***
	(0.15)	(0.16)	(0.16)
ln(Own Precip.)		0.04	0.05
		(0.17)	(0.17)
ln(Neighbor Precip., lag)			0.30
			(0.40)
Observations	526	526	526
County×Year FE	✓	✓	✓
AP F: Price	20.6	18.8	23.1
AP F: Fixed Fee	67.3	69.8	68.7

Notes: Dependent variable is log residential water use per capita; Arizona utilities, 2014–2019. Two-endogenous IV (volumetric price and fixed fee instrumented by lagged leave-one-out county-peer values), County×Year FE, SEs clustered by utility. Columns (2)–(3) add own and lagged neighbor precipitation. AP F is the Angrist–Pischke first-stage F. Significance: ***p<0.01, **p<0.05, *p<0.10.

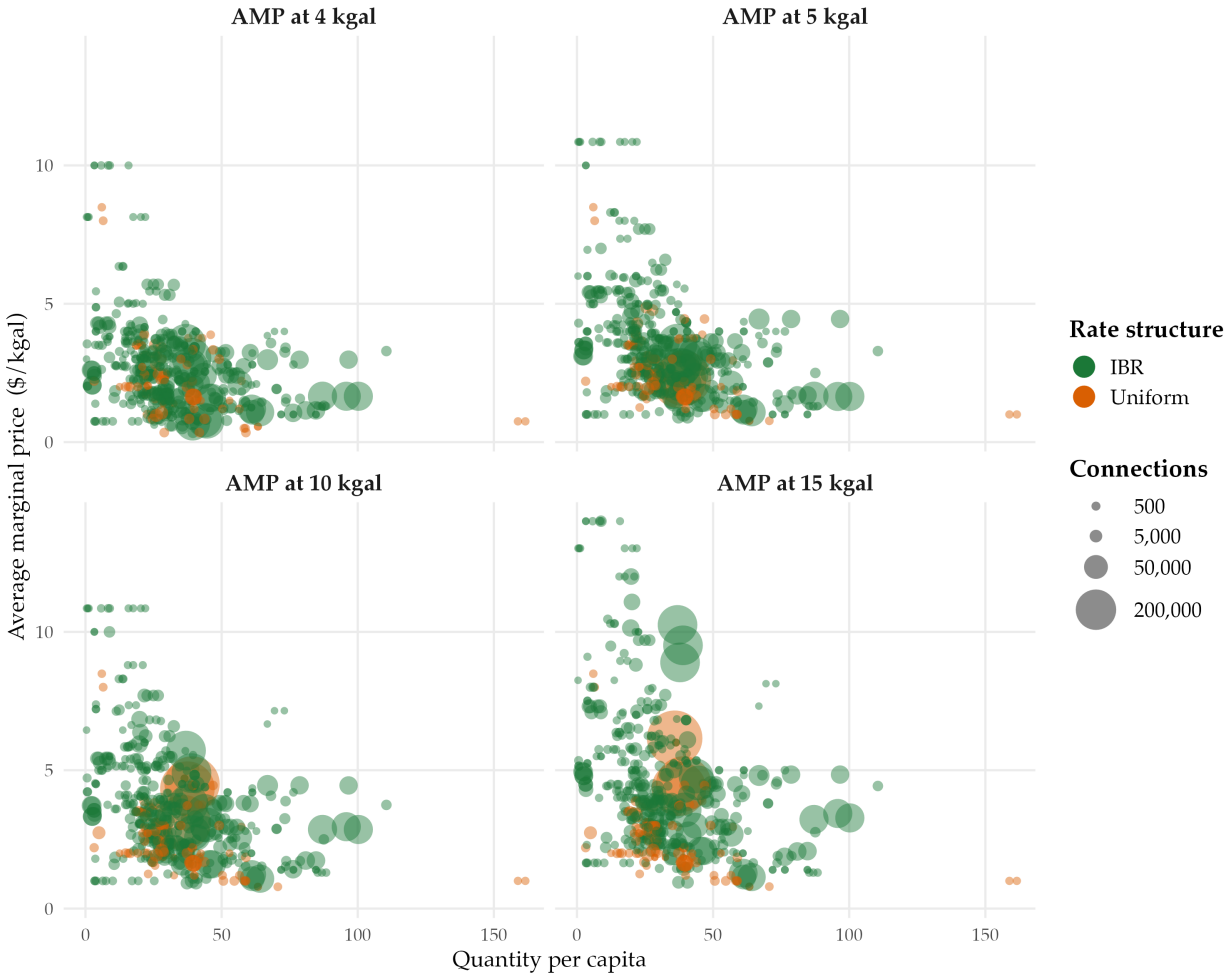
Table A7: Non-Residential Demand: Exclusion-Restriction Robustness to Drought Controls

	Base	+ Own precip.	+ Own & neighbor
	(1)	(2)	(3)
$\ln(\widehat{AMP_{ict}^{com}})$	-0.87**	-0.71**	-0.72*
	(0.36)	(0.36)	(0.38)
$\ln(\widehat{\text{Fixed Fee}})$	0.64*	0.60*	0.62*
	(0.37)	(0.36)	(0.37)
ln(Payroll per Estab.)	-0.26	-0.28	-0.28
	(0.28)	(0.28)	(0.28)
ln(Median Income)	-2.0***	-1.9***	-1.9***
	(0.57)	(0.59)	(0.59)
ln(Own Precip.)		-0.87	-0.88
		(0.65)	(0.67)
ln(Neighbor Precip., lag)			-0.19
			(1.2)
Observations	392	392	392
County×Year FE	✓	✓	✓
AP *F*: Price	28.6	26.2	26.6
AP *F*: Fixed Fee	32.5	33.4	34.6

Notes: Dependent variable is log non-residential water use per establishment; Arizona utilities, 2014–2019. Two-endogenous IV (volumetric price and fixed fee instrumented by lagged leave-one-out county-peer values), County×Year FE, SEs clustered by utility. Columns (2)–(3) add own and lagged neighbor precipitation. AP *F* is the Angrist–Pischke first-stage *F*. Significance: ***p<0.01, **p<0.05, *p<0.10.

B.4 Price and Quantity Variation

Figure A1: Residential Price and Quantity Variation by Consumption Tier



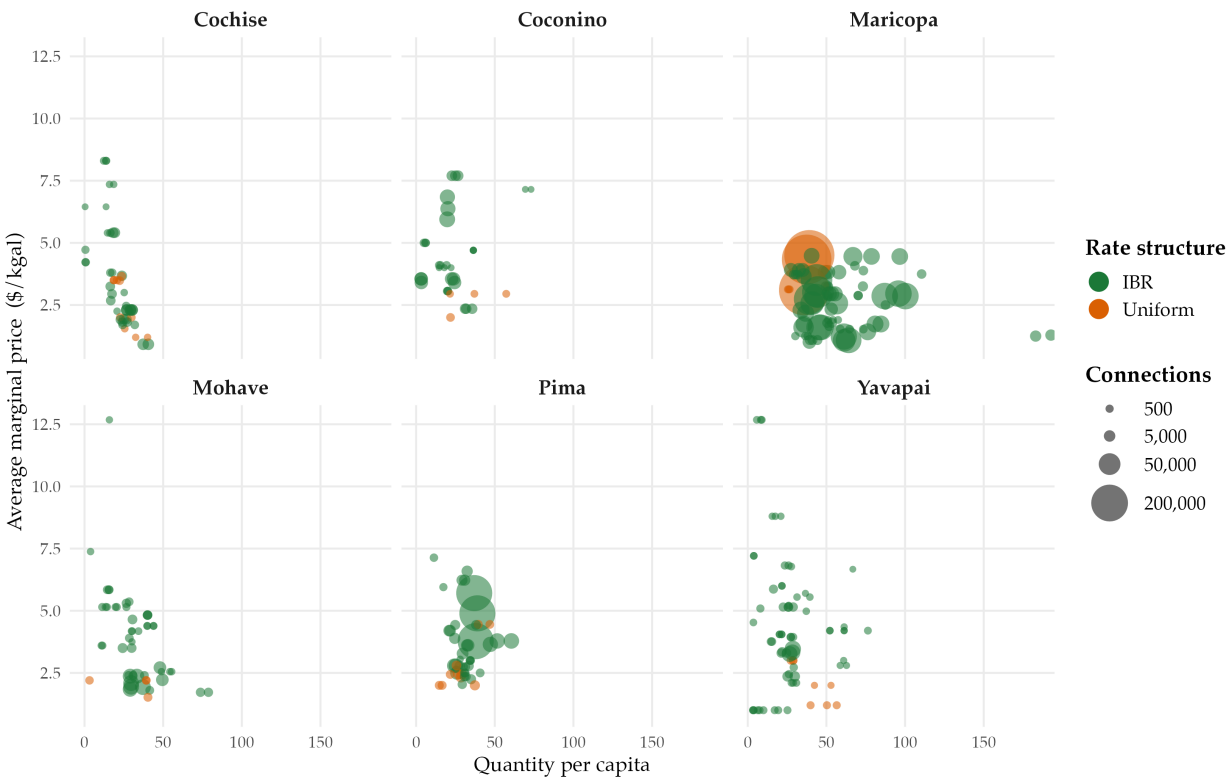
Notes: Each panel plots residential water quantity per capita (kgal per person per year) against the average marginal price at the indicated monthly consumption level (4, 5, 10, and 15 kgal), across the residential demand estimation sample (2015, 2017, and 2019). Point color denotes rate structure (green: IBR; purple: uniform) and point size is the number of water connections. Points are raw utility-year observations; no fitted line or fixed-effect residualization is applied. Prices are clipped at the 1st and 99th percentiles within each tier for readability.

Figure A2: Non-Residential Price and Quantity Variation by Consumption Tier



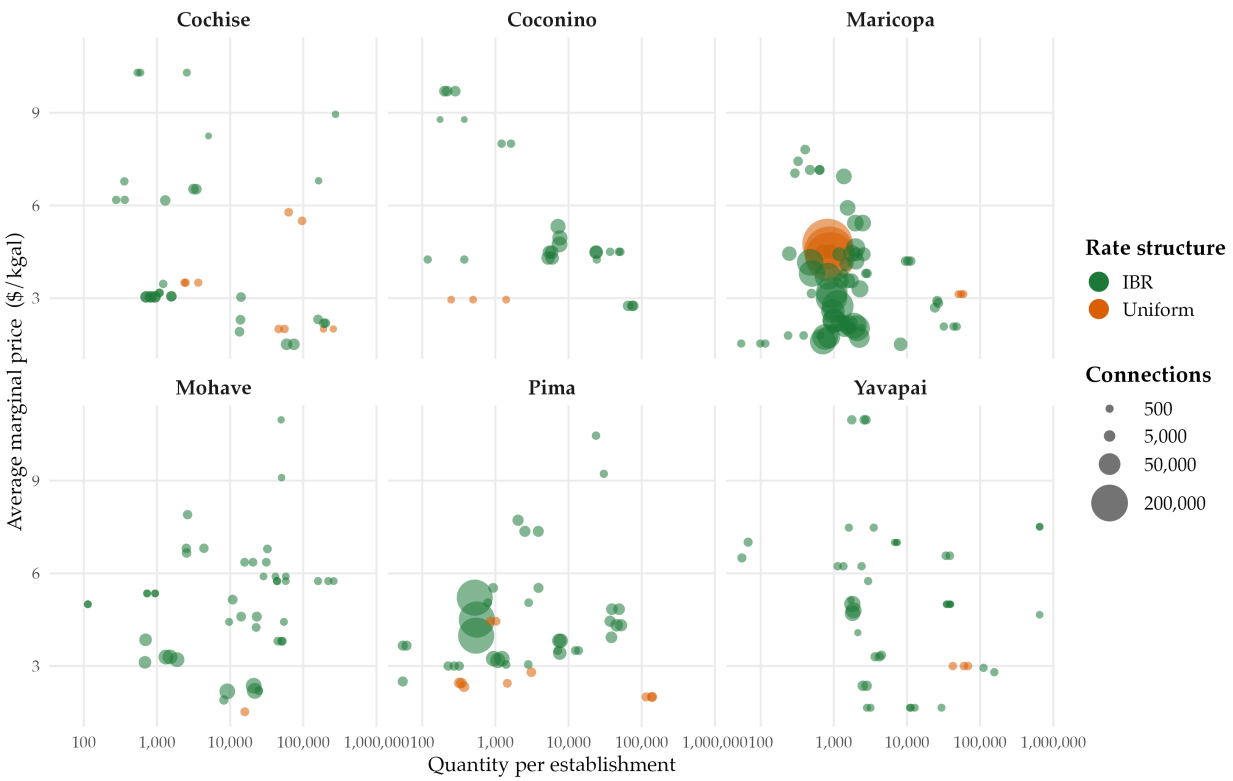
Notes: Each panel plots non-residential water quantity per establishment (kgal per establishment per year, log scale) against the average marginal price in the indicated consumption tier (10, 25, 50, and 100 kgal per month), across the commercial demand estimation sample (2015, 2017, and 2019). Point color denotes rate structure (green: IBR; purple: uniform) and point size is the number of water connections. Points are raw utility-year observations; no fitted line or fixed-effect residualization is applied. Prices and quantities are clipped at the 1st and 99th percentiles within each tier for readability.

Figure A3: Residential Price and Quantity Variation Within County, at the Estimation Consumption Level



Notes: Each panel plots residential water quantity per capita against the average marginal price at 10 kgal per month for the six counties contributing the most utility-years to the estimation sample. Point color denotes rate structure (green: IBR; purple: uniform) and point size is the number of water connections. Axes are common across panels; quantity is clipped at the 98th percentile for readability. Points are raw utility-year observations; no fitted line or fixed-effect residualization is applied. The within-county spread in price at a common quantity is the variation the leave-one-out county instrument exploits.

Figure A4: Non-Residential Price and Quantity Variation Within County, at the Estimation Consumption Level



Notes: Each panel plots non-residential water quantity per establishment (log scale) against the average marginal price in the 25–50 kgal per month tier for the same six counties. Point color denotes rate structure (green: IBR; purple: uniform) and point size is the number of water connections. Points are raw utility-year observations; no fitted line or fixed-effect residualization is applied. Axes are common across panels.

B.5 Price Gap to Optimal

Table A8: Marginal-Price Underpricing Relative to the Scarcity-Inclusive Optimum, by Consumption Tier (percent)

Consumption (kgal/mo)	Unif. CAP	Unif. non-CAP	IBR CAP	IBR non-CAP
Panel A: Residential				
4	24%	48%	43%	43%
5	23%	33%	41%	19%
10	9%	26%	28%	18%
15	5%	25%	-18%	-5%
25	18%	25%	-24%	-5%
Panel B: Non-residential				
10	-53%	-124%	-26%	-116%
25	-11%	-72%	-24%	-94%
50	1%	-59%	-33%	-85%
100	7%	-51%	-32%	-80%
250	8%	-49%	-29%	-81%

Notes: Each cell is the median across utilities (2019) of the marginal-price shortfall relative to the scarcity-inclusive optimum, $100 \times [\, p^* - (AC_{it} - MC_{it}) - \text{Avg. MP}_{it} \,] / [\, p^* - (AC_{it} - MC_{it}) \,]$, at the consumption level shown. A positive value means the price lies that many percent below the optimum. Sample: IBR and uniform-rate utilities with $AC_{it} \leq \$30$/kgal; years 2019.

B.6 Rate Structure Balance

Table A9: Characteristics of CAP Utilities by Rate Structure

	IBR ($n = 60$ util-yrs)	Uniform ($n = 24$ util-yrs)	p
Connections	22691 [4452, 87000]	15850 [11400, 54928]	0.66
Residential price ($/kgal)	2.99 [2.17, 3.82]	2.91 [1.28, 3.68]	0.11
Non-residential price ($/kgal)	3.92 [3.36, 5.41]	3.45 [2.67, 3.81]	0.01
Average cost ($/kgal)	4.40 [3.76, 5.56]	3.90 [3.39, 4.25]	0.07
Marginal cost ($/kgal)	4.11 [3.52, 5.19]	3.64 [3.16, 3.97]	0.07
Residential quantity (mil. kgal/yr)	2.9 [0.4, 17.9]	2.25 [1.36, 9.76]	0.73
Non-residential quantity (mil. kgal/yr)	1.38 [0.35, 7.19]	0.70 [0.36, 3.67]	0.27
CAP entitlement (mil. kgal/yr)	3.5 [0.7, 18.6]	2.21 [0.77, 7.36]	0.18
Non-CAP supply (mil. kgal/yr)	1.81 [0.00, 4.00]	1.76 [0.99, 2.85]	0.92
LTSA water stock (mil. kgal, 2019)	24 [2, 106]	6.5 [2.4, 44.8]	0.59
Revenue ($M)	26.9 [6.8, 52.5]	18.8 [11.8, 36.5]	0.72
Service-area median income ($K)	74.1 [59.7, 84.3]	72.6 [54.2, 89.4]	0.83
Rate changes in sample (count)	3.00 [1.50, 3.00]	0.00 [0.00, 1.50]	0.05

Notes: Medians; interquartile range in brackets. Sample is all utility-years for the 21 CAP utilities (2014, 2015, 2017, 2019; $n = 60$ IBR across 15 utilities, $n = 24$ Uniform across 6), with structure fixed at the 2019 value. LTSA water stock is reported for 2019 only; rate changes count survey years with a residential rate change (utility level, $n = 21$). p-values from two-sided Wilcoxon rank-sum tests.

B.7 Welfare under an Observed Status-Quo Benchmark

The headline welfare tables benchmark optimal pricing against a status quo that prices at observed tariffs but purchases CAP water under the same optimal rule, so the comparison isolates the pricing effect. Table A10 instead fixes both observed prices and observed CAP deliveries, the fully observed status quo. The welfare gains are substantially larger under this benchmark, and far more so in the stressed regime, because much of the apparent gain is then a correction of the utilities' observed CAP procurement rather than of their prices. The demand reductions are identical to the headline because they are driven by prices, while the change in CAP purchases is several times larger, confirming that the procurement margin, not pricing, accounts for the difference between the two benchmarks.

Table A10: Welfare and Quantity Change under an Observed Status Quo

	5-year		100-year	
	Mild	Stressed	Mild	Stressed
ΔW (\$M/yr)	31.4	123.3	32.1	48.9
ΔQ (AF/yr)	-75,574	-95,929	-85,101	-146,455
ΔX CAP (AF/yr)	-137,223	-337,817	-139,663	-175,466

Notes: Sample total across all 21 CAP utilities under an alternative status quo that fixes *both* observed prices and observed CAP deliveries, mixing pricing and procurement effects. $\Delta W = \Delta CS + \Delta NR$ is the flow welfare change as a per-year annuity (\$M/yr); ΔQ (demanded) and ΔX (CAP) are annual quantities (AF/yr; negative = reduction).

B.8 Model Robustness Variants

Table A11: Welfare and Quantity Change from Optimal Pricing: Model Robustness Variants (Sample Total, 100-Year Horizon)

	Headline	Deterministic	No bonds	Single ε	AP ε
Panel A: Mild regime					
ΔW (\$/conn/yr)	16.7	16.6	10.8	15.7	18.8
ΔW (\$M/yr)	21.0	20.8	13.6	19.7	23.6
ΔQ (AF/yr)	-85,101	-84,738	-116,653	-83,665	-103,491
ΔX CAP (AF/yr)	-89,655	-88,752	-96,229	-91,427	-100,107
Panel B: Stressed regime					
ΔW (\$/conn/yr)	30.6	30.9	27.6	22.5	41.9
ΔW (\$M/yr)	38.5	38.8	34.7	28.3	52.6
ΔQ (AF/yr)	-146,455	-105,473	-184,568	-116,816	-131,900
ΔX CAP (AF/yr)	-106,533	-78,705	-146,042	-90,347	-87,303

Notes: ΔW (per-year annuity) is reported per connection-year (connection-weighted) and as a sample total; ΔQ and ΔX are horizon-average annual water demand and CAP purchases (negative = reduction). Variants: *Deterministic* fixes CAP delivery at the regime mean; *No bonds* removes the revenue reserve; *Single* ε applies -0.53 to both sectors; *AP* ε uses the average-price elasticities (-0.67/ -0.70).

Table A12: Sector Decomposition of the Welfare Change from Optimal Pricing (single residential elasticity -0.53, AMP marginal price, fee-inclusive)

	All CAP		IBR		Uniform	
	Mild	Stressed	Mild	Stressed	Mild	Stressed
ΔCS, residential	\$26.8	\$-32.3	\$29.8	\$-41.8	\$10.5	\$19.2
ΔCS, non-residential	\$16.1	\$-3.1	\$23.7	\$0.7	\$-25.1	\$-23.5
ΔNR (utility)	\$-27.2	\$58.0	\$-40.0	\$54.7	\$42.3	\$75.7
ΔW (total)	\$15.7	\$22.7	\$13.5	\$13.7	\$27.7	\$71.3

Notes: Connection-weighted mean welfare change (\$/conn/yr) across the 21 CAP utilities at the 100-year horizon (per-year annuity), policy-own terminal-cap total-surplus measure. The status quo charges observed AMP prices plus the fixed fee, split across sectors by each utility's residential connection share. $\Delta W = \Delta CS + \Delta NR$; IBR $n = 15$, uniform $n = 6$.

Table A13: Monthly Bill Change under a Single Pooled Demand Elasticity

	Mild		Stressed	
Consumption level	$/mo	%	$/mo	%
Residential				
5 kgal	$-0.00	-0.0%	$0.44	2.1%
10 kgal	$4.16	11.1%	$5.03	13.4%
15 kgal	$-1.95	-3.0%	$-0.64	-1.0%
Non-residential				
25 kgal	$1.25	1.2%	$3.43	3.3%
50 kgal	$3.71	1.8%	$8.07	3.9%
100 kgal	$5.84	1.4%	$14.56	3.5%

Notes: Bill change at the stated consumption when the observed 2019 schedule is replaced by the utility's optimal all-volumetric price under a single pooled demand elasticity ($\varepsilon = -0.53$ for both sectors), pooled across all 21 CAP utilities. $/mo is the connection-weighted mean; % the connection-weighted aggregate change. Directly comparable to Table 9.

Table A14: Monthly Bill Change from Moving to Optimal Pricing, by Rate Structure

	IBR		Uniform	
Consumption level	$/mo	%	$/mo	%
Panel A: Normal (mild) optimal price				
Residential				
5 kgal	$1.96	9.7%	$-9.03	-36.2%
10 kgal	$6.33	16.8%	$-4.51	-12.4%
15 kgal	$-1.37	-2.0%	$-0.43	-0.9%
Non-residential				
25 kgal	$-1.63	-1.5%	$1.86	2.5%
50 kgal	$-4.49	-2.1%	$18.16	13.5%
100 kgal	$-13.50	-3.1%	$50.75	20.0%
Panel B: Shortage (stressed) optimal price				
Residential				
5 kgal	$2.80	14.0%	$-8.74	-35.0%
10 kgal	$8.02	21.3%	$-3.92	-10.8%
15 kgal	$1.17	1.7%	$0.44	0.9%
Non-residential				
25 kgal	$4.19	3.9%	$4.34	5.8%
50 kgal	$7.14	3.3%	$23.11	17.2%
100 kgal	$9.74	2.2%	$60.66	23.9%

Notes: Change in the monthly bill at the stated consumption (q kgal) when the observed 2019 schedule is replaced by the utility's optimal all-volumetric price p^*, for increasing-block (IBR, $n = 15$) and uniform-rate ($n = 6$) CAP utilities. Panels A and B use the mild- and stressed-regime optimal price. $/mo is the connection-weighted mean; % is the connection-weighted aggregate change. A positive value is a bill increase.

B.9 Timmins-style Alternative Objective Function

Table A15: Welfare and Quantity Change under the Group-Weighted (Timmins-Style) Objective

	5-year		100-year	
	Mild	Stressed	Mild	Stressed
ΔW (\$M/yr)	1.1	6.5	16.9	37.1
ΔQ (AF/yr)	-88,093	-104,496	-91,916	-147,468
ΔX CAP (AF/yr)	10,359	-20,539	-94,548	-106,509

Notes: Sample total across all 21 CAP utilities under the group-weighted (Timmins-style) objective. $\Delta W = \Delta CS + \Delta NR$ is the flow welfare change as a per-year annuity (\$M/yr); ΔQ (water demanded) and ΔX (CAP purchased) are annual quantities (AF/yr; negative = reduction). Welfare is measured with equal (unweighted) surplus so it is comparable to the headline Table 7; only the policy-generating objective is weighted. Benchmark: observed prices, optimal CAP.

Table A16: Sector Decomposition of the Welfare Change under the Group-Weighted (Timmins-Style) Objective

	All CAP		IBR		Uniform	
	Mild	Stressed	Mild	Stressed	Mild	Stressed
ΔCS, residential	\$33.3	\$9.8	\$39.4	\$10.6	\$0.2	\$5.2
ΔCS, non-residential	\$4.6	\$-49.8	\$8.1	\$-54.7	\$-14.5	\$-23.2
ΔNR (utility)	\$-24.5	\$69.6	\$-37.0	\$64.7	\$43.0	\$96.3
ΔW (total)	\$13.4	\$29.6	\$10.6	\$20.6	\$28.7	\$78.4

Notes: Connection-weighted mean welfare change per connection-year (\$/conn/yr) across the 21 CAP utilities at the 100-year horizon, under the group-weighted objective $w_1 CS_1 + w_2 CS_2 + \beta V$ with each utility's recovered residential and non-residential welfare weights. The status quo charges observed AMP marginal prices plus the observed fixed fee; the optimum recovers fixed costs through the volumetric price. IBR $n = 15$, uniform $n = 6$. Directly comparable to the equal-weight Table 8.

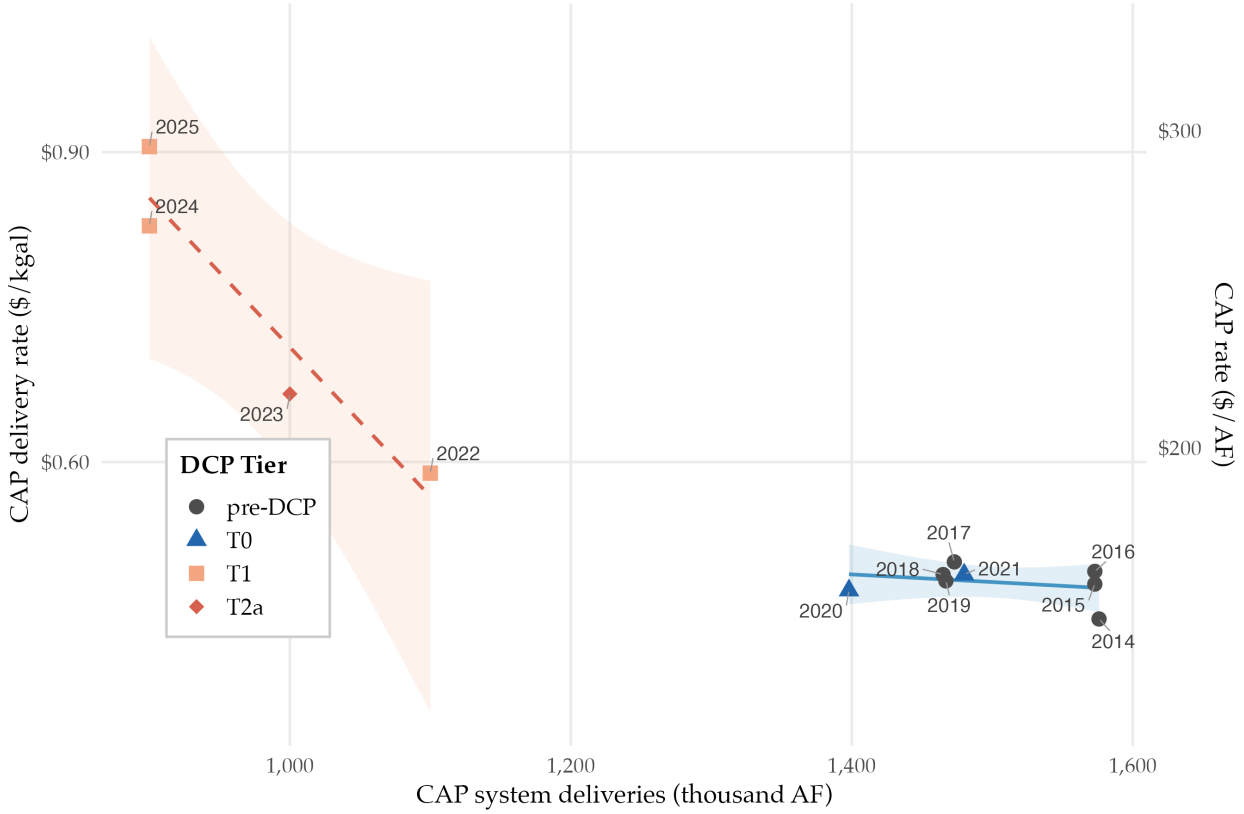
C Additional Figures

Figure A5: Central Arizona Project System Map



Notes: The Central Arizona Project aqueduct system diverts Colorado River water from Lake Havasu (western Arizona) eastward and southward through a 336-mile canal system to serve the Phoenix and Tucson metropolitan areas. Source: Central Arizona Project (reproduced with permission); `https://www.cap-az.com`.

Figure A6: Central Arizona Project Delivery Volume and Price Relationship



Notes: The figure displays OLS linear regressions of CAP delivered water price on total water delivered for two periods. Each point represents one annual observation of total CAP M&I deliveries and the corresponding delivery rate. Y-axis: CAP M&I delivery rate ($/kgal). X-axis: total annual CAP water delivered (kgal). Blue dashed line: pre- and early-DCP period (2014–2021, $n = 8$); orange dashed line: post-shortage period (2022–2025, $n = 4$). In the early period there is no statistically significant relationship between delivery price and volume. For years 2022–2025, there is a strong negative correlation between CAP rate and amount delivered, determined by DCP shortage tier. Source: Central Arizona Water Conservation District delivery records.