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

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

August 2026

Abstract

Municipal water utilities are mandated to recover costs, balance equity concerns, and send conservation signals. Yet the degree to which observed prices reflect the scarcity value of water remains largely unexplored. This paper develops a stochastic dynamic program in which the optimal two-sector price decomposes into a Ramsey markup, a scarcity rent on the water stock, and sector-specific marginal cost. Residential and non-residential demand elasticities of -0.59 and -0.82, respectively, are estimated from a novel panel of 194 Arizona water utilities (2014–2019). Simulated model-implied optimal prices show that increasing block rate (IBR) utilities price closer to the dynamic optimum than uniform-rate utilities. The welfare cost of suboptimal pricing is approximately \$5 per connection larger under uniform rates than IBR in mild supply conditions and approximately \$21 per connection larger under shortage stress. At the median 10 kgal consumption tier, uniform-rate CAP utilities price approximately 25 percent below their model-implied optimum while IBR CAP utilities price approximately 11 percent below theirs, a gap roughly twice as large under uniform rates. In this sample, IBR utilities more fully incorporate the scarcity rent into observed prices than uniform-rate utilities, and uniform-rate utilities retain a welfare gap of approximately \$19 per connection under severe shortage conditions.

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

1 Introduction

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). Water utilities and regulators must adequately design and set rates to incorporate the scarcity and uncertainty of supplies to maximize social benefit.

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 as compared to water-scarce regions like the western United States.

In this paper, I study the extent to which scarcity rents have been incorporated into water prices of the recent past and the implications rate structures have on natural resource value and social welfare. I first develop a dynamic Ramsey pricing model for a two-sector (residential and non-residential) municipal water utility facing stochastic supply shocks. I extend Sağlam (2019) by centering 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

[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).

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. I apply the model to Arizona and estimate parameters using a novel panel of 194 water utilities over 2014–2019, assembled from utility-level rate schedules, supply and withdrawal records, financial filings, and purchased delivery data to provide evidence that observed prices deviate from optimal. Finally, I simulate the optimal pricing policy for different utilities to investigate the magnitude of pricing inefficiencies by rate structure.

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 resources are not scarce. To my 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. I fill this gap by calibrating a stochastic dynamic program to utility-level data and testing how rate structure type affects alignment between observed and optimal prices over simulated resource scarcity scenarios.

I make three contributions. First, I provide simulation evidence that IBR utilities in this sample price observed rates closer to the dynamic Ramsey optimum than uniform-rate utilities, and that this alignment is larger under more scarce water scenarios. Municipal water ratemaking requires formal regulatory approval and generates political friction, so observed rates often lag optimal prices by five years or more (Hanemann, 1998; McRae, 2015). One explanation for this cross-sectional pattern is that a flat rate, once set, cannot differentiate scarcity signals across consumption levels between rate reviews; an IBR's tiered schedule embeds a price gradient that partially approximates the scarcity rent even when the overall rate level cannot be adjusted. I find that in the recent past, the median IBR CAP utility prices approximately 11 percent below its model-implied optimum at the 10 kgal consumption tier, while the median uniform-rate CAP utility prices approximately 25 percent below its optimum — a gap roughly twice as large under uniform rates. Second, I show that uniform-rate utilities face larger welfare losses from suboptimal pricing than IBR utilities under both mild and severe supply shock scenarios. I simulate five-year paths under mild and severe supply shock scenarios and find that welfare gains for utilities with uniform rate structures moving to optimal prices are much larger than those operating at increasing block rates. Third, I provide estimates of price elasticity of water demand for both residential and non-residential sectors. I estimate both from the same panel using a cross-market instrument in the style of Hausman et al. (1994) and Nevo (2001) for the endogenous marginal price.[2] The empirical water demand literature has focused almost exclusively on the residential sector (Dalhuisen et al., 2003; Olmstead, 2010), leaving non-residential demand — water sold to businesses, institutions, and other non-household customers, which accounts for a substantial share of total municipal withdrawals — largely unexamined at this scale. I show that non-residential water demand is more price-elastic than residential. Restricting demand estimation to the residential sector would overstate the quantity demanded by non-residential customers at higher prices and thereby understate the welfare gain from price increases.

The empirical contribution rests on a data set assembled from multiple public sources and those requiring public records requests to access. I combine utility-level rate schedules from University of North Carolina-Chapel Hill's Environmental Finance Center for four survey years (2014, 2015, 2017, 2019), annual supply and withdrawal quantities by source type from the Arizona Department of Water Resources, financial data from Arizona Corporation Commission annual reports, and Central Arizona Project (CAP) delivery records by entity and contract tier. The resulting panel covers 194 utilities supplying approximately 5.6 million of Arizona's 7.6 million residents. The CAP shortage mechanism translates Lake Mead elevations directly into delivery reductions by contract type and provides the empirical analog to the supply shock in the model.

I estimate residential price elasticity of demand at -0.51 to -0.59 using an instrumental variables strategy that instruments for the marginal price with the average price charged by neighboring utilities in the same county, lagged one period. This strategy follows from modern applications and recent discussion on the validity of Hausman instruments (Autor et al., 2013; Berry and Haile, 2021; MacKay and Miller, 2025) and is in a similar range of recent water demand elasticity estimates (Mansur and Olmstead, 2012; Wichman, 2014; Szabó, 2015; Brent and Ward, 2019; Wichman, 2024; Wietelman et al., 2025). I provide the first large-scale, urban non-residential water demand estimates at -0.78 to -0.83.[3] This range is on the more elastic end of residential water demand

[2] Berry and Haile (2021) and MacKay and Miller (2025) detail the conditions which must hold for this instrument to identify the price parameter. The key condition is that the demand shock in the focal market is mean-independent of the price in the other market conditional on observables.

[3] 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, insti-

estimates but in line with recent evidence from Burlig et al. (2021) on agricultural water demand estimates.

Following Wichman (2024), I estimate marginal costs and recover utility-year estimates that reflect differences in regulation and use of external purchased water. I combine these marginal cost estimates with the full data to recover optimal prices using the model for each utility-year combination. I find that utilities with increasing block rates price water substantially closer to the model-derived optimal regardless of CAP water use. The median IBR utility prices residential water approximately 43 cents per thousand gallons (kgal) closer to the model-implied optimal prices compared to the median uniform rate utility. I further explore this result by simulating five years of the model for all CAP water purchasing utilities in my sample and show that the welfare cost of uniform-rate underpricing substantially exceeds that of IBR underpricing. Welfare change is measured as the difference in value functions evaluated at status quo prices and those from the dynamic optimal where optimal prices are chosen to maximize consumer surplus, utility net revenues, and stock of water for future use. Among utilities at typical underpricing levels, over 500 simulations of 5-year paths, the median IBR utility's current pricing is approximately at or above the model-implied optimum, yielding a welfare change near zero (-\$1.8 per connection under mild conditions; -\$1.6 under stress). By contrast, the median uniform-rate utility prices substantially below the optimum, implying a welfare gain of \$3.2 per connection under mild shortage conditions and \$19.1 per connection under severe shortage conditions. The larger welfare cost of uniform underpricing under severe scarcity illustrates how the flat-rate structure provides no mechanism to embed a scarcity signal as supply tightens between rate reviews.

I build most directly on the stochastic dynamic programming literature for water resource 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. I also connect this pricing literature to the natural capital accounting framework of Fenichel and Abbott (2014) and Fenichel et al. (2016), who show that the scarcity rent of a resource stock is the theoretically

tutional, and industrial customers) served by municipal systems.

correct measure of its natural capital price and demonstrate recovery of stock prices from observed behavior for groundwater. I provide the first empirical implementation of a Saglam-style stochastic dynamic program using utility-level data and connect the recovered scarcity rent directly to the natural capital literature, situating the rate structure comparison within a broader accounting of water as a stock resource. Under greater supply uncertainty, scarcity rents rise slightly due to utilities having less access to purchase and store water for future periods even if they have revenue surplus.[4]

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 trade-off 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) shows that replacing quantity rationing with market-clearing prices raises welfare substantially. I depart from this literature by asking not how demand responds to observed rate structures, but how well observed rate designs approximate the optimal price which incorporates the scarcity rent. This question has gone unanswered and connects directly to whether rate design captures the natural capital value of the water stock as supply conditions change between rate reviews.

The remainder of this paper is organized as follows. Section 2 presents the dynamic pricing model and derives optimal prices analytically. Section 3 provides institutional background on the Central Arizona Project and the shortage declaration mechanism. Section 4 describes the data. Section 5 estimates residential and non-residential demand. Section 6 estimates marginal costs. Section 7 presents results comparing observed and optimal prices across rate structures and quantifies the welfare loss from suboptimal pricing. Section 8 concludes.

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

2 Model

In this model, a benevolent water manager chooses prices and 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:

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

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

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

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

$$w \geq 0, \tag{5}$$

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 (5). 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 $\{x(\theta) \geq 0\}$, 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 $\eta(\theta)$ denote the Lagrange multiplier on the purchase limit (4). Equation (5) is the non-negativity constraint on the water stock.

2.2 Timing

I consider residential (sector 1) and non-residential (sector 2). Let θ denote a supply-side stochastic shock to the model. I 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:

$$w'(\theta) + q_1(p_1) + q_2(p_2) \leq w + w_f + x(\theta); \quad \forall\theta, \tag{6}$$

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 scaled by the resource 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)[5], and the additional purchased water $(x(\theta))$. I assume that the water manager can only purchase this water and not sell it, therefore $(x(\theta) \geq 0)$.

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

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:

$$q_1(p_1) + q_2(p_2) > w_f + x(\theta). \tag{7}$$

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 pricing water higher to signal scarcity to users.

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), I assume marginal costs may differ between sectors due to differences in production costs or accessibility[6]. The water manager's budget constraint is:

$$b'(\theta) + p_x x(\theta) \leq (p_1 - c_1)q_1(p_1) + (p_2 - c_2)q_2(p_2) + Rb - \tau, \tag{8}$$

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

2.5 Optimal Policy Functions

The optimal policy functions follow from the first-order (FOC) and envelope (EC) conditions of the dynamic program (1)–(5). 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. I 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.

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

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:

$$\text{FOC:} \quad \frac{\beta\, \partial V(w'(\theta),\, b'(\theta))}{\partial w'} = \lambda(\theta), \tag{9}$$

$$\text{EC:} \quad \frac{\partial V(w,\, b)}{\partial w} = \mathbb{E}_\theta[\lambda(\theta)]\,. \tag{10}$$

Iterating equation (10) forward one period and substituting (9) yields the Euler equation for the carryover water stock:

$$\lambda(\theta) = \beta\, \mathbb{E}_{\theta'}\left[\lambda'(\theta')\right]. \tag{11}$$

Equation (11) 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:

$$\text{FOC:} \quad \frac{\beta\, \partial V(w'(\theta),\, b'(\theta))}{\partial b'} = \mu(\theta), \tag{12}$$

$$\text{EC:} \quad \frac{\partial V(w,\, b)}{\partial b} = R\, \mathbb{E}_\theta[\mu(\theta)]\,, \tag{13}$$

which gives:

$$\mu(\theta) = \beta R\, \mathbb{E}_{\theta'}\left[\mu'(\theta')\right]. \tag{14}$$

The shadow value of a dollar of carryover revenue today equals the discounted expected shadow value next period scaled 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).

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 (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:

$$0 = \frac{\partial CS_i}{\partial p_i} + \mathbb{E}_\theta\left[-\lambda(\theta)\frac{\partial q_i}{\partial p_i} + \mu(\theta)\left((p_i - c_i)\frac{\partial q_i}{\partial p_i} + q_i(p_i)\right)\right]. \tag{15}$$

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 (15) simplifies to:

$$(\bar{\mu} - 1)\, q_i(p_i) + \left(\bar{\mu}(p_i - c_i) - \bar{\lambda}\right)\frac{\partial q_i}{\partial p_i} = 0. \tag{16}$$

Because demand $q_i(p_i)$ and its derivative $\partial q_i/\partial p_i$ are non-stochastic—a consequence of prices being set before θ is observed—the expectations in (15) reduce to the scalar multipliers $\bar{\lambda}$ and $\bar{\mu}$. Solving (16) for p_i yields the following result.

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

$$p_i = \underbrace{\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right)\left(\frac{-q_i(p_i)}{\partial q_i/\partial p_i}\right)}_{\text{Ramsey markup}} + \underbrace{\frac{\bar{\lambda}}{\bar{\mu}}}_{\text{scarcity rent}} + \underbrace{c_i}_{\text{marginal cost}}, \tag{17}$$

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 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 (17). 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.

Because both sectors share the same scarcity rent $\bar{\lambda}/\bar{\mu}$, prices differ across sectors only through their Ramsey markups, which depend on sector-specific demand elasticities. With $|\hat{\varepsilon}_2| = 0.82 > |\hat{\varepsilon}_1| = 0.59$, non-residential demand is more elastic than residential demand, so the inverse-elasticity rule assigns the residential sector the larger Ramsey markup. The scarcity rent, however, 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 $x(\theta)$:

$$\lambda(\theta) - p_x\, \mu(\theta) = \eta(\theta) \geq 0, \tag{18}$$

with complementary slackness $\eta(\theta)(\bar{x}(\theta) - x(\theta)) = 0$. Equation (18) 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 $x(\theta) = \bar{x}(\theta) = \bar{x} \cdot \theta$.

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 are not subject to variation year-to-year. 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 6 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. The amount of CAP's allotment is determined by the level of Lake Mead 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

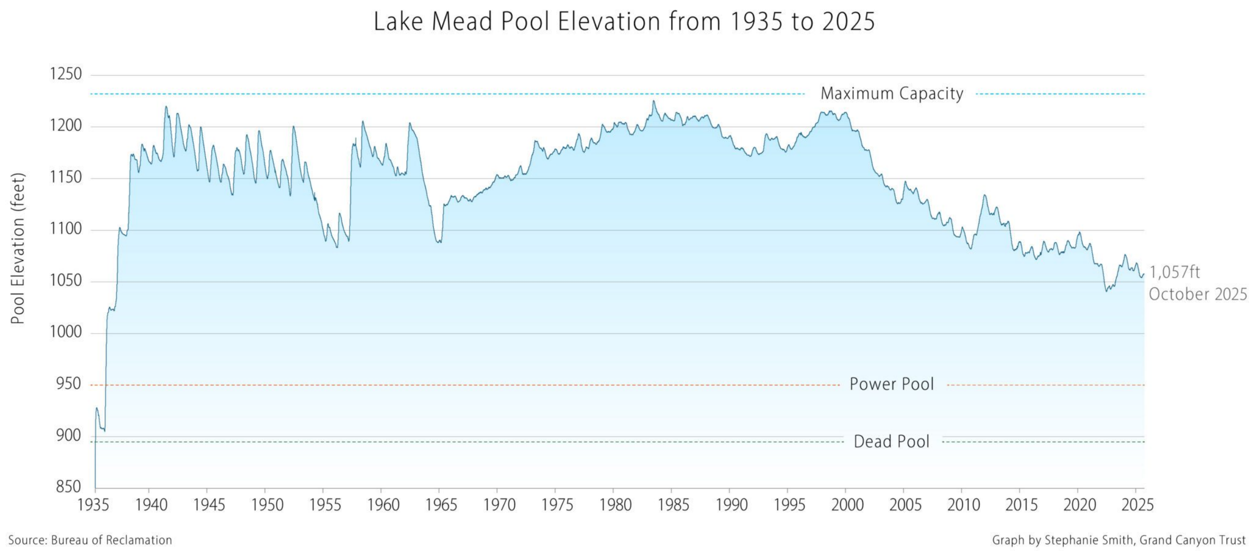


Figure 1: Lake Mead Pool Elevation, 1935–2025

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).

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 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 Credits (LTSCs), which can be withdrawn in future periods. Through 2024, the AWBA had accumulated approximately 4.4 million acre-feet of LTSCs, with roughly 2.3 million acre-feet designated specifically to firm M&I subcontract supplies during shortage events (Arizona Water Banking Authority, 2024).

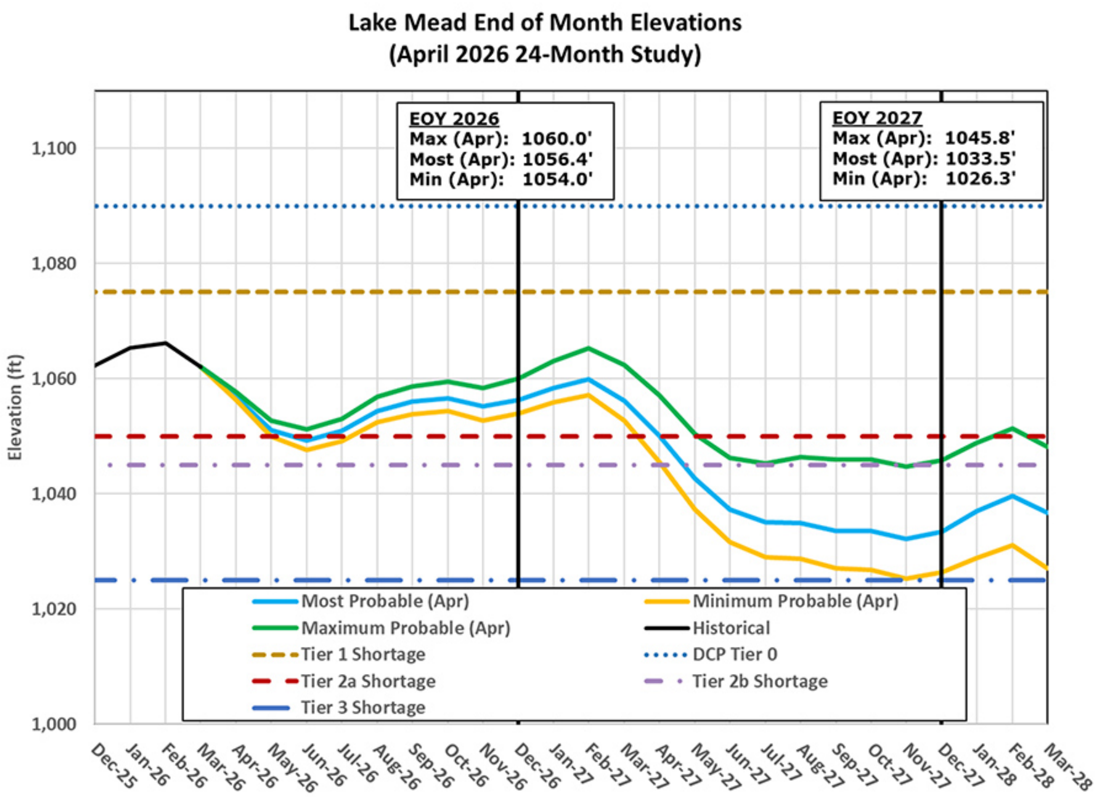


Figure 2: Lake Mead End-of-Month Elevation Projections, 2026–2028

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`.

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 trade-off at the center of the Ramsey pricing problem.

4 Data

I 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.

4.1 Water Rates Data

I 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 I 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 I detail how these bill amounts are used to compute average marginal prices.

4.2 Revenue and Expenditure

I compile revenue and expenditure data from several sources. I include these data from a public records request for community water service (CWS) providers.[7] For non-CWS utilities, I use operating ratio[8] if available from EFC data[9] to approximate revenue and expenditure. Otherwise, I manually collect PDF files of city and county annual financial reports through Google searches as well as the Arizona Financial Transparency Portal.[10]

[7] 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.

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

[9] For years 2014, 2015, and 2017 I observe the utility's operating ratio, the ratio of revenue to operating expenditure in the EFC data. For later years I compile revenue and expenditure data by hand from annual reports provided by the Arizona Corporation Commission.

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

4.3 Arizona Supply and Withdrawal Quantities

I 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. I also observe the quantity of water withdrawn by source type. Source type is broken out to groundwater (GW), surface water (SW), Colorado River, water from the Central Arizona Project (CAP), and other.[11] I match 194 utilities with rates and supply data for all four years.

4.4 Central Arizona Project

I 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 request.

I 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. I 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 roughly equal in priority and 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.

[11] Other includes reclaimed water.

4.5 Variable Construction

From the rates data, I 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 me to capture approximate marginal price changes across different levels of consumption. I include larger AMP thresholds than Wichman (2024) because several utilities in my data set have an average monthly demand per connection greater than 25 kgal and I 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 I observe utility-level annual quantity supplied. As shown in the table, about 84 percent of utilities' water supply comes from groundwater, another 10% from the Colorado River, 2% from other surface water sources, and the remaining 4% from reclaimed water on average. There is wide variation across utilities in average monthly supply per connection. However, the mean and median fall in the range of average family monthly water consumption.[12] Utilities in my sample have revenues approximately 9 percent above operating costs on average.

The sample is primarily made up of small, groundwater-dependent utilities. The median utility serves 874 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 8,432 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. The Colorado River share

[12] 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).

distribution is similarly bimodal: the median utility holds no CAP contract at all, while the 90th percentile reaches 47 percent. I 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, where the stochastic supply problem is empirically operative and the scarcity rent is nonzero. 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 trade-off 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
Service Population	26,953	124,443	215	2,088	38,016	822
# of Connections	8,432	34,512	100	874	16,799	822
Total Annual Supply (AF)	5,469	26,239	19	296	6,820	842
Supply per Connection (kgal/month)	12.32	15.64	2.69	8.84	22.05	822
Operating Ratio (Rev./Exp.)	1.09	0.25	0.77	1.08	1.35	321
AMP_{4kgal} ($/kgal)	2.50	1.71	0.80	2.25	4.00	841
AMP_{5kgal} ($/kgal)	3.27	2.04	1.25	3.00	5.48	841
AMP_{10kgal} ($/kgal)	3.56	2.26	1.42	3.11	5.87	841
AMP_{15kgal} ($/kgal)	4.38	3.33	1.65	3.61	7.72	841
AMP_{25kgal} ($/kgal)	4.46	3.49	1.65	3.72	7.74	842
AMP_{50kgal} ($/kgal)	4.71	3.85	1.75	3.85	8.00	842
$AMP_{100kgal}$ ($/kgal)	4.85	4.05	1.80	3.94	8.95	842
Groundwater	0.84	0.33	0.11	1.00	1.00	839
Surface Water	0.02	0.11	0.00	0.00	0.00	839
Colorado Riv. Share	0.10	0.26	0.00	0.00	0.47	839

Notes: Utility-year observations from the Arizona water utility panel, 2014–2019. N varies across variables due to missing financial and supply records. AMP_X denotes the average marginal price (in $/kgal) between consecutive consumption thresholds at X kgal/month; see Section 4 for construction. Supply shares (groundwater, surface water, Colorado River) sum to approximately 0.96 on average; the remaining $\approx$4 percent is from reclaimed water, which is not shown separately. The distribution of groundwater dependence is highly bimodal: 57 percent of utility-years report a groundwater share of one, and the 90th percentile also equals one. CAP-dependent utility-years are a distinct minority concentrated in the Phoenix and Tucson metropolitan service areas.

5 Demand Estimation

I 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, so 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, satisfying the exclusion restriction. The leave-one-out construction, 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). I 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 general form of the instrument for sector s in utility i, county c, and year t is:

$$\bar{P}_{s,c,t-1} = \frac{\sum_{j \neq i} P_{s,j,c,t-1}}{N_{ct} - 1} \quad \forall \quad j \neq i$$

where N is the number of utilities in county c at year $t-1$.

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

$$q_{sict} = \alpha_1 \widehat{P}_{sict} + X'_{ict}\gamma_1 + \delta_{ct} + \epsilon_{ict} \tag{19}$$

where q_{sict} is the equilibrium water quantity (in AF) sold in sector s of utility i in county c in year t, $\widehat{P}_{sict}$ is the fitted value of price from the first stage regression for sector s of utility i in county c in year t, X_{ict} are controls, δ_{ct} are county-by-year fixed effects, and ϵ_{ict} is the error term. (Note: α_1 and α_2 below denote demand regression coefficients and are distinct from both the model discount factor β and the cost-function elasticities $\hat{\beta}_1$, $\hat{\beta}_2$ in Section 6.)

The first stage regression to obtain the fitted values of water prices is:

$$P_{sict} = \alpha_2 \bar{P}_{s,c,t-1} + X'_{ict}\gamma_2 + \psi_{ct} + v_{sict} \tag{20}$$

where both P_{sict} and $\bar{P}_{s,c,t-1}$ are in logs. I instrument for water price with $\bar{P}_{s,c,t-1}$, the average sectoral water price within that utility's county lagged one period. ψ_{ct} are county-by-year fixed effects and v_{sict} is the error term.

5.1 Residential Water Demand

Table 2 presents estimates of the residential demand equation. I treat both volumetric price (AMP_{ict} and AP_{ict}) and the 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 price (AP_{ict}) at 10 kgal of consumption yield similar elasticity estimates of -0.59 and -0.54, respectively. The Cragg-Donald minimum eigenvalue statistic is 84.8 in Column (2) and 84.6 in Column (3). The Angrist and Pischke (2009) marginal F-statistic, which tests each excluded instrument alone in its first stage holding the other constant, is 11.3 for the price instrument and 71.1 for the fixed-fee instrument in Column (2), and 10.5 and 65.1 in Column (3).[13] The fixed fee elasticity estimates are statistically insignificant and I 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. I use $\hat{\varepsilon}_1 = -0.59$ 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 in absolute value 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.

[13]The Cragg-Donald F is the minimum eigenvalue statistic of Cragg and Donald (1993); the Stock-Yogo (2005) 10% maximal-bias critical value is 7.03 for K=L=2. The AP F is the Angrist and Pischke (2009) marginal F-statistic (cluster-robust), testing each excluded instrument alone in its first stage holding the other constant. In just-identified specifications with two endogenous regressors ($N = K = 2$), the mean 2SLS bias does not exist and neither statistic has a formal theoretical justification as a weak-instrument diagnostic; the appropriate procedure is the $g_{\min}$ statistic of Lewis and Mertens (2025) evaluated against median-bias critical values. I report both standard diagnostics for comparability with the existing literature. The Lewis and Mertens (2025) $g_{\min}$ statistics are 11.60 for the residential specification and 20.86 for the non-residential specification. The exact critical value against which to evaluate these in the just-identified case ($K = N = 2$) is not computable from the existing distributional results for this test; see Lewis and Mertens (2025) for discussion.

Table 2: Residential Water Demand

	(1) OLS	(2) IV	(3) IV
AMP_{ict}	-0.51***		
	(0.12)		
$\widehat{AMP}_{ict}$		-0.59**	
		(0.26)	
$\widehat{AP}_{ict}$			-0.54**
			(0.26)
Fixed Fee$_{ict}$	0.10		
	(0.10)		
$\widehat{\text{Fixed Fee}}_{ict}$		-0.03	-0.11
		(0.19)	(0.17)
ln(Median Income)	0.93***	0.93***	0.94***
	(0.16)	(0.15)	(0.15)
Observations	526	526	526
County×Year FE	✓	✓	✓
Cragg-Donald F	—	84.8	84.6
AP F: Price	—	11.3	10.5
AP F: Fixed Fee	—	71.1	65.1

Notes: The dependent variable is the natural logarithm of annual residential water delivered per capita (in AF/yr). Column (1) is OLS using the observed, uninstrumented price. Columns (2) and (3) instrument for both the marginal price and fixed fee using lagged leave-one-out county averages of neighboring utilities' prices; hats ($\widehat{\cdot}$) denote fitted values from the first-stage regression. AMP is the average marginal price at 10 kgal/month; AP is the average volumetric price at 10 kgal/month. Standard errors clustered by utility in parentheses. Years: 2014, 2015, 2017, 2019. Significance: ***p<0.01, **p<0.05, *p<0.10.

5.2 Non-Residential Water Demand

I 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 (1) presents the OLS baseline. All three IV specifications yield price elasticities in a tight range: -0.82, -0.78, and -0.83, respectively, with Column (2) significant at the 10 percent level, Column (3) significant at the 5 percent level, and Column (4) significant at the 5 percent level. The Cragg-Donald minimum eigenvalue statistic ranges from 46.5 to 50.3 across the three IV specifications; the Angrist and Pischke (2009) marginal F-statistic is 28.2 to 31.3 for the price instrument and 38.8 to 39.9 for the fixed-fee instrument, both well above conventional thresholds. The same $N = K = 2$ caveat noted above applies; the Lewis-Mertens $g_{\min}$ values for both specifications are reported in the footnote to Section 5.1. I use $\hat{\varepsilon}_2 = -0.82$ from the preferred Column (2) specification in the model calibration for non-residential water demand elasticity.

Non-residential demand is more price-elastic than residential demand at -0.82 versus -0.59. 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 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 welfare comparisons in Section 7.4. The fixed-fee elasticity of 0.66 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 is -2.0 across all specifications, indicating that utilities in higher-income counties have substantially lower non-residential water use per establishment. This is better interpreted as a compositional effect than a true income response: wealthier Arizona counties (Maricopa, Pima) are dominated by offices, healthcare, and retail — activities with low water intensity — while lower-income rural counties contain more food processing, light manufacturing, and construction supply operations. County median household income serves here as a reduced-form proxy for industrial composition rather than an income shifter for non-residential demand per se. I 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 at $N = 392$ and yields a price elasticity of -0.90, a shift of -0.07 from the baseline of -0.82, within the range of sampling variability across the three IV price-tier specifications. The income coefficient is essentially unchanged at -2.00 versus the baseline -2.02. 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 estimate of -0.82 is therefore robust to direct controls for service-area industrial composition. The income coefficient is not used in model calibration; only $\hat{\varepsilon}_2 = -0.82$ enters the SDP.

As a sensitivity check on the spatial allocation method, I implement three alternatives to the area-weighted ZIP-to-service-area allocation; Appendix Table 10 summarizes the results. The first two replace the area-fraction weight with population-based weights derived from census geography centroids: one using 2020 decennial census blocks and one using ACS 5-year census block groups. In both cases the matched sample falls well below the area-weighted baseline of $N = 392$: the census block approach yields 126 utility-years and the block group approach yields 236. The reductions arise because centroid-based assignment systematically excludes small and rural utilities whose service area polygons contain few or no census unit centroids. The price elasticity estimates from these restricted samples are -0.99 and -0.09, respectively — far from each other and from the baseline -0.82, indicating that the observed shifts reflect changes in sample composition rather than sensitivity to the weighting method. The third alternative replaces the area fraction with a hybrid weight proportional to each utility's geographic overlap with the ZCTA multiplied by its service-area population ($\text{overlap}_{km^2} \times \text{ServicePop}$), normalised to sum to one across all utilities overlapping each ZCTA. This preserves the full sample ($N = 387$, nearly identical to the baseline) and maintains strong first-stage diagnostics (Cragg-Donald $F = 50.9$, AP $F = 30.2$ and 41.7), but produces an elasticity of $+0.08$, a sign flip from the baseline. The reversal reflects a mechanical distortion: population-based weighting systematically shifts establishments toward large urban utilities, which already have high per-connection water use, altering the per-establishment quantity measure in a way that breaks the price-quantity relationship rather than refining it. Across all three alternatives, the instruments remain strong but the elasticity is unstable, ranging from -0.99 to $+0.08$. This range reflects sensitivity to the assumed distribution of establishments within ZCTAs rather than a genuine replication of the area-weighted result. The area-weighted specification is retained as the main estimate because the uniform-density assumption — establishments are distributed in proportion to geographic area — is the most natural baseline when establishment locations within ZCTAs are unobserved, and it is the only approach that maintains a representative cross-section of Arizona utilities across the full size distribution.

As a robustness check, appendix Tables 11 and 12 present results restricting to county-years in which at least five utilities report valid prices. The leave-one-out instrument for each utility is constructed from at least four peers, in which case 19 residential and 28 non-residential utility-years are excluded. The residential elasticity estimates are nearly identical to the full-sample estimates. The AP F-statistics also increase representing more similar peer pricing behavior. The non-residential estimates shift to -0.90, -0.97, and -0.96 across the three specifications, reflecting heterogeneity in non-residential water use between smaller rural county markets (which are excluded by the restriction) and larger multi-utility counties.

6 Marginal Cost Estimation

The optimal pricing formula in equation (17) requires utility-year estimates of marginal cost c_i. I 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.

Table 3: Non-Residential Water Demand (Per Establishment)

	(1) OLS	(2) IV	(3) IV	(4) IV
AMP_{ict}^{com}	-0.72**			
	(0.33)			
$\widehat{AMP}_{ict}^{com}$ (25–50 tier)		-0.82*		
		(0.44)		
$\widehat{AMP}_{ict}^{com}$ (10–25 tier)			-0.78**	
			(0.36)	
$\widehat{AP}_{ict}^{com}$ (avg, 50 kgal)				-0.83**
				(0.41)
Fixed Fee_{ict}^{com}	0.32			
	(0.34)			
$\widehat{\text{Fixed Fee}}_{ict}^{com}$		0.66*	0.43	0.58
		(0.38)	(0.39)	(0.38)
ln(Payroll/Estab)	-0.29	-0.24	-0.27	-0.27
	(0.28)	(0.29)	(0.30)	(0.29)
ln(Median HH Income)	-2.05***	-2.0***	-2.1***	-2.1***
	(0.58)	(0.57)	(0.58)	(0.58)
Observations	392	392	392	392
County×Year FE	✓	✓	✓	✓
Cragg-Donald F	—	50.3	46.5	49.6
AP F: Price	—	31.3	28.2	31.8
AP F: Fixed Fee	—	39.9	38.8	39.8

Notes: The dependent variable is the natural logarithm of non-residential water delivered per non-residential establishment (in kgal/yr). Establishment counts are from ZIP Business Patterns (ZBP), allocated to utility service areas via area-proportional ZCTA intersection. Column (1) is OLS using the average marginal price at the 50 kgal threshold (AMP_{50}). Columns (2)–(4) instrument for both the volumetric price and fixed fee using lagged leave-one-out county averages; column headers identify the price tier used for each IV specification. Average payroll per establishment (ZBP, $1,000s) controls for establishment size. Standard errors clustered by utility. Years: 2014, 2015, 2017, 2019. Significance: ***p<0.01, **p<0.05, *p<0.10.

6.1 Specification

I 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, I 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:

$$\ln(E_{it}) = \beta_1 \ln(Q_{it}) + \beta_2 \left[\ln(Q_{it}) \times \mathbf{1}(\text{CAP}_i)\right] + \eta_{ct} + \varepsilon_{it}, \tag{21}$$

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, and η_{ct} are county-by-year fixed effects that absorb common variation in input costs, drought severity, and regulatory conditions. 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 (21), the implied marginal cost is:

$$\widehat{MC}_{it} = \hat{\beta}_{it} \cdot \frac{E_{it}}{Q_{it}} = \hat{\beta}_{it} \cdot AC_{it}, \tag{22}$$

where $\hat{\beta}_{it} = \hat{\beta}_1 + \hat{\beta}_2 \cdot \mathbf{1}(\text{CAP}_i)$ is the estimated cost elasticity for utility i and $AC_{it} = E_{it}/Q_{it}$ is the observed average cost. Marginal cost is therefore the product of the cost elasticity and average cost. The fixed cost component that must be recovered through the Ramsey markup is:

$$\hat{\tau}_{it} = (1 - \hat{\beta}_{it}) \cdot E_{it}, \tag{23}$$

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 $\hat{\beta}_1 = 0.902$, implying that a 1 percent increase in total supply is associated with a 0.9 percent increase in operating expenditure. The CAP interaction adds $\hat{\beta}_2 = 0.031$ ($SE = 0.013$), so the cost elasticity for AMA utilities that purchase CAP water is 0.933, reflecting the higher variable costs associated with CAP water purchases relative to groundwater extraction. Non-AMA utilities (Column 3) exhibit greater economies of scale, with a cost elasticity of $\hat{\beta}_1 = 0.724$, 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 \$3.57 per kgal for CAP utility-years and \$5.15 per kgal for non-AMA utility-years. For AMA utilities, the small difference between marginal and average cost ($\hat{\beta}_1 = 0.902$ implies $MC \approx 0.90 \times AC$) 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 ($\hat{\beta}_1 = 0.724$) means the Ramsey markup would account for roughly 28 percent of average cost.

Expenditure-quantity simultaneity biases $\hat{\beta}$ upward: utilities that plan for larger output volumes also authorize larger capital and operating budgets, so higher Q predicts higher E partly through budget-setting rather than through production technology alone. An upward bias on $\hat{\beta}$ implies a downward bias on $(1 - \hat{\beta}) \cdot E$, understating the fixed cost $\hat{\tau}$ that enters the VFI budget constraint. With $\hat{\tau}$ too small, the VFI budget constraint is less binding than in the true cost environment: the model-optimal price p^* is calibrated to a smaller fixed-cost recovery requirement, producing a lower optimal price and a smaller model-derived scarcity rent $\bar{\lambda}/\bar{\mu} = p^* - AC$. The estimated $\bar{\lambda}/\bar{\mu} = \$0.44$ per kgal should therefore be interpreted as a conservative lower bound on the true scarcity rent.

Table 4: Cost Function Estimates

	Pooled (1)	AMA (2)	Non-AMA (3)
ln(Q)	0.745***	0.902***	0.724***
	(0.035)	(0.050)	(0.044)
ln(Q)×CAP user	0.057***	0.031**	
	(0.011)	(0.013)	
Observations	581	156	425
R^2	0.91	0.96	0.82
County-Year fixed effects	✓	✓	✓

Notes: Dependent variable: ln(annual operating expenditure, \$). Q_{it} is total annual supply in kgal. County×Year fixed effects. Standard errors clustered by utility in parentheses. Implied marginal cost: $\widehat{MC}_{it} = \hat{\beta}_1 \times AC_{it}$; AMA CAP utilities use $(\hat{\beta}_1 + \hat{\beta}_2) \times AC_{it}$, where $AC_{it} = E_{it}/Q_{it}$. Years: 2014, 2015, 2017, 2019. Of the 776 potential utility-year observations (194 utilities × 4 survey years), 581 have non-missing expenditure data; attrition is concentrated among small utilities that do not file annual reports with the Arizona Corporation Commission. Significance: ***p<0.01, **p<0.05, *p<0.10.

7 Results

7.1 Optimal Prices

Table 5 reports the distribution of model-implied optimal prices for the residential and non-residential sectors across the four rate structure and CAP status groups. Residential optimal prices are evaluated at 10 kgal per connection per month and non-residential at 25 kgal per connection per month, matching the consumption benchmarks used throughout the demand estimation. For CAP utilities, p_1^* and p_2^* are solutions to the stochastic dynamic program evaluated at each utility-year's observed water stock; for non-CAP utilities, the scarcity component is zero and the benchmark collapses to average cost. The median optimal residential price for uniform-rate CAP utilities is \$3.81 per kgal and \$3.94 for non-residential, compared to medians of \$4.56 and \$4.68 for their IBR counterparts. Non-CAP utilities face substantially higher benchmarks, with medians of \$5.08 and \$5.67 per kgal for uniform and IBR utilities respectively, reflecting the higher average costs that prevail outside the CAP service area where groundwater extraction is the marginal supply source. These figures are economically plausible: median observed prices across the sample run \$3.06 to \$3.47 per kgal, placing the optimal benchmark 25 to 65 percent above what utilities currently charge.

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

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Residential optimal, p_1^*	3.74	0.94	2.52	3.81	4.92	15
Non-residential optimal, p_2^*	3.85	0.97	2.47	3.94	5.02	15
Panel B: *Uniform rates, non-CAP utilities*						
Residential optimal, p_1^*	6.71	5.37	2.00	5.08	14.11	134
Non-residential optimal, p_2^*	6.71	5.37	2.00	5.08	14.11	134
Panel C: *IBR rates, CAP utilities*						
Residential optimal, p_1^*	5.09	2.08	3.48	4.56	9.12	31
Non-residential optimal, p_2^*	5.12	1.97	3.39	4.68	8.86	31
Panel D: *IBR rates, non-CAP utilities*						
Residential optimal, p_1^*	7.27	5.49	2.54	5.67	12.98	174
Non-residential optimal, p_2^*	7.27	5.49	2.54	5.67	12.98	174

Notes: Model-implied optimal prices in \$/kgal. For CAP utilities (Panels A and C), p_1^* and p_2^* are the residential and non-residential solutions to the stochastic dynamic program evaluated at the observed water stock (w_{it}, b_{it}). For non-CAP utilities (Panels B and D), the scarcity component is zero and the optimal price equals the average cost benchmark: $p_1^* = p_2^* = AC_{it}$. Sample restricted to IBR and uniform-rate utilities with $AC_{it} \leq \$30$/kgal. Years: 2014, 2015, 2017, 2019.

Comparing observed prices to the model-implied optimum across the full utility-year panel reveals a consistent gap that differs sharply by rate structure. At the 10 kgal per month consumption tier, the median adjusted residential price gap (Table 16) is \$0.95 per kgal for uniform-rate CAP utilities and \$0.52 per kgal for IBR CAP utilities. Expressed as a fraction of each group's median optimal price from Table 5, uniform-rate CAP utilities price approximately 25 percent below their model-implied optimum (0.95/\$3.81) while IBR CAP utilities price approximately 11 percent below theirs (0.52/\$4.56). The uniform-rate gap is more than twice the IBR gap on this metric — IBR utilities are roughly 54 percent closer to the scarcity-inclusive optimum than uniform-rate utilities when the gap is expressed as a share of each structure's optimal price.

The residential and non-residential optimal prices track closely within each panel, differing by less than 15 cents at the median. Both are governed by the same theoretical decomposition — marginal cost plus a Ramsey markup plus the scarcity rent, which is identical across sectors at a given utility-year, reflecting the fact that a unit of water drawn by any user carries the same opportunity cost to the utility. The sector-specific Ramsey markups differ because demand elasticities differ ($|\hat{\varepsilon}_2| = 0.82$ versus $|\hat{\varepsilon}_1| = 0.59$): the inverse-elasticity rule implies a smaller per-unit markup for the more elastic non-residential sector. That the two prices remain close despite this elasticity differential is a consequence of the shared scarcity rent dominating the markup difference at the parameter values estimated for Arizona utilities. That p_1^* and p_2^* differ at all is a result of the two-sector structure: a single-sector model would assign the same optimal price to all users, abstracting from the differential conservation incentives that sector-specific elasticities imply.

7.2 Observed Prices Relative to the Optimal Benchmark

The large majority of Arizona utilities price water below the cost-recovery benchmark, and virtually all price below the scarcity-inclusive optimum I derive from the dynamic program. Among the 241 non-AMA utility-years in the trimmed sample, 91 percent charge prices below average cost, confirming the systematic cost-recovery failures documented by McRae (2015) and Wichman (2024). The median observed price is \$3.47 per kgal against a median average cost of \$6.19, a shortfall of \$2.88 per kgal. These utilities face no formal groundwater constraint in the model and no CAP exposure, so the benchmark is cost recovery alone; their true optimal price would be higher still once groundwater depletion externalities are incorporated.

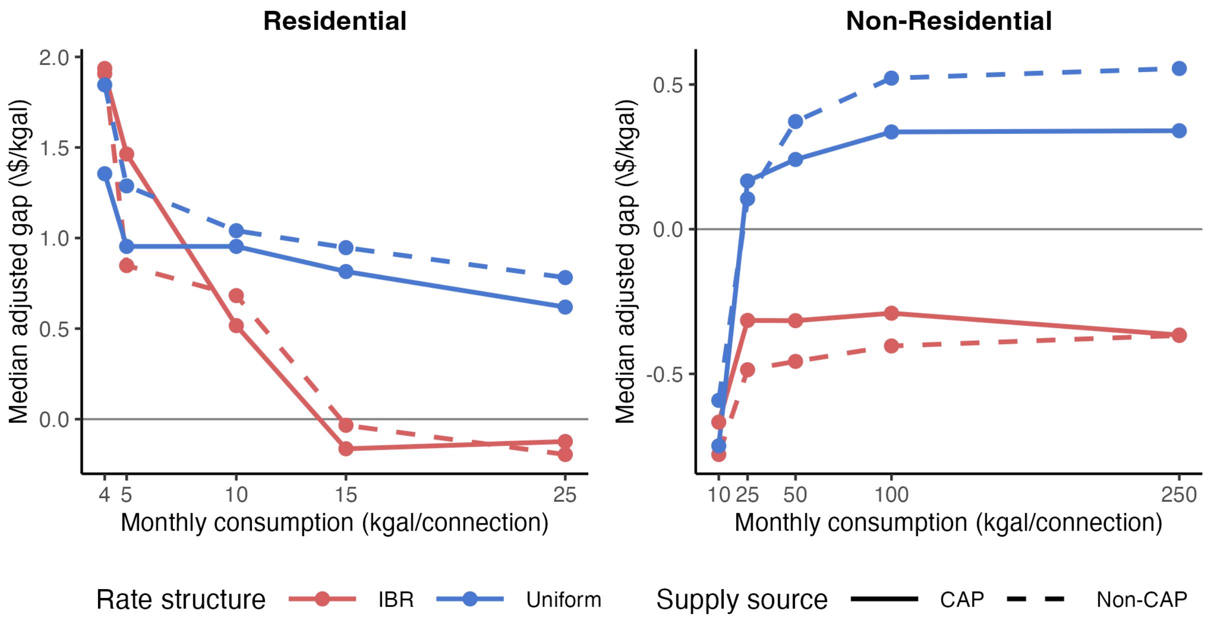
AMA utilities without CAP contracts price closer to average cost but still fall short. Among the 67 AMA non-CAP utility-years, 63 percent price below average cost, with a median gap of \$0.75 per kgal. The Arizona Department of Water Resources' 100-year assured water supply requirement disciplines investment decisions but does not directly bind annual rate-setting, which explains why regulation reduces but does not close the gap.

The 46 CAP utility-years, for which I compute the full scarcity-inclusive optimal from the dynamic program, reveal the additional cost of ignoring the resource constraint. The median observed price is \$3.06 per kgal against a median average cost of \$3.82 and a median model-implied optimum of \$4.26. Treating the observed price as a proxy for marginal cost, the \$1.20 per kgal gap between observed and optimal can be decomposed into a fixed cost-recovery component of \$0.76 and a scarcity rent component of \$0.44, where the latter equals the model-implied $\bar{\lambda}/\bar{\mu}$ evaluated at the observed water stock. This decomposition is exact within the theoretical structure of equation (17); in practice it is subject to estimation uncertainty in both the cost function and the VFI solution. Optimal price equals marginal cost plus the Ramsey markup plus the shadow value of water; the Ramsey markup is identified from the value function iteration (VFI) solution, and the residual is the scarcity rent. The typical CAP utility recovers its fixed costs only partially and embeds essentially none of the shadow value of CAP availability into observed prices.

7.3 Price Gaps by Consumption Tier and Rate Structure

Figure 3 plots the median adjusted price gap 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^* — distinct from the exact Ramsey markup in equation (17), but equal to it when observed prices approximate marginal cost — and subtracts it before comparing to the observed average marginal price, yielding a volumetrically comparable benchmark of approximately $MC +$ scarcity component. Full distributional detail appears in Tables 16 and 19 in the appendix; Tables 17 and 20 report the corresponding markup of observed prices relative to marginal cost.

Figure 3: Median Adjusted Price Gap by Consumption Level, Rate Structure, and CAP Status



Notes: Each line plots the median adjusted gap $= p^* - (AC - MC) -$ Avg. MP in \$/kgal across utility-years in the indicated group, where a positive value indicates underpricing relative to the scarcity-inclusive benchmark. $(AC - MC)$ proxies the fixed-cost recovery component of the optimal price; it equals the exact Ramsey markup from equation (17) when observed prices approximate marginal cost. Residential gaps (left) are evaluated at 4, 5, 10, 15, and 25 kgal per connection per month; non-residential gaps (right) at 10, 25, 50, 100, and 250 kgal per connection per month. Line coding: solid dark blue = uniform-rate CAP utilities; dashed light blue = uniform-rate non-CAP utilities; solid dark red = IBR CAP utilities; dashed light red = IBR non-CAP utilities. For non-CAP utilities the scarcity component is zero and the benchmark reduces to MC. Sample restricted to IBR and uniform-rate utilities with $AC \leq \$30$/kgal.

For uniform-rate utilities, the residential gap is positive and roughly flat across all consumption levels, declining only modestly from +\$1.36 per kgal at 4 kgal per month to +\$0.62 at 25 kgal per month. This flatness is mechanical: a uniform rate produces the same average marginal price at every consumption level, so the observed price falls short of the adjusted optimum by a roughly constant amount regardless of usage. IBR utilities exhibit a sharply different pattern. The median residential gap for IBR CAP utilities opens at +\$1.91 per kgal at 4 kgal per month and compresses to −\$0.16 by 15 kgal, a swing of more than \$2 per kgal. Negative gaps at upper tiers indicate that IBR upper-block prices exceed the adjusted optimum, effectively overpricing high-use residential consumption to cross-subsidize the heavily discounted first block. This pattern reveals a deliberate equity trade-off embedded in IBR design: low-volume residential users pay below the scarcity-inclusive price, while the conservation burden is concentrated at higher tiers.

The non-residential panel tells a different story. IBR utilities price non-residential customers above the adjusted optimum across all consumption levels: the median adjusted gap is negative at every tier (ranging from −\$0.67 per kgal at 10 kgal to −\$0.32 at 50 kgal), indicating that observed non-residential prices exceed the scarcity-inclusive benchmark after removing the fixed-cost recovery component. Rather than the tier-dependent compression seen in the residential panel, the non-residential overpricing for IBR utilities is broadly flat across consumption levels, suggesting that IBR schedules calibrated for residential conservation send a stronger-than-optimal volumetric signal to non-residential accounts. Uniform-rate utilities show a contrasting pattern: non-residential prices are above the adjusted optimum at low consumption (median gap −\$0.75 at 10 kgal) but fall below it at moderate and higher tiers (median gaps of +\$0.17 to +\$0.34 at 25 to 250 kgal). For uniform-rate utilities, the conservation shortfall for non-residential customers is concentrated at higher consumption levels, precisely where the higher price elasticity ($\hat{\varepsilon}_2 = -0.82$) implies the largest demand response to a price increase.

7.4 Welfare Simulation: Status Quo versus Optimal Pricing

The tier-table results establish that IBR utilities price high-use consumers closer to the scarcity-inclusive optimum on average, but they do not answer the welfare question: how much does persistent underpricing cost, and how do the costs differ across rate structure types? I address this with a 5-year forward simulation that computes the discounted expected utility gain from transitioning to model-optimal pricing for each of the 21 CAP utilities in the sample. Appendix Table 13 confirms that IBR and uniform utilities in this group are broadly comparable on observable characteristics: connections, costs, quantities, revenue, and county 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.

The simulation proceeds as follows. Starting from 20 percent of observed 2019 total demand as the initial water stock and 20 percent of observed 2019 revenue as the initial financial reserve, I draw sequences of CAP supply shocks $\{\theta_t\}_{t=1}^{5}$ from a calibrated Beta distribution and propagate the state forward under each pricing scenario. Under the status quo scenario, prices are held fixed at observed 2019 levels and the utility purchases CAP water according to the VFI policy function evaluated at the status quo state. Under the optimal scenario, I evaluate the VFI policy functions at each realized state (w_t, b_t) to obtain the period-t optimal prices and CAP purchase decision. The simulation is repeated for 500 shock sequences. I compute results 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.[14] I evaluate the welfare change as $V^*(w_t^{\text{OPT}}, b_t^{\text{OPT}}) - V^*(w_t^{\text{SQ}}, b_t^{\text{SQ}})$ at the post-transition state in each period, then divide by the number of connections to obtain a per-connection figure in dollars.

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 follows the 2019 observed M&I delivery rate of \$0.485 per kgal grown forward at 2.5 percent per year to account for structural cost inflation, yielding a path-independent schedule $p_x(t) = 0.485 \times 1.025^t$ for simulation years $t = 1, \ldots, 5$. 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 7 in the appendix. As that figure illustrates, 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 in year one. 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 reports median welfare gains per connection across all 21 CAP utilities. In mild supply

[14]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.

Table 6: Discounted Expected Utility Gain from Optimal Pricing

	Mild Scenario	Stressed Scenario
IBR ($n = 15$)	\$7.2 [\$-18.2, \$27.8]	\$7.2 [\$-5.3, \$44.6]
Uniform ($n = 6$)	\$3.2 [\$-0.7, \$23.1]	\$19.1 [\$0.1, \$32.3]

Notes: Entries are medians; interquartile range in brackets. Discounted expected utility gain per connection (\$/conn) from switching to optimal pricing, averaged across 500 Monte Carlo paths and 5 simulation years. 21 unique CAP utilities ($n = 15$ IBR, $n = 6$ uniform). Mild regime: mean CAP delivery $\approx$ 98% of contract, low variance. Stressed regime: mean delivery = 80% of contract, higher variance consistent with Tier 3 shortage conditions.

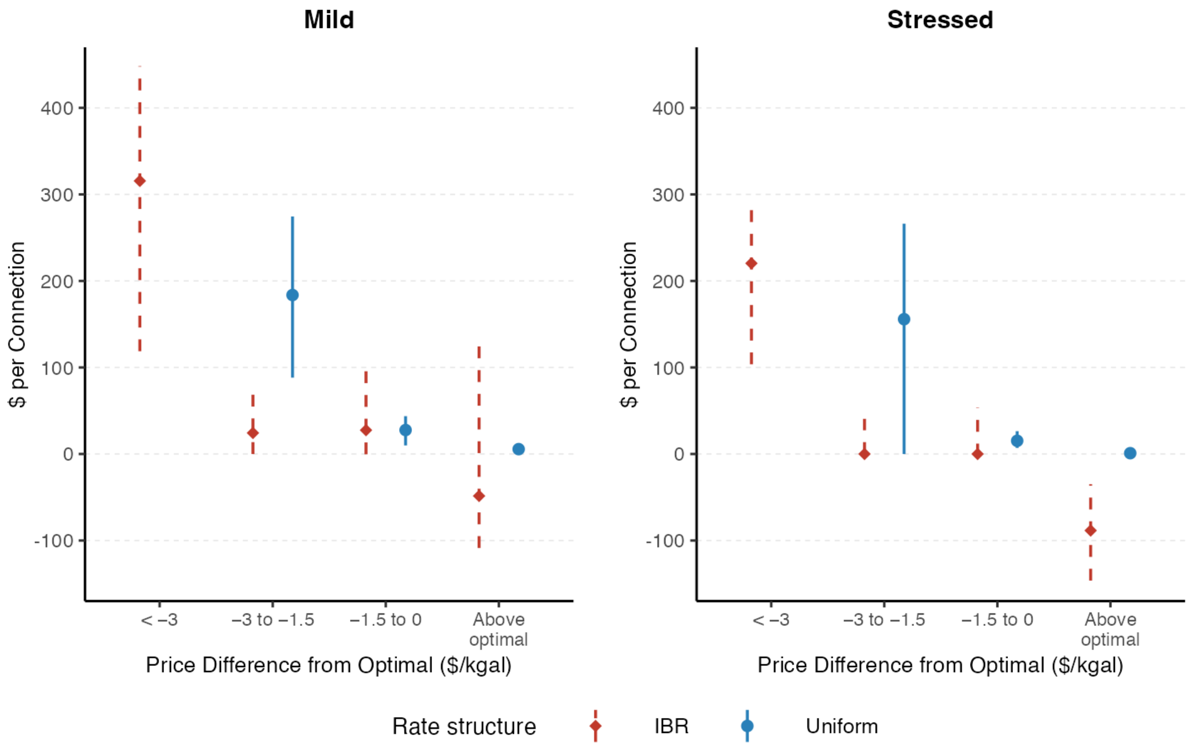
conditions, the IBR full-sample median is \$7.2 per connection against a uniform median of \$3.2; under stressed conditions the ordering reverses, with the uniform median rising to \$19.1 while the IBR median remains at \$7.2. Both comparisons, however, are shaped by a small number of severely underpriced IBR utilities whose gains are large by construction; Figure 4 disaggregates the distributions by the magnitude of each utility's observed price deviation, clarifying where the full-sample contrast originates.

Figure 4 organizes the results by how far each utility's observed residential price deviates from the model optimum. The x-axis divides the 21 utilities into four bins based on the signed gap $p_1^{\text{obs}} - p_1^*$: below $-\$3$, $-\$3$ to $-\$1.50$, $-\$1.50$ to \$0, and above the optimum. The y-axis reports the median discounted welfare gain per connection, with interquartile range bars; red diamonds represent IBR utilities and blue circles represent uniform-rate utilities. The left and right panels display the mild and stressed regimes respectively.

Under the mild regime, the welfare gains are broadly comparable across rate structure types for utilities whose observed prices fall within \$1.50 of the optimum. In the $-\$1.50$ to \$0 bin, IBR and uniform-rate utilities achieve nearly identical median gains of about \$28 per connection. In the $-\$3$ to $-\$1.50$ bin, by contrast, the two uniform-rate utilities (Paradise Valley and Oro Valley) achieve a median of \$184 per connection against \$24 for the three IBR utilities in that bin (Eloy, Surprise, and Sun City). Four IBR utilities fall in the above-optimal bin, meaning the model predicts their observed prices exceed the efficient level;[15] their median gain is $-\$49$, reflecting the cost of moderate overpricing in mild conditions.

[15] Tucson, Agua Fria, Mesa, and Vail.

Figure 4: Discounted Expected Welfare Gain from Optimal Pricing, by Price Deviation and Rate Structure



Notes: Each point plots the median discounted welfare gain per connection ($V^*(w_t^{\mathrm{OPT}}, b_t^{\mathrm{OPT}}) - V^*(w_t^{\mathrm{SQ}}, b_t^{\mathrm{SQ}})$, in $/connection) across 500 Monte Carlo paths and 5 simulation years, for utilities grouped by the signed gap $p_1^{\mathrm{obs}} - p_1^*$ (in $/kgal). Vertical bars show the interquartile range across simulation paths. Red diamonds: IBR utilities ($n = 15$ unique utilities). Blue circles: uniform-rate utilities ($n = 6$ unique utilities). Left panel: mild regime, $\theta \sim \mathrm{Beta}(50, 1)$, mean delivery $\approx 98\%$ of contract. Right panel: stressed regime, $\theta \sim \mathrm{Beta}(4, 1)$, mean delivery $= 80\%$ of contract. The $< -\$3$ bin for uniform-rate utilities is omitted because no uniform-rate utilities fall in that range.

The stressed regime reveals a sharper pattern. In the −\$1.50 to \$0 bin, IBR utilities' median gain falls to \$0 while uniform-rate utilities retain a median of \$15 per connection. In the −\$3 to −\$1.50 bin, the IBR median collapses from \$24 in the mild case to essentially \$0, while the uniform-rate median declines from \$184 to \$156. Across all moderate deviation bins, uniform-rate utilities preserve a meaningful positive welfare gain under stressed supply conditions while IBR utilities do not.

Three IBR utilities fall in the leftmost bin, with observed prices between \$4.80 and \$6.13 per kgal below the model optimum.[16] The resulting median welfare gain is \$316 per connection under mild conditions, driven by the severity of structural underpricing rather than rate structure type. These three utilities account for the elevated IBR median in Table 6 and motivate the trimmed comparison below. The three excluded utilities are not borderline cases: their price deviations of \$4.80 to \$6.13 per kgal are well in excess of the \$3 trimming threshold, and no utility in the sample falls within \$0.50 of the boundary, so the exclusion is not sensitive to the precise threshold chosen.

Table 7: Discounted Expected Utility Gain from Optimal Pricing (Excluding Severely Underpriced Utilities)

	Mild Scenario	Stressed Scenario
IBR ($n = 12$)	\$-1.8 [\$-22.4, \$24.0]	\$-1.6 [\$-9.1, \$30.0]
Uniform ($n = 6$)	\$3.2 [\$-0.7, \$23.1]	\$19.1 [\$0.1, \$32.3]

Notes: Entries are medians; interquartile range in brackets. Discounted expected utility gain per connection (\$/conn) from switching to optimal pricing, averaged across 500 Monte Carlo paths and 5 simulation years. 3 IBR utilities (Cave Creek, Florence, Goodyear) with observed prices more than \$3/kgal below optimal are excluded; their gains are driven by the severity of structural underpricing rather than rate structure per se. Uniform utilities unchanged ($n = 6$ unique utilities). Mild regime: mean CAP delivery $\approx$ 98% of contract, low variance. Stressed regime: mean delivery = 80% of contract, higher variance consistent with Tier 3 shortage conditions.

Restricting attention to utilities with price deviations greater than −\$3 per kgal, the sample that enables a cleaner comparison at more typical underpricing levels, reverses the ordering. The trimmed sample comprises 12 IBR and 6 uniform-rate utilities. Under the mild regime, the IBR median falls to −\$1.8 per connection against a uniform median of \$3.2. The divergence is more pronounced under the stressed regime: the IBR median remains near zero at −\$1.6 while the uni-

[16] Cave Creek, Florence, and Goodyear. Model-implied optimal prices range from \$7.04 (Florence) to \$9.38 (Cave Creek) against observed prices of \$2.19 to \$4.48.

form median rises to \$19.1. Appendix Table 22 repeats this comparison at a tighter $-\$2$ per kgal exclusion threshold, removing two additional IBR utilities whose observed prices fall between \$2 and \$3 below optimal. At this narrower trim the welfare cost of uniform underpricing grows larger while the IBR median turns more negative, reinforcing the main finding. Appendix Table 23 applies the $-\$4$ threshold and produces results identical to the $-\$3$ case, confirming that the three outlier IBR utilities drive all of the threshold sensitivity — once they are removed, tightening or loosening the trim has no effect.

The pattern that welfare gains are smaller under the stressed distribution than the mild distribution applies to both rate structure types and reflects the intertemporal mechanism through which optimal pricing generates value. Under mild supply conditions, a utility pricing at the model optimum can accumulate water stocks and financial reserves during favorable years, building the resource cushion that provides insurance against future shortfall. A utility on the status quo path, holding prices below the optimum, accumulates less stock and enters future periods with a weaker reserve position. The welfare gain from transitioning to optimal pricing captures this divergence through the dynamic program's value function. Although the simulation runs for five years, V^* is the solution to the infinite-horizon Bellman equation, so evaluating $V^*(w_t^{\mathrm{OPT}}, b_t^{\mathrm{OPT}}) - V^*(w_t^{\mathrm{SQ}}, b_t^{\mathrm{SQ}})$ at each period's realized state incorporates the full continuation value of accumulated stocks and reserves—including at the terminal period, where the state (w_5, b_5) is valued at $V^*(w_5, b_5)$ rather than discarded, with V^* capturing the continuation value of all future net benefits beyond the simulation window. Under the stressed distribution, CAP deliveries are persistently compressed and the opportunity for early-period stockpiling is limited, so the insurance mechanism operates under tighter constraints and the welfare gap between the two paths narrows.

The near-zero IBR welfare gain under the stressed regime also reflects how the simulation represents IBR pricing, and is itself a substantive finding. The welfare calculation uses the average marginal price at 10 kgal per connection per month as the single observed price for each IBR utility. When that rate already approximates the model-optimal level, a pattern the tier-table analysis in Section 7.3 confirms for moderate consumption levels, the aggregate gain from transitioning to the uniform-equivalent optimum is correspondingly small. IBR pricing, by concentrating higher prices at upper tiers, brings the marginal rate for moderate consumers closer to the scarcity-inclusive optimum than a flat uniform rate set far below the optimal level. What the simulation does not resolve is how the gains and losses from a transition to optimal uniform pricing would be distributed across the full tier schedule. Low-use residential customers, who pay a subsidized first-block price below the scarcity-inclusive optimum, would incur a consumer surplus loss under the transition; high-use customers, whose upper-block prices often exceed the optimal rate, would see their consumer surplus rise. Whether the net distributional effect favors or disfavors the transition depends on the number of customers at each tier, which is left for future research.

Two considerations bear 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 described above. Neither consideration alters the core finding: the figures in Table 7 measure the welfare gain from moving each utility's observed price to the scarcity-inclusive optimum, and uniform-rate utilities capture larger gains than IBR utilities at typical underpricing levels. The equity question of how those gains and losses are distributed across the customer base remains an important open question and a natural extension of this work.

7.5 Robustness

7.5.1 Observable Controls for the IBR–Uniform Comparison

A potential concern with comparing welfare gains across rate structure types is that IBR and uniform utilities may differ on observable characteristics that independently drive welfare outcomes. Appendix Tables 14 and 15 address this by regressing each utility's median welfare gain on the IBR indicator together with the observable characteristics described in Appendix Table 13. Table 14 presents the full sample of 21 utilities (Panels A and B) and Table 15 the trimmed sample of 18 excluding the three severely underpriced IBR utilities (Panels A and B), with all coefficients estimated by OLS and HC2-robust standard errors throughout.

In the full sample, the unconditional IBR coefficient is small and imprecise in both regimes (Column 1), reflecting the opposing forces that the three severely underpriced IBR utilities introduce: their large gains push the IBR mean upward, partially offsetting the lower gains of the remaining IBR utilities. Adding income and size controls (Column 2) tips the IBR coefficient slightly positive — county income enters significantly and positively, and once variation across utility markets is absorbed, the outliers' influence reverses the sign — before returning negative when cost structure and CAP exposure are controlled for (Columns 3–4). This sign instability in the full sample is itself informative: it confirms that the unconditional full-sample IBR average is driven by the outlier utilities rather than by the rate structure.

In the trimmed sample, the IBR coefficient is negative in every specification across both regimes. The unconditional gap is approximately $-\$54$ per connection, and it narrows to $-\$22$ to $-\$31$ after absorbing income, size, cost, and supply controls but never changes sign. Given $n = 18$, no individual coefficient reaches conventional significance, but the uniform direction across all eight trimmed specifications is consistent with the headline finding that the welfare cost of uniform-rate underpricing exceeds that of IBR underpricing at typical deviation levels, and that this pattern is not explained by observable differences between the two groups.

7.5.2 2019 Observed Initial States

The main simulation assigns each utility the same proportional starting point: 20 percent of observed 2019 total demand as the initial water stock and 20 percent of observed 2019 revenue as the initial financial reserve. This equal weighting reflects the absence of a single theoretically motivated baseline and ensures that no utility enters the simulation with an artificially favorable or unfavorable position. As a robustness check, I replace these proportional initializations with utility-specific 2019 observed values. Water stock is taken from each utility's reported long-term storage account (LTSA) balance at the end of 2019, converted from acre-feet to kgal; for utilities without an LTSA account the starting stock is zero. Financial reserve is set to each utility's operating surplus, defined as revenue minus total expenditure and clipped to zero for the four utilities with operating deficits. The resulting distribution of starting states is substantially more dispersed than the main simulation's proportional grid: the median 2019 water stock is 16.1 million kgal against a median annual demand of approximately 18 million kgal, ranging from zero for utilities without LTSA balances to 165.9 million kgal for the largest. The median financial reserve is \$3.5 million, ranging from zero to \$41.8 million.

Tables 8 and 9 report the welfare gains. The full-sample medians in Table 8 show that IBR and Uniform utilities achieve similar gains in mild conditions (\$63.4 and \$57.6 per connection, respectively), with IBR marginally higher. Under stress the ordering reverses: the Uniform median rises to \$37.8 against \$23.5 for IBR. As in the main simulation, the full-sample mild comparison is influenced by the three severely underpriced IBR utilities. Applying the same exclusion criterion removes Cave Creek, Florence, and Goodyear; the trimmed IBR median falls to \$9.5 per connection in the mild regime and \$2.5 under stress (Table 9), while the Uniform median remains at \$57.6 and \$37.8 — gaps of approximately \$48 and \$35 per connection, substantially wider than the main simulation's trimmed differential of \$3.2 against $-$\$1.8 in mild conditions and \$19.1 against $-$\$1.6 under stress. The qualitative finding is unchanged: Uniform utilities capture substantially larger welfare gains at typical underpricing levels, and the margin widens rather than narrows under realistic 2019 initial conditions.

Table 8: Discounted Expected Utility Gain from Optimal Pricing — 2019 Observed Initial States

	Mild Scenario	Stressed Scenario
IBR ($n = 15$)	\$63.4 [\$1.2, \$183.0]	\$23.5 [\$-12.5, \$98.4]
Uniform ($n = 6$)	\$57.6 [\$7.4, \$157.7]	\$37.8 [\$6.2, \$65.3]

Notes: Entries are medians; interquartile range in brackets. Discounted expected utility gain per connection (\$/conn) from switching to optimal pricing, simulated from 2019 observed initial water stock and financial reserve (revenue minus expenditure, clipped to zero for utilities with operating deficits). 21 unique CAP utilities ($n = 15$ IBR, $n = 6$ uniform). Mild regime: mean CAP delivery $\approx$ 98% of contract, low variance. Stressed regime: mean delivery $=$ 80% of contract, higher variance consistent with Tier 3 shortage conditions.

Table 9: Discounted Expected Utility Gain from Optimal Pricing — 2019 Observed Initial States (Excluding Severely Underpriced Utilities)

	Mild Scenario	Stressed Scenario
IBR ($n = 12$)	\$9.5 [\$-2.0, \$81.5]	\$2.5 [\$-33.9, \$58.8]
Uniform ($n = 6$)	\$57.6 [\$7.4, \$157.7]	\$37.8 [\$6.2, \$65.3]

Notes: Entries are medians; interquartile range in brackets. Discounted expected utility gain per connection (\$/conn) from switching to optimal pricing, simulated from 2019 observed initial water stock and financial reserve. 3 IBR utilities (Cave Creek, Florence, Goodyear) with observed prices more than \$3/kgal below optimal are excluded; their gains are driven by the severity of structural underpricing rather than rate structure per se. Uniform utilities unchanged ($n = 6$ unique utilities). Mild regime: mean CAP delivery $\approx$ 98% of contract, low variance. Stressed regime: mean delivery $=$ 80% of contract, higher variance consistent with Tier 3 shortage conditions.

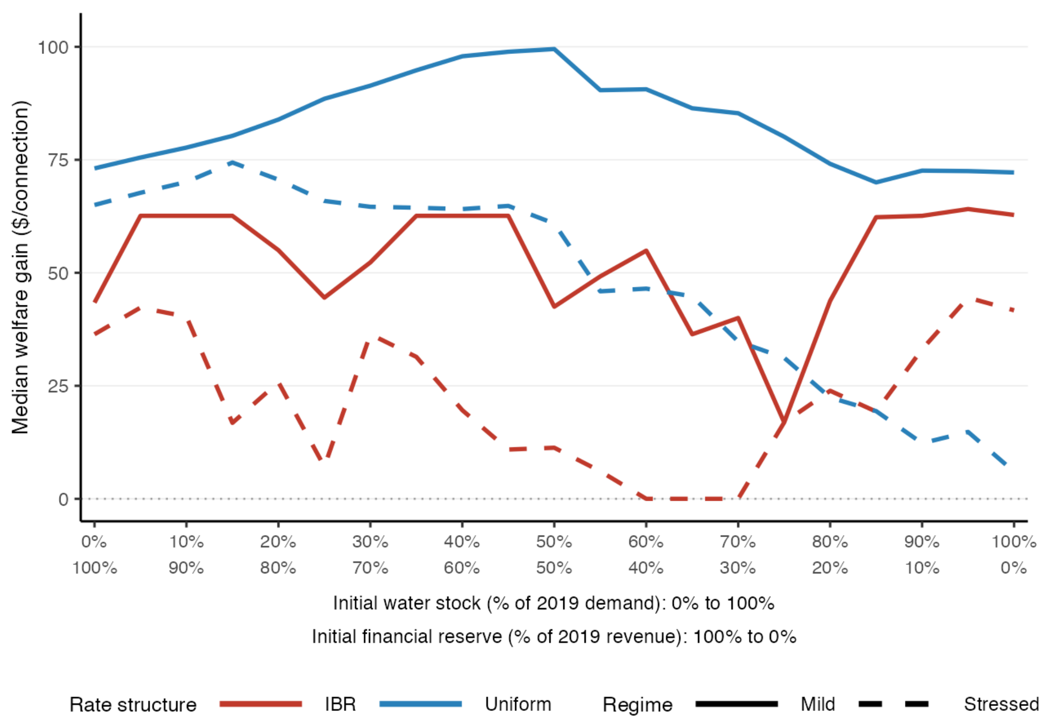
The wider welfare gap for Uniform utilities under 2019 observed starting states reflects the heterogeneity of LTSA balances across rate structure types. Several IBR utilities — including Phoenix, Mesa, Chandler, and Avondale — enter 2019 with large water stocks but no reported financial surplus, either because revenue data are unavailable or because expenditures equal or exceed revenues. Several others (EPCOR Paradise Valley, El Mirage, Vail, Sun City, and Cave Creek) hold no LTSA balance at all, beginning the simulation at zero water stock regardless of their financial position. The model's value function assigns the largest gains to utilities that begin near the policy function's most responsive region, which for many IBR utilities in 2019 corresponds to a high-stock, zero-reserve corner where the optimal and status quo paths diverge less sharply. The Uniform utilities in the sample begin at more evenly distributed positions along both state dimensions, where the gap between optimal and status quo trajectories is persistently larger.

7.5.3 Sliding Initial Stock and Reserve

To trace the sensitivity of the welfare ranking across the full space of possible initial conditions, I vary the starting state systematically along a one-dimensional grid. I define a parameter $\alpha \in \{0, 0.05, \ldots, 1.00\}$ and initialize each utility with water stock equal to α times its 2019 annual demand and financial reserve equal to $(1 - \alpha)$ times its 2019 revenue. As α rises, utilities begin the simulation with progressively more physical water but progressively less financial cushion. Figure 5 traces the cross-utility median welfare gain for each of the four rate structure–regime combinations across the 21 grid points.

Uniform utilities exhibit larger welfare gains from transitioning to optimal pricing than IBR utilities under mild supply conditions at nearly all initial states, with median gains peaking near \$100 per connection around $\alpha = 0.50$ before declining as the financial reserve is exhausted. These larger gains reflect the larger current cost of uniform underpricing: uniform-rate utilities price further below the scarcity-inclusive optimum, so the simulation's transition to optimal pricing delivers a correspondingly larger benefit. The stressed-regime comparison reveals a more structured pattern. Uniform utilities' stressed gains begin comparably to their mild gains when α is low (large financial reserve), but fall sharply above $\alpha = 0.50$ as the initial reserve diminishes, reaching \$6 per connection at $\alpha = 1.00$. IBR utilities' stressed gains follow the opposite trajectory: near zero or below for moderate α values between 0.55 and 0.70, they recover to \$41.7 per connection at $\alpha = 1.00$. Under stressed supply conditions with depleted financial reserves, the cost of IBR underpricing exceeds that of uniform underpricing, a reversal of the typical ordering. The mechanism is structural: IBR schedules automatically collect higher revenue from upper-tier consumers even at a given average

Figure 5: Welfare Gain Sensitivity to Initial State Allocation



Notes: Each line plots the cross-utility median discounted welfare gain per connection ($/connection) as a function of α, which indexes the initial-state trade-off: water stock $= \alpha \times$ 2019 annual demand; financial reserve $= (1 - \alpha) \times$ 2019 revenue. Moving left to right increases the initial water stock and decreases the initial financial reserve. Solid lines: mild regime ($\theta \sim \text{Beta}(50, 1)$, mean delivery $\approx$ 98% of contract). Dashed lines: stressed regime ($\theta \sim \text{Beta}(4, 1)$, mean delivery $=$ 80% of contract). Red: IBR utilities ($n = 15$). Blue: Uniform utilities ($n = 6$).

price, providing a budget buffer that a single uniform rate set below the optimum cannot replicate. Since the median 2019 CAP utility in the sample sits near $\alpha \approx 0.50$ on this grid — substantial water stocks, modest financial reserves — the main simulation and observed-initial-state results both fall in the region where the cost of uniform underpricing exceeds that of IBR, while the two structures impose broadly comparable welfare costs under stress. The full reversed-ranking region, $\alpha > 0.70$ under stress, corresponds to a financially depleted starting position that, while possible in prolonged shortage scenarios, is not representative of the sample in 2019.

7.5.4 Rate Structure Self-Selection and the Scope of the Welfare Comparison

The IBR–uniform welfare comparison is descriptive rather than causal. Rate structure adoption is self-selected: a utility chooses to implement IBR or maintain a uniform tariff, and that choice may reflect unobserved demand characteristics, political constraints, or regulatory history that independently influence optimal prices and welfare outcomes. I make no claim that switching from uniform to IBR pricing, or vice versa, would produce the welfare differences documented above. The comparison instead answers a narrower question: conditional on the observed pricing structure, how do the gains from transitioning to the model-optimal price differ across rate structure types?

Appendix Table 13 reveals that the 21 CAP utilities differ on two observable dimensions after conditioning on county income and system size: IBR utilities charge higher non-residential prices (median \$3.92 versus \$3.45 per kgal) and have changed their residential rates more frequently over the sample period. The observable-controls regression in Appendix Table 15 adds these characteristics to the IBR indicator and finds that the IBR coefficient is negative across all eight trimmed specifications — indicating lower welfare cost of IBR underpricing relative to the uniform baseline — and never reverses sign, providing evidence that the pattern is not driven by observable differences in non-residential pricing or rate-revision frequency alone.

The sliding-initial-state result offers the strongest available evidence that the comparison is not an artifact of the particular initialization. The pattern of larger welfare costs for uniform underpricing holds at nearly all 21 points on the α grid under mild supply, from the financially rich corner ($\alpha = 0$) through the midpoint approximating observed 2019 conditions ($\alpha \approx 0.20$) and well beyond. This robustness to initialization reinforces the interpretation that the welfare differential reflects the structural difference in how the two rate forms price incremental consumption relative to the scarcity-inclusive optimum, rather than a sensitivity of the simulation to the starting state. A causal interpretation would require quasi-experimental variation in rate structure adoption—such as a staggered rollout of IBR mandates or natural discontinuities in regulatory eligibility—which is left for future research.

8 Conclusion

This paper asks which rate structures best embed the scarcity rent of water into prices, and by how much Arizona utilities fall short of the Ramsey-optimal benchmark. I address both questions with three complementary components. First, I assemble what is, to my knowledge, the most comprehensive utility-level water pricing panel for any U.S. state, linking billing schedules, supply quantities, financial records, and CAP delivery data for 194 Arizona utilities over 2014–2019.[17] Second, I estimate residential and non-residential price elasticities of -0.59 and -0.82, using variation in neighboring utilities' prices as an instrument for average marginal price. 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. The results reveal systematic and substantial underpricing: 91 percent of non-AMA utility-years set prices below average cost, and the median CAP utility prices \$1.20 per kgal below the model optimum, with \$0.76 attributable to unrecovered fixed costs and \$0.44 to a scarcity rent that is essentially absent from observed rate schedules. Rate structure determines how this shortfall is distributed across consumption tiers. IBR utilities price moderate consumers substantially closer to the scarcity-inclusive optimum than their uniform-rate counterparts. This proximity means the welfare cost of IBR underpricing is small: the gap between the IBR status quo and the model-optimal path is narrow, leaving little value to recover from a transition to optimal pricing. By contrast, uniform-rate utilities price further below the optimum, so the welfare cost of their status quo is substantially larger. Among utilities at typical underpricing levels, the potential welfare gain from transitioning to optimal pricing is approximately zero for IBR utilities under the stressed scenario while the uniform-rate median reaches \$19.1 per connection (among the 6 uniform-rate CAP utilities in the trimmed sample), reflecting the larger and more persistent distance of a flat rate from the optimal.

[17] Wichman (2024) assembles a panel of similar scale for utilities in Georgia and North Carolina but does not observe quantities supplied directly, let alone broken out by customer group; instead, he infers approximate total quantity delivered from observed revenues and prices. I observe annual supply quantities by source type and by customer group (residential, non-residential) for each utility-year directly from ADWR public records.

The analysis proceeds at the utility level, and two limitations follow from this. The welfare figures measure the aggregate gain from transitioning from observed to optimal pricing at the average marginal rate. 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 extent of this tradeoff and the precise conservation vs 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 to this study and left for future work.

The findings carry a direct message for water regulators and utility managers. In most Arizona jurisdictions, rates are set through administrative proceedings that are infrequent and politically costly to revise. A utility that sets prices too low in one rate case may be unable to correct the error for several years. The choice of rate structure at the time of ratesetting therefore has lasting consequences for how well prices signal conservation, recover costs, and preserve the financial resilience needed to absorb supply shocks. IBR utilities in this sample provide a closer approximation of the scarcity-inclusive optimum than uniform-rate utilities. By concentrating higher prices at upper tiers, the IBR structure can partially embed a scarcity signal without requiring the overall rate level to reflect the full shadow cost; a uniform rate, once set below the optimum, provides no such tier-differentiated gradient. This proximity proves most consequential under stressed supply conditions, precisely when the scarcity rent is highest and the cost of underpricing is largest. As Colorado River shortage declarations deepen and CAP delivery uncertainty grows through the remainder of the decade, the rate structure chosen today will determine how much of that cost falls on utilities, ratepayers, and the water stock itself.

Two extensions follow directly from these findings. The most immediate is an analysis of the distributional consequences of rate reform. Quantifying the consumer surplus effects for low-, medium-, and high-use customers within each utility requires customer-level consumption data and would make it possible to evaluate the equity-efficiency tradeoff that IBR's tier structure implicitly makes. A second extension involves deriving the second-best IBR schedule, which would sharpen the welfare comparison between IBR and uniform pricing as institutional forms and provide regulators with a principled basis for tier design. Together, these extensions would translate the aggregate efficiency results here into distributional guidance that regulators can use in rate cases.

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.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6), 2121–68.

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

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.

Brent, Daniel A and Michael B Ward, "Price perceptions in water demand," *Journal of Environmental Economics and Management*, 2019, *98*, 102266.

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.

Chu, Long and R. Quentin Grafton, "Dynamic water pricing and the risk adjusted user cost (RAUC)," *Water Resources and Economics*, 2021, *35*, 100181.

Dalhuisen, Jasper M., Raymond J.G.M. Florax, Henri L.F. de Groot, and Peter Nijkamp, "Price and Income Elasticities of Residential Water Demand: A Meta-Analysis," *Land Economics*, 2003, *79* (2), 292–308.

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.

_ , _ , Jude Bayham, Whitney Boone, Erin M. K. Haacker, and Lisa Pfeiffer, "Measuring the value of groundwater and other forms of natural capital," *Proceedings of the National Academy of Sciences*, 2016, *113* (9), 2382–2387.

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.

Lewis, Daniel J. and Karel Mertens, "A Robust Test for Weak Instruments for 2SLS with Multiple Endogenous Regressors," *Review of Economic Studies*, 2025.

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.

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.

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

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.

A Demand Estimation Robustness

Table 10: Non-Residential Demand Elasticity Across Spatial Allocation Methods

Allocation method	N	$\hat{\varepsilon}_2$	(SE)	CD F
Area-weighted (baseline)	392	-0.820^{*}	(0.44)	50.3
Census block centroids	126	-0.990	—	—
Block group centroids	236	-0.090	—	—
Population $\times$ area	387	0.080	(0.836)	50.9

Notes: Each row reports the IV price elasticity of non-residential water demand from the preferred AMP (25–50 kgal tier) specification. The dependent variable is log non-residential water delivered per non-residential establishment. All specifications include log payroll per establishment, log county median household income, and county$\times$year fixed effects; instruments are lagged leave-one-out county-mean price and fixed fee. CD F: Cragg-Donald minimum eigenvalue statistic. Baseline (row 1) is the area-proportional ZCTA intersection method applied to ZIP Business Patterns data; the result matches Column (2) of Table 3. Census block centroid and block group centroid methods assign ZCTA establishments to utilities by matching census geography centroids to service area polygons; both reduce sample coverage substantially by excluding small and rural utilities. Population $\times$ area method weights each utility's share of a ZCTA's establishments proportionally to overlap area $\times$ service population, normalised to sum to one; the sample size is similar to the baseline but the elasticity flips sign (SE in parentheses; not significant). Significance: $^{*}p < 0.10$.

Table 11: Residential Water Demand: Thin-County Robustness (Counties with $\geq$ 5 Utilities)

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

Notes: Sample restricted to county-years (in the instrument period $t-1$) with at least 5 utilities reporting valid prices, so the leave-one-out instrument for each utility is constructed from at least 4 peers. Specifications mirror Table 2: Column (1) instruments for the average marginal price (AMP) at 10 kgal and the fixed fee; Column (2) uses the volumetric average price (AP). The dependent variable is ln(residential water per capita, kgal/yr). County×Year fixed effects included. Standard errors clustered by utility in parentheses. Years: 2014, 2015, 2017, 2019. Significance: ***p<0.01, **p<0.05, *p<0.10.

Table 12: Non-Residential Water Demand: Thin-County Robustness (Counties with ≥ 5 Utilities)

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

Notes: Sample restricted to county-years (in the instrument period $t-1$) with at least 5 utilities reporting valid non-residential prices, so the leave-one-out instrument is constructed from at least 4 peers. Specifications mirror Table 3: Columns (1)–(2) use average marginal price (AMP) in the 25–50 and 10–25 kgal/month tiers; Column (3) uses the volumetric average price (AP = $(\text{Bill}_{50} - \text{Fixed Fee})/50$). The dependent variable is ln(non-residential water per establishment, kgal/yr). County×Year fixed effects included. Standard errors clustered by utility in parentheses. Years: 2014, 2015, 2017, 2019. Significance: ***p<0.01, **p<0.05, *p<0.10.

B Rate Structure Balance and Welfare Controls

Table 13: Characteristics of CAP Utilities by Rate Structure

	IBR ($n = 60$)	Uniform ($n = 24$)	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 ($/kgal)	3.82 [3.43, 5.15]	3.66 [3.26, 4.04]	0.18
Marginal cost ($/kgal)	3.57 [3.20, 4.81]	3.41 [3.05, 3.77]	0.18
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)	2.9 [0.7, 14.2]	1.71 [0.77, 3.36]	0.05
Non-CAP supply (mil. kgal/yr)	2.8 [0.0, 11.7]	1.82 [0.99, 9.76]	0.94
LTSA water stock (mil. kgal, 2019)	17.2 [1.0, 38.6]	4.8 [0.4, 38.8]	0.67
Revenue ($M)	26.9 [6.8, 52.5]	18.8 [11.8, 36.5]	0.72
County median income ($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: Entries are medians; interquartile range in brackets. Sample comprises all utility-years for the 21 CAP utilities with converged VFI policies used in the welfare simulation (years 2014, 2015, 2017, and 2019; $n = 60$ IBR utility-years across 15 utilities, $n = 24$ Uniform utility-years across 6 utilities). IBR/Uniform classification is fixed at the 2019 value used in the simulation. Connections from EPA SDWIS. LTSA water stock is reported for 2019 only (the year from which stock data are drawn). Rate changes counts the number of survey years (2014, 2015, 2017, 2019) in which the utility reported a residential rate change since the prior survey; reported at the utility level ($n = 21$). p-values from two-sided Wilcoxon rank-sum tests.

Table 14: Welfare Gain and Rate Structure: OLS with Observable Controls, Full Sample ($n = 21$)

	(1)	(2)	(3)	(4)
Panel A: Mild regime				
IBR	1.5	28.3	-13.5	-24.9
	(59.4)	(40.9)	(37.8)	(45.9)
log(Connections)	—	-43.9*	11.8	17.9
		(23.1)	(20.1)	(21.9)
County income ($K)	—	1.8***	2.3***	2.0***
		(0.5)	(0.3)	(0.5)
CAP entitlement (mil. kgal)	—	—	-3.1	-2.3
			(1.9)	(2.0)
Average cost ($/kgal)	—	—	56.4***	39.3**
			(10.9)	(16.7)
Price deviation ($/kgal)	—	—	—	-28.9
				(25.0)
N	21	21	21	21
R^2	0.00	0.34	0.64	0.70
Panel B: Stressed regime				
IBR	-5.8	16.3	-21.8	-33.4
	(52.7)	(31.7)	(34.8)	(35.3)
log(Connections)	—	-37.2*	-8.8	-0.8
		(18.2)	(15.1)	(14.6)
County income ($K)	—	1.5***	1.8***	1.5***
		(0.4)	(0.4)	(0.5)
CAP entitlement (mil. kgal)	—	—	-0.6	0.1
			(1.2)	(1.3)
Average cost ($/kgal)	—	—	38.2***	19.7*
			(10.6)	(11.2)
Price deviation ($/kgal)	—	—	—	-30.6**
				(12.8)
N	21	21	21	21
R^2	0.00	0.40	0.66	0.77

Notes: Dependent variable is the cross-path median discounted welfare gain per connection ($/connection) for each utility under the indicated regime. *IBR* $= 1$ for increasing-block-rate utilities, 0 for uniform-rate. *Price deviation* $= p_1^{\text{obs}} - p_1^*$ (negative = underpriced; Column 4 adds price deviation to the full Column 3 specification). HC2-robust standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

Table 15: Welfare Gain and Rate Structure: OLS with Observable Controls, Trimmed Sample ($n = 18$)

	(1)	(2)	(3)	(4)
Panel A: Mild regime				
IBR	-55.2	-22.8	-22.1	-22.2
	(51.1)	(30.9)	(39.9)	(41.1)
log(Connections)	—	-19.3	-52.8	-48.5
		(14.1)	(41.4)	(39.8)
County income ($K)	—	1.9***	1.6***	1.6**
		(0.3)	(0.5)	(0.6)
CAP entitlement (mil. kgal)	—	—	2.9	2.8
			(3.5)	(3.6)
Average cost ($/kgal)	—	—	-34.0	-33.2
			(39.7)	(38.6)
Price deviation ($/kgal)	—	—	—	-11.3
				(26.3)
N	18	18	18	18
R^2	0.07	0.47	0.51	0.52
Panel B: Stressed regime				
IBR	-52.1	-26.1	-31.1	-31.2
	(45.8)	(21.4)	(21.6)	(20.2)
log(Connections)	—	-15.6	-65.1**	-60.2***
		(10.9)	(21.4)	(18.9)
County income ($K)	—	1.5**	1.1***	1.1***
		(0.6)	(0.4)	(0.3)
CAP entitlement (mil. kgal)	—	—	4.7**	4.7**
			(1.8)	(1.8)
Average cost ($/kgal)	—	—	-42.6**	-43.7***
			(16.1)	(12.8)
Price deviation ($/kgal)	—	—	—	-16.1
				(12.5)
N	18	18	18	18
R^2	0.11	0.57	0.70	0.73

Notes: Excludes Cave Creek, Florence, and Goodyear (observed prices more than $3/kgal below optimal), matching the trimmed sample in Table 7. Dependent variable is the cross-path median discounted welfare gain per connection ($/connection). $IBR = 1$ for increasing-block-rate utilities, 0 for uniform-rate. *Price deviation* $= p_1^{\text{obs}} - p_1^*$ (negative = underpriced; Column 4 adds price deviation to the full Column 3 specification). IBR coefficient is negative in every specification and never reverses sign. HC2-robust standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

C Price Gap Tables by Consumption Tier

Table 16: Residential Price Gap from Optimal, Adjusted for Two-Part Tariff Structure

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Avg. MP at 4 kgal/mo	1.45	1.05	0.17	1.36	2.81	15
Avg. MP at 5 kgal/mo	1.19	0.96	0.17	0.95	2.56	15
Avg. MP at 10 kgal/mo	0.82	1.14	-0.52	0.95	2.36	15
Avg. MP at 15 kgal/mo	0.74	1.14	-0.52	0.82	2.36	15
Avg. MP at 25 kgal/mo	0.59	1.07	-1.03	0.62	1.88	15
Panel B: *Uniform rates, non-CAP utilities*						
Avg. MP at 4 kgal/mo	2.74	3.55	-0.47	1.85	7.91	134
Avg. MP at 5 kgal/mo	2.22	3.34	-0.82	1.29	7.11	134
Avg. MP at 10 kgal/mo	2.02	3.33	-0.82	1.04	6.46	134
Avg. MP at 15 kgal/mo	1.92	3.34	-0.85	0.95	6.29	134
Avg. MP at 25 kgal/mo	1.84	3.33	-1.00	0.78	6.29	134
Panel C: *IBR rates, CAP utilities*						
Avg. MP at 4 kgal/mo	2.34	1.74	0.47	1.91	5.52	31
Avg. MP at 5 kgal/mo	1.78	2.04	-0.28	1.46	5.52	31
Avg. MP at 10 kgal/mo	1.29	1.94	-0.44	0.52	4.42	31
Avg. MP at 15 kgal/mo	0.06	2.33	-2.03	-0.16	2.80	31
Avg. MP at 25 kgal/mo	0.39	1.75	-1.48	-0.12	3.14	31
Panel D: *IBR rates, non-CAP utilities*						
Avg. MP at 4 kgal/mo	2.57	3.18	0.13	1.94	5.46	174
Avg. MP at 5 kgal/mo	1.57	3.16	-0.92	0.85	4.48	174
Avg. MP at 10 kgal/mo	1.35	3.14	-1.14	0.68	4.11	174
Avg. MP at 15 kgal/mo	0.20	3.42	-2.68	-0.03	2.55	174
Avg. MP at 25 kgal/mo	0.11	3.41	-2.79	-0.20	2.55	174

Notes: Adjusted gap $= p_1^* - (AC_{it} - MC_{it}) -$ Avg. MP_{it}, in \$/kgal. Avg. MP_{it} is the average marginal price for utility i in year t at the consumption level shown. MC_{it} is the utility-year specific marginal cost estimate from the log-linear cost regression (Section 6); AC_{it} is average cost. The model-optimal price p_1^* embeds a Ramsey markup $(AC_{it} - MC_{it})$ to recover fixed costs through the volumetric rate, because the model uses volumetric-only pricing. Observed utilities recover fixed costs through a fixed service charge, so their volumetric price is not expected to cover the Ramsey component. Subtracting the Ramsey markup from p_1^* yields the fair volumetric benchmark MC_{it} + scarcity component. Positive values indicate the observed price falls below the adjusted optimum (underpricing); negative values indicate overpricing. For non-CAP utilities (Panels B and D), the scarcity component is zero and the benchmark reduces to MC_{it}. Sample restricted to IBR and uniform-rate utilities with $AC_{it} \leq \$30$/kgal. Years: 2014, 2015, 2017, 2019. See Table 18 for the unadjusted gap (appendix).

Table 17: Residential Price Markup above Marginal Cost by Consumption Tier and Rate Structure

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Avg. MP at 4 kgal/mo	-0.39	0.30	-0.71	-0.35	-0.02	15
Avg. MP at 5 kgal/mo	-0.30	0.26	-0.70	-0.23	-0.00	15
Avg. MP at 10 kgal/mo	-0.18	0.35	-0.64	-0.17	0.19	15
Avg. MP at 15 kgal/mo	-0.16	0.36	-0.64	-0.16	0.19	15
Avg. MP at 25 kgal/mo	-0.11	0.32	-0.41	-0.13	0.35	15
Panel B: *Uniform rates, non-CAP utilities*						
Avg. MP at 4 kgal/mo	-0.26	0.90	-0.80	-0.49	0.20	134
Avg. MP at 5 kgal/mo	-0.11	1.10	-0.72	-0.35	0.46	134
Avg. MP at 10 kgal/mo	-0.06	1.08	-0.70	-0.30	0.38	134
Avg. MP at 15 kgal/mo	-0.04	1.08	-0.70	-0.27	0.46	134
Avg. MP at 25 kgal/mo	-0.02	1.09	-0.71	-0.25	0.52	134
Panel C: *IBR rates, CAP utilities*						
Avg. MP at 4 kgal/mo	-0.42	0.28	-0.72	-0.45	-0.13	31
Avg. MP at 5 kgal/mo	-0.27	0.35	-0.64	-0.36	0.21	31
Avg. MP at 10 kgal/mo	-0.15	0.31	-0.55	-0.07	0.21	31
Avg. MP at 15 kgal/mo	0.15	0.45	-0.40	0.14	0.58	31
Avg. MP at 25 kgal/mo	0.08	0.31	-0.40	0.13	0.45	31
Panel D: *IBR rates, non-CAP utilities*						
Avg. MP at 4 kgal/mo	-0.34	0.51	-0.70	-0.43	-0.03	174
Avg. MP at 5 kgal/mo	-0.10	0.69	-0.60	-0.19	0.30	174
Avg. MP at 10 kgal/mo	-0.05	0.69	-0.56	-0.15	0.37	174
Avg. MP at 15 kgal/mo	0.20	0.86	-0.46	0.01	0.88	174
Avg. MP at 25 kgal/mo	0.21	0.85	-0.45	0.05	0.88	174

Notes: Markup = $(\text{Avg. MP}_{it} - MC_{it})/MC_{it}$, where MC_{it} is the utility-year specific marginal cost estimate from the log-linear cost regression (Section 6). Negative values indicate pricing below marginal cost; positive values indicate a markup above variable cost. See notes to the corresponding price-gap table for sample and panel definitions.

Table 18: Residential Price Gap from Optimal by Consumption Tier and Rate Structure

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Avg. MP at 4 kgal/mo	1.69	1.07	0.39	1.61	3.08	15
Avg. MP at 5 kgal/mo	1.43	0.98	0.39	1.10	2.81	15
Avg. MP at 10 kgal/mo	1.05	1.17	-0.28	1.10	2.61	15
Avg. MP at 15 kgal/mo	0.98	1.16	-0.28	1.07	2.61	15
Avg. MP at 25 kgal/mo	0.82	1.10	-0.79	0.77	2.18	15
Panel B: *Uniform rates, non-CAP utilities*						
Avg. MP at 4 kgal/mo	4.48	5.03	0.35	2.89	11.80	134
Avg. MP at 5 kgal/mo	3.96	4.81	-0.19	2.50	10.99	134
Avg. MP at 10 kgal/mo	3.76	4.79	-0.25	2.34	10.31	134
Avg. MP at 15 kgal/mo	3.66	4.79	-0.48	2.30	9.82	134
Avg. MP at 25 kgal/mo	3.59	4.78	-0.58	2.30	9.82	134
Panel C: *IBR rates, CAP utilities*						
Avg. MP at 4 kgal/mo	2.64	1.83	0.84	2.22	6.00	31
Avg. MP at 5 kgal/mo	2.09	2.13	-0.03	1.77	6.00	31
Avg. MP at 10 kgal/mo	1.60	2.03	-0.07	0.77	5.04	31
Avg. MP at 15 kgal/mo	0.37	2.39	-1.69	0.07	3.42	31
Avg. MP at 25 kgal/mo	0.69	1.83	-1.20	0.11	3.76	31
Panel D: *IBR rates, non-CAP utilities*						
Avg. MP at 4 kgal/mo	4.42	4.65	0.77	3.21	9.23	174
Avg. MP at 5 kgal/mo	3.43	4.55	-0.03	2.08	8.12	174
Avg. MP at 10 kgal/mo	3.21	4.50	-0.23	1.83	7.73	174
Avg. MP at 15 kgal/mo	2.05	4.40	-1.56	1.12	6.48	174
Avg. MP at 25 kgal/mo	1.97	4.38	-1.56	1.01	6.26	174

Notes: Price gap $= p_1^* -$ Avg. MP_{it}, in \$/kgal. Avg. MP_{it} is the average marginal price for utility i in year t at the consumption level shown in each row. MC_{it} is the utility-year specific marginal cost estimate from the log-linear cost regression (Section 6). Positive values indicate the observed price falls below the scarcity-inclusive optimum (underpricing). For CAP utilities (Panels A and C), p_1^* is the solution to the stochastic dynamic program evaluated at the observed water stock. For non-CAP utilities (Panels B and D), $p_1^* = AC_{it}$, the average cost (cost-recovery lower bound). Sample restricted to IBR and uniform-rate utilities with $AC_{it} \leq \$30$/kgal. Years: 2014, 2015, 2017, 2019. See Table 16 for the two-part-tariff-adjusted gap (main results).

Table 19: Non-Residential Price Gap from Optimal, Adjusted for Two-Part Tariff Structure

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Avg. MP at 10 kgal/mo	-0.81	0.91	-2.08	-0.75	0.27	15
Avg. MP at 25 kgal/mo	0.09	0.75	-0.96	0.17	1.03	15
Avg. MP at 50 kgal/mo	0.36	0.81	-0.81	0.24	1.44	15
Avg. MP at 100 kgal/mo	0.49	0.86	-0.59	0.34	1.66	15
Avg. MP at 250 kgal/mo	0.52	0.86	-0.44	0.34	1.70	15
Panel B: *Uniform rates, non-CAP utilities*						
Avg. MP at 10 kgal/mo	0.08	3.57	-2.51	-0.59	4.66	134
Avg. MP at 25 kgal/mo	1.14	3.32	-1.72	0.10	5.58	134
Avg. MP at 50 kgal/mo	1.43	3.34	-1.40	0.37	5.93	134
Avg. MP at 100 kgal/mo	1.52	3.38	-1.35	0.52	6.11	134
Avg. MP at 250 kgal/mo	1.50	3.40	-1.40	0.56	6.17	134
Panel C: *IBR rates, CAP utilities*						
Avg. MP at 10 kgal/mo	-0.48	1.18	-1.84	-0.67	1.38	31
Avg. MP at 25 kgal/mo	0.06	1.35	-0.96	-0.32	2.12	31
Avg. MP at 50 kgal/mo	0.03	1.27	-1.21	-0.32	1.71	31
Avg. MP at 100 kgal/mo	-0.11	1.05	-1.36	-0.29	1.64	31
Avg. MP at 250 kgal/mo	-0.34	0.96	-1.43	-0.37	1.24	31
Panel D: *IBR rates, non-CAP utilities*						
Avg. MP at 10 kgal/mo	-0.41	3.12	-3.09	-0.78	1.90	174
Avg. MP at 25 kgal/mo	-0.10	3.19	-2.66	-0.49	2.18	174
Avg. MP at 50 kgal/mo	-0.06	3.32	-2.72	-0.46	2.35	174
Avg. MP at 100 kgal/mo	-0.06	3.34	-2.83	-0.40	2.43	174
Avg. MP at 250 kgal/mo	-0.16	3.37	-2.92	-0.37	2.30	174

Notes: Adjusted gap $= p_2^* - (AC_{it} - MC_{it}) -$ Avg. MP_{it}, in \$/kgal. Avg. MP_{it} is the average marginal price for utility i in year t at the consumption level shown. MC_{it} is the utility-year specific marginal cost estimate from the log-linear cost regression (Section 6); AC_{it} is average cost. The model-optimal price p_2^* embeds a Ramsey markup $(AC_{it} - MC_{it})$ to recover fixed costs through the volumetric rate, because the model uses volumetric-only pricing. Observed utilities recover fixed costs through a fixed service charge, so their volumetric price is not expected to cover the Ramsey component. Subtracting the Ramsey markup from p_2^* yields the fair volumetric benchmark MC_{it} + scarcity component. Positive values indicate the observed price falls below the adjusted optimum (underpricing); negative values indicate overpricing. For non-CAP utilities (Panels B and D), the scarcity component is zero and the benchmark reduces to MC_{it}. Sample restricted to IBR and uniform-rate utilities with $AC_{it} \leq \$30$/kgal. Years: 2014, 2015, 2017, 2019. See Table 21 for the unadjusted gap (appendix).

Table 20: Non-Residential Price Markup above Marginal Cost by Consumption Tier and Rate Structure

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Avg. MP at 10 kgal/mo	0.35	0.32	0.09	0.27	0.83	15
Avg. MP at 25 kgal/mo	0.07	0.27	-0.16	0.02	0.54	15
Avg. MP at 50 kgal/mo	-0.01	0.28	-0.27	-0.06	0.49	15
Avg. MP at 100 kgal/mo	-0.04	0.30	-0.33	-0.09	0.42	15
Avg. MP at 250 kgal/mo	-0.05	0.30	-0.35	-0.09	0.38	15
Panel B: *Uniform rates, non-CAP utilities*						
Avg. MP at 10 kgal/mo	0.60	1.87	-0.50	0.15	1.26	134
Avg. MP at 25 kgal/mo	0.23	1.38	-0.59	-0.04	0.79	134
Avg. MP at 50 kgal/mo	0.13	1.24	-0.65	-0.09	0.71	134
Avg. MP at 100 kgal/mo	0.10	1.17	-0.69	-0.14	0.70	134
Avg. MP at 250 kgal/mo	0.11	1.18	-0.69	-0.14	0.68	134
Panel C: *IBR rates, CAP utilities*						
Avg. MP at 10 kgal/mo	0.28	0.37	-0.22	0.28	0.75	31
Avg. MP at 25 kgal/mo	0.16	0.31	-0.32	0.20	0.58	31
Avg. MP at 50 kgal/mo	0.16	0.31	-0.30	0.20	0.47	31
Avg. MP at 100 kgal/mo	0.18	0.30	-0.19	0.20	0.52	31
Avg. MP at 250 kgal/mo	0.22	0.30	-0.05	0.25	0.54	31
Panel D: *IBR rates, non-CAP utilities*						
Avg. MP at 10 kgal/mo	0.40	1.00	-0.25	0.20	1.01	174
Avg. MP at 25 kgal/mo	0.28	0.90	-0.33	0.12	0.81	174
Avg. MP at 50 kgal/mo	0.27	0.91	-0.36	0.12	0.88	174
Avg. MP at 100 kgal/mo	0.26	0.92	-0.38	0.10	0.92	174
Avg. MP at 250 kgal/mo	0.28	0.94	-0.38	0.11	0.97	174

Notes: Markup $=$ (Avg. $\text{MP}_{it} - MC_{it})/MC_{it}$, where MC_{it} is the utility-year specific marginal cost estimate from the log-linear cost regression (Section 6). Negative values indicate pricing below marginal cost; positive values indicate a markup above variable cost. See notes to the corresponding price-gap table for sample and panel definitions.

Table 21: Non-Residential Price Gap from Optimal by Consumption Tier and Rate Structure

	Mean	SD	10th %ile	50th %ile	90th %ile	n
Panel A: *Uniform rates, CAP utilities*						
Avg. MP at 10 kgal/mo	-0.57	0.93	-1.84	-0.54	0.53	15
Avg. MP at 25 kgal/mo	0.33	0.78	-0.74	0.35	1.35	15
Avg. MP at 50 kgal/mo	0.60	0.84	-0.59	0.40	1.77	15
Avg. MP at 100 kgal/mo	0.73	0.89	-0.35	0.56	1.97	15
Avg. MP at 250 kgal/mo	0.76	0.89	-0.21	0.59	2.00	15
Panel B: *Uniform rates, non-CAP utilities*						
Avg. MP at 10 kgal/mo	1.82	4.90	-2.03	0.45	8.29	134
Avg. MP at 25 kgal/mo	2.88	4.75	-1.07	1.53	9.03	134
Avg. MP at 50 kgal/mo	3.17	4.78	-0.90	1.83	9.40	134
Avg. MP at 100 kgal/mo	3.26	4.81	-0.88	2.02	9.61	134
Avg. MP at 250 kgal/mo	3.24	4.83	-0.87	2.01	9.72	134
Panel C: *IBR rates, CAP utilities*						
Avg. MP at 10 kgal/mo	-0.18	1.24	-1.45	-0.39	1.71	31
Avg. MP at 25 kgal/mo	0.36	1.43	-0.68	-0.10	2.61	31
Avg. MP at 50 kgal/mo	0.34	1.34	-0.98	-0.08	2.16	31
Avg. MP at 100 kgal/mo	0.19	1.11	-1.04	-0.03	1.89	31
Avg. MP at 250 kgal/mo	-0.03	0.97	-1.16	-0.01	1.50	31
Panel D: *IBR rates, non-CAP utilities*						
Avg. MP at 10 kgal/mo	1.44	4.38	-1.92	0.43	4.99	174
Avg. MP at 25 kgal/mo	1.76	4.31	-1.68	0.65	5.68	174
Avg. MP at 50 kgal/mo	1.79	4.36	-1.59	0.67	5.93	174
Avg. MP at 100 kgal/mo	1.79	4.34	-1.65	0.78	6.00	174
Avg. MP at 250 kgal/mo	1.70	4.32	-1.82	0.71	5.70	174

Notes: Price gap $= p_2^* - \text{Avg. MP}_{it}$, in \$/kgal. Avg. MP_{it} is the average marginal price for utility i in year t at the consumption level shown in each row. MC_{it} is the utility-year specific marginal cost estimate from the log-linear cost regression (Section 6). Positive values indicate the observed price falls below the scarcity-inclusive optimum (underpricing). For CAP utilities (Panels A and C), p_2^* is the solution to the stochastic dynamic program evaluated at the observed water stock. For non-CAP utilities (Panels B and D), $p_2^* = AC_{it}$, the average cost (cost-recovery lower bound). Sample restricted to IBR and uniform-rate utilities with $AC_{it} \leq \$30$/kgal. Years: 2014, 2015, 2017, 2019. See Table 19 for the two-part-tariff-adjusted gap (main results).

D Welfare Simulation Robustness

Table 22: Discounted Expected Welfare Gain from Optimal Pricing, Trimmed at $-\$2$ per kgal

	Mild Regime ($/connection)	Stressed Regime ($/connection)
IBR ($n = 10$)	$-13.4 [$-28.2, $7.0]	$-2.1 [$-11.2, $5.0]
Uniform ($n = 5$)	$2.1 [$-1.6, $4.4]	$7.7 [$-2.5, $30.5]

Notes: Entries are medians; interquartile range in brackets. Utilities with observed residential price more than $2 per kgal below the model-implied optimum are excluded. Mild regime: $\theta \sim \text{Beta}(50, 1)$, mean delivery $\approx$ 98% of contract. Stressed regime: $\theta \sim \text{Beta}(4, 1)$, mean delivery $=$ 80% of contract. $n =$ unique CAP utilities.

Table 23: Discounted Expected Welfare Gain from Optimal Pricing, Trimmed at $-\$4$ per kgal

	Mild Regime ($/connection)	Stressed Regime ($/connection)
IBR ($n = 12$)	$-1.8 [$-22.4, $24.0]	$-1.6 [$-9.1, $30.0]
Uniform ($n = 6$)	$3.2 [$-0.7, $23.1]	$19.1 [$0.1, $32.3]

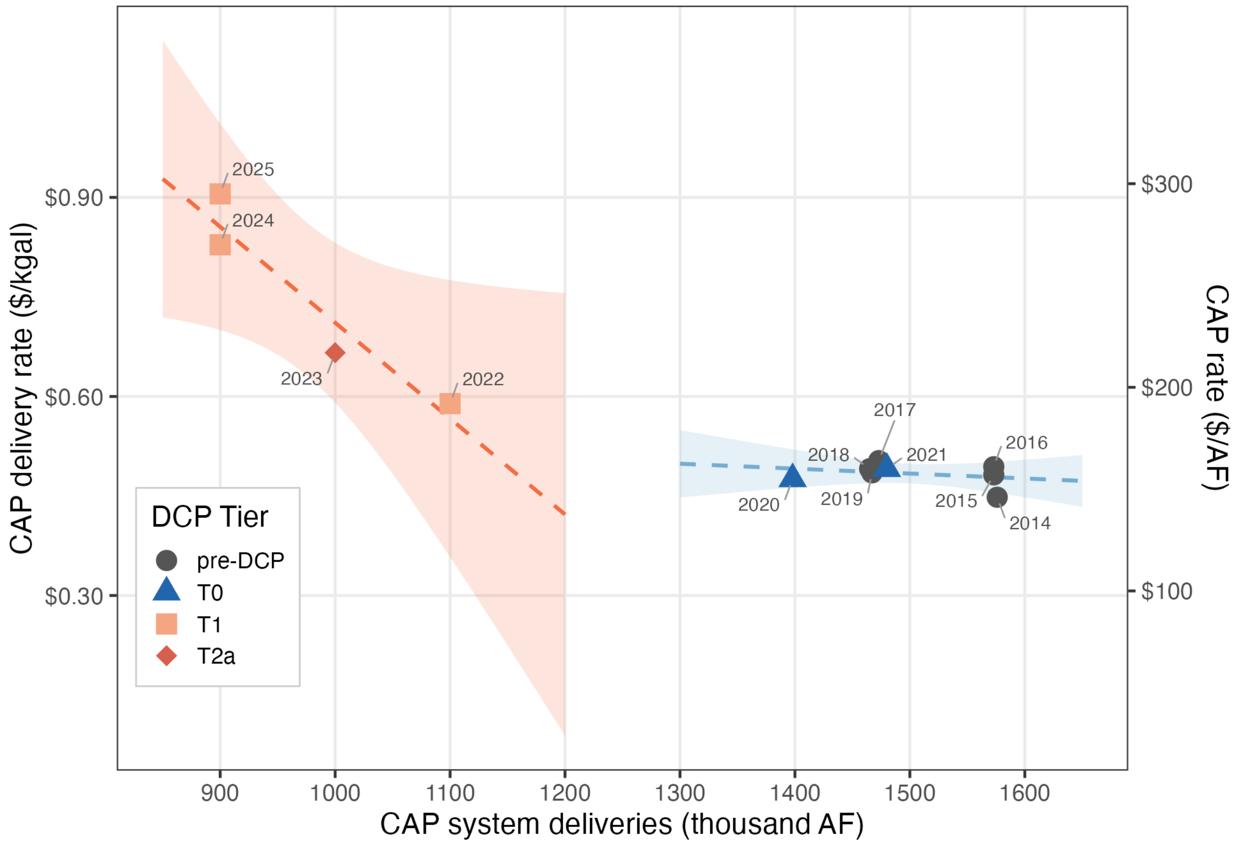
Notes: Entries are medians; interquartile range in brackets. Utilities with observed residential price more than $4 per kgal below the model-implied optimum are excluded. Mild regime: $\theta \sim \text{Beta}(50, 1)$, mean delivery $\approx$ 98% of contract. Stressed regime: $\theta \sim \text{Beta}(4, 1)$, mean delivery $=$ 80% of contract. $n =$ unique CAP utilities.

E Additional Figures

Figure 6: 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 7: 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: 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.