

Regulation Under Asymmetric Information in Water Utilities*

Isabelle Brocas

University of Southern California

Department of Economics, Los Angeles, CA 90089-0253

Phone: (213) 740-8842, Fax: (213) 740-8543, email: brocas@usc.edu

Kitty Chan

Brattle Group

1133 Twentieth Street NW, Washington, DC 20036

Phone: (202) 955-5050 , Fax: (202) 955-5059 , email: kitty.chan@brattle.com

Isabelle Perrigne (corresponding author)

Pennsylvania State University

Department of Economics, University Park, PA 16802

Phone: (814) 863-2157, Fax: (814) 863-4775 , email: perrigne@psu.edu

SESSION TITLE: RESTRUCTURING AND REGULATION OF PUBLIC UTILITIES

Regulation Under Asymmetric Information in Water Utilities

by I. Brocas, K. Chan and I. Perrigne*

Water utilities are reminiscent of network industries and are characterized by important fixed costs. These factors contribute to a single firm serving an area justifying public intervention on pricing. About one fourth of U.S. water utilities is private and subject to regulation. Regulators are unlikely to be perfectly informed and regulation is unlikely to be costlessly implemented. These inherent imperfections have lead economists to consider the incentive properties of regulatory procedures using the economics of information. See David Baron (1989). The empirical literature on regulation has focused on evaluating the effects of regulation on prices, firms' costs, efficiency and innovation in sectors such as airline, electricity and energy as surveyed by Paul Joskow and Nancy Rose (1989). Few of these empirical studies rely on the so-called theory of regulation. Regarding the water industry, there is an abundant literature on residential water demand, firms' cost and their efficiency given their public versus private nature. Relying on a model with asymmetric information and a sample of California water utilities, Frank Wolak (1994) assesses the consumer welfare loss due to asymmetric information and shows that the model with asymmetric information provides a superior description of the cost and demand data than the model under perfect information. Analyzing pricing for residential water is an important policy issue as the sector experienced recently price increases. The problem is even more acute in California because of a high residential demand for water coupled with population growth, water scarcity and the probability of severe droughts.

Relying on a new data set of 32 districts in California over 1995-2000, we analyze regulation of private water utilities. For every district, the California Public Utilities Commission (CPUC) chooses a price for water, an access fee per meter and a rate of return on capital to satisfy firm's revenue requirements. We assume that the CPUC is imperfectly informed about firms' labor efficiency. Following David Besanko (1984) and Wolak (1994), we develop a model in which the firm's capital is used as a screening variable. In particular, the model has the features of a rate-of-return regulation. We show how the rate of return and the access fee can be determined optimally to control firms' rents. We then adopt a structural

approach to analyze the data. A multistep estimator allows us to estimate the key parameters of the model. The empirical results show price inelasticity, an income effect, slightly decreasing returns to scale and a concentration of efficient firms. The computation of the optimal rate of return and access fee shows that the CPUC would tend to be cautious by allowing a lower than optimal rate and access fee. Relying on the estimated parameters, a first experiment evaluates the cost of asymmetric information. The price would be significantly lower resulting in a gain of consumer surplus. A second experiment consists in simulating the outcome of an optimal price cap following Farid Gasmi, Mark Kennet, Jean-Jacques Laffont and William Sharkey (2002) model. Price cap has become a popular regulatory tool in the eighties such as for electricity though the incentives resulting in price cap regulation have been questioned by economists. The counterfactual simulations show a price increase, which results in a significant loss in consumer surplus. The increase in firms' profit does not, however, counterbalance this loss supporting the relevance of the actual mechanism.

A first section briefly presents the data and the regulatory process. A second section introduces the model and the econometric modeling, while a third section is devoted to the empirical results and policy simulations.

I. Data

Class A utilities are serving more than 10,000 residential connections and are required to submit annual reports on their capital stock including water sources and operating costs. The rate cases provide information on water price, meter price or access fee, rate of return as well as the interest rate on investment paid by the utility. See Wolak (1994) for the regulatory process used by the CPUC. The CPUC is in charge of protecting the consumers' interests by challenging any claim made by utilities through the general rate cases. The CPUC also exercises some control over the utility's capital stock. The important role played by the capital suggests that the regulator uses the utility's capital stock as a screening variable in the spirit of Besanko (1984) model. There is an important heterogeneity in the quality of labor across the districts. It is expected that the utilities are better informed about the labor quality than the regulator. Annual reports do not provide complete information on labor costs. Out of the 58 districts regulated by the CPUC, 32 districts are served by the

California Water Company and the Southern California Water Company. Because additional information could be obtained from both companies, we focus on these 32 districts from 1995 to 2000 making a total of 192 observations. Data on income per capita and rainfall have been collected from the Bureau of Economic Analysis, Department of Commerce and the Water Resources of California, respectively. All the data have been deflated and are expressed in 1995 dollars. Table 1 provides summary statistics on the key variables. The data present an important variability over districts. An hedonic price model shows that the price of water tends to be larger in Southern California and rural areas than in Northern California and urban areas. The former can be explained by the scarcity of water sources, while the latter is due to additional capital in pipes. A comparison of capital stock with variable cost shows a large capital requirement. Energy expenses represent an important proportion of operating costs because of water pumping. The energy price is measured as the ratio of the energy expenses per foot of pipe. The rate of return, which varies from 0.089 and 0.113, is always larger than the price of capital, which varies from 0.079 and 0.103. Districts have applied for rate cases on average 5.47 times, i.e. almost once a year, suggesting frequent revisions of prices and rates of return by the CPUC.

II. The Model and Econometric Modeling

We consider a model in the spirit of Baron and Roger Myerson (1982) and Besanko (1984). The plant is privately informed about its labor efficiency or type θ distributed as $F(\cdot)$ on $[\underline{\theta}, \bar{\theta}]$, where $\underline{\theta}$ ($\bar{\theta}$) denotes the most (least) efficient plant. The plant chooses a level of capital K that is observable and can be contracted upon. Investing K costs δK . The cost is $C(\theta, K, q, \epsilon_c)$ where q is the quantity demanded by consumers and ϵ_c is a stochastic shock. It is increasing in θ and q and decreasing in K . The demand $q(p, \epsilon_d)$ is subject to a stochastic shock ϵ_d . The regulator offers a scheme $[p(K), T(K), R(K)]$ where $p(K)$ is the price per unit, $T(K)$ is the total access fee and $R(K)$ is the rate-of-return on capital. Then, the plant's expected payoff is $U = p(K)E[q(p(K), \epsilon_d)] - E[C(\theta, K, q(p(K), \epsilon_d), \epsilon_c)] - \delta K + T(K)$ where $E(\cdot)$ denotes expectations with respect to ϵ_d and ϵ_c . The rate-of-return satisfies the revenue requirement $E p(K)[q(p(K), \epsilon_d)] - E[C(\theta, K, q(p(K), \epsilon_d), \epsilon_c)] + T(K) = R(K)K$. Given the scheme, a plant with type θ chooses a level of capital $K(\theta)$ and in equilibrium, it is allowed to charge

$p(K(\theta)) \equiv p(\theta)$ and $T(K(\theta)) \equiv T(\theta)$ and is entitled to $R(K(\theta)) \equiv R(\theta)$. Because the CPUC protects consumers' interests, the regulator's objective is to maximize consumers' surplus while satisfying the revenue requirement and anticipating the plant's strategic response.

Proposition 1: *The rate-of-return regulation yields $K^*(\theta)$, $p^*(\theta)$, $T^*(\theta)$ and $R^*(\theta)$ satisfying*

$$\begin{aligned}
p^*(\theta)E[q_p(p^*(\theta), \epsilon_d)] &= E[C_q(\theta, K^*(\theta), q(p^*(\theta), \epsilon_d), \epsilon_c)q_p(p^*(\theta), \epsilon_d)] \\
&\quad + E[C_{\theta q}(\theta, K^*(\theta), q(p^*(\theta), \epsilon_d), \epsilon_c)q_p(p^*(\theta), \epsilon_d)]\frac{F(\theta)}{f(\theta)}, \\
\delta &= -E[C_K(\theta, K^*(\theta), q(p^*(\theta), \epsilon_d), \epsilon_c)] \\
&\quad - E[C_{\theta K}(\theta, K^*(\theta), q(p^*(\theta), \epsilon_d), \epsilon_c)]\frac{F(\theta)}{f(\theta)}, \\
T^*(\theta) &= -E[q(p^*(\theta), \epsilon_d)]p^*(\theta) + E[C(\theta, K^*(\theta), q(p^*(\theta), \epsilon_d), \epsilon_c)] \\
&\quad + R^*(\theta)K^*(\theta), \\
R^*(\theta)K^*(\theta) &= \delta K^*(\theta) + \int_{\theta}^{\bar{\theta}} E[C_{\theta}(s, K^*(s), q(p^*(s), \epsilon_d), \epsilon_c)]ds + [R^*(\bar{\theta}) - \delta]K^*(\bar{\theta}),
\end{aligned}$$

where $R^*(\bar{\theta})$ is a constant left to the discretion of the regulator.¹

In the presence of asymmetric information, the regulator must grant rents to induce the plant to make optimal decisions. The revenue requirement can be rewritten as $U = [R(\theta) - \delta]K(\theta)$, and therefore the rate-of-return determines by which amount total revenues must exceed cost and guarantees each plant is provided correct incentives. Overall, the rate-of-return regulation is the optimal second-best policy. Note, however, that the policy does not include a non-negative profit requirement and the regulator can choose arbitrarily $R^*(\bar{\theta})$ so that the least efficient plants make losses.

Given its popularity over the past twenty years, we also consider the price cap regulation as in Gasmi et al. (2002). In that case, the regulator sets a price cap \bar{p} and the plant is free to choose K as well as any price below \bar{p} . In the optimal price cap, the most efficient plants charge their monopoly prices and only less efficient plants are constrained by the cap.

Proposition 2: *In the pure price cap regulation,*

(i) *there exists $\hat{\theta}$ such that $\forall \theta < \hat{\theta}$, the plant charges $p^M(\theta)$ and invests $K^M(\theta)$ where*

$$p^M(\theta) E[q_p(p^M(\theta), \epsilon_d)] + E[q(p^M(\theta), \epsilon_d)] = E[C_q(\theta, K^M(\theta), q(p^M(\theta), \epsilon_d), \epsilon_c)q_p(p^M(\theta), \epsilon_d)]$$

$$\delta = -E[C_K(\theta, K^M(\theta), q(p^M(\theta), \epsilon_d), \epsilon_c)]$$

and $\forall \theta \geq \hat{\theta}$, the plant charges \bar{p} and invests $\bar{K}(\theta, \bar{p})$ where $\delta = -E[C_K(\theta, \bar{K}(\theta, \bar{p}), q(\bar{p}, \epsilon_d), \epsilon_c)]$,

(ii) \bar{p} is such that $E[q(\bar{p}, \epsilon_d)]\bar{p} - E[C(\bar{\theta}, \bar{K}(\bar{\theta}, \bar{p}), q(\bar{p}, \epsilon_d), \epsilon_c)] - \delta\bar{K}(\bar{\theta}, \bar{p}) = 0$,

(iii) $\hat{\theta}$ is such that $p^M(\hat{\theta}) = \bar{p}$.

We adopt a structural approach to estimate the model, i.e. the econometric model is directly derived from Proposition 1. The error terms are given by the random shocks ϵ_d and ϵ_c and θ , which can be interpreted as a term of unobserved heterogeneity. We follow Perrigne (2002) identification strategy and estimation procedure. The basic idea is to parameterize the model structure $[q(\cdot), C(\cdot), F(\cdot)]$, while minimizing the assumptions on ϵ_d and ϵ_c and exploiting the independence of ϵ_d and θ . We consider a demand function with constant price elasticity, $q(p, \epsilon_d) = \exp(d_0)\mathbf{Z}_d^{d_1}p^{d_2}\exp(\epsilon_d)$, where \mathbf{Z}_d is a vector of exogenous variables. We consider a Cobb-Douglas technology, in which the adverse selection variable affects labor efficiency. This gives the variable cost function $C(\theta, q, K, \epsilon_c) = \exp(\beta_0)\exp(\beta_L\theta)p_L^{\beta_L}p_E^{\beta_E}q^{\beta_y}K^{-\beta_K}\mathbf{Z}_c^{\beta_c}\exp(\epsilon_c)$, where p_E and p_L denote the price for energy and labor, respectively and \mathbf{Z}_c is a vector of exogenous variables. Homogeneity of degree 1 can be imposed. Regarding the firms' type density, we choose a Gamma density for its flexibility, i.e. $f(\theta; r, \gamma) = \gamma(\gamma\theta)^{r-1}\exp(-\gamma\theta)/\Gamma(r)$, where r is a positive integer and $\Gamma(r) = \int_0^\infty x^{r-1}\exp(-x)dx$.

Given the multiplicative random shocks, it is natural to assume $E[\exp(\epsilon_d)|\mathbf{Z}] = 1$ and $E[\exp(\epsilon_c)|\mathbf{Z}] = 1$, where \mathbf{Z} is a vector of exogenous variables. We do not make any distributional assumption on ϵ_d and ϵ_c . Identification of the model relies on the independence of θ and ϵ_d conditionally on \mathbf{Z} , i.e. $\theta \perp \epsilon_d|\mathbf{Z}$. In addition to the demand and cost, the price and capital equations in Proposition 1 define the econometric model.² Because $p(\theta)$ and $K(\theta)$ can create a problem of endogeneity, we need to solve the system in p and K given θ . Given the potential measurement error on capital, we prefer to use the price equation. Thus, the econometric model is made of the demand, cost and price equations while the error terms are ϵ_d , ϵ_c and θ . It gives

$$q_i = \exp(d_0)\mathbf{Z}_{di}^{d_1}p_i^{d_2}\exp(\epsilon_{di}) \quad (D)$$

$$C_i = \exp(\beta_0)p_{Li}^{\beta_L}p_{Ei}^{\beta_E}q_i^{\beta_y}K_i^{-\beta_K}\mathbf{Z}_{ci}^{\beta_c}\exp(\beta_L\theta_i)\exp(\epsilon_{ci}), \quad (C)$$

$$\log p_i = \frac{1}{(d_2 + 1)(\beta_K + 1) - d_2\beta_y} \left\{ \eta + \beta_L \log p_{Li} + \beta_E \log p_{Ei} + \beta_c \log \mathbf{Z}_{ci} \right.$$

$$+(d_1\beta_y - d_1\beta_K - d_1) \log \mathbf{Z}_{di} + \beta_K \log \delta_i + \xi(\theta_i; r, \gamma) \Big\}, \quad (P)$$

$i = 1, \dots, 192$, where $\eta = -\beta_K \log \beta_K + (\beta_K + 1) \log \beta_y + \beta_0 + d_0(\beta_y - \beta_K - 1) + \log E_P$, $E_P = E[\exp(\beta_y \epsilon_d + \epsilon_c)]$, and $\xi(\theta) = \beta_L \theta + \log[1 + \beta_L(F(\theta; r, \gamma)/f(\theta; r, \gamma))]$. Note that p_i in (D) is not endogenous because $p(\theta_i)$ and $\theta_i \perp \epsilon_{di} | \mathbf{Z}$.

The parameters to be estimated are $(d_0, d_1, d_2, \beta_0, \beta_L, \beta_E, \beta_K, \beta_c, r, \gamma)$, while the observables are $(q_i, \mathbf{Z}_{di}, C_i, p_{Li}, p_{Ei}, \delta_i, \mathbf{Z}_{ci}, i = 1, \dots, 192)$. The estimation method is multistep. Using $E(\exp(\epsilon_{di}) | \mathbf{Z}_i) = 1$, the first step estimates d_0, d_1 and d_2 using a Nonlinear GMM estimator. Using the orthogonality condition $\theta_i \perp \epsilon_{di} | \mathbf{Z}_i$, the second step estimates $\beta_L, \beta_E, \beta_c$ and β_K using a Linear GMM estimator. The third step estimates $\beta_0 + \log E_P$ and (r, γ) . Note that (P) provides $\log p_i = \psi(\theta_i; \beta_0 + \log E_P, r, \gamma)$, since the other terms in (P) are either observed or estimated. For $r = 1, 2, \dots$, a Method of Simulated Moments is used to estimate $(\beta_0 + \log E_P, \gamma)$. The computation of additional moments allows us to assess the best adjustment for r . Using these estimates, we can recover $\theta_i, i = 1, \dots, 192$ from (P).³

III. Estimation Results and Policy Experiments

Estimation results are displayed in Table 2. The demand parameters have the expected magnitude. The demand is quite inelastic with a price elasticity at -0.29 as found in previous studies. There is a significant revenue effect with a revenue elasticity at 0.52. Households living in Northern California tend to consume less water than in Southern California. The ratio population by pipe footage is used as a proxy for population density. Thus, larger population density areas tend to consume less water with an elasticity equal to -0.20. The production process tends to exhibit decreasing returns to scale though they could be considered as constant. Since homogeneity is imposed, the parameter for capital is equal to 0.3249. Rainfall has a negative impact on costs as districts with more rainfall are expected to have water sources in proximity thereby reducing costs. The case $r = 1$ provides the best fit with γ equal to 1.04 giving $E(\theta) = 0.9617$ suggesting a concentration of efficient firms. These results can be used to assess whether the regulator has implemented the optimal regulation through T and R . After elementary algebra, $R^*(\theta) = \delta + [(\int_{\bar{\theta}}^{\theta} E[C_\theta(s, K^*(s), q(p^*(s), \epsilon_d), \epsilon_c)] ds + U(\bar{\theta})) / K^*(\theta)]$, where $U(\bar{\theta})$ is the profit value for the least efficient firm and $T^*(\theta)$ is given in Proposition 1. The optimal monthly access fee and the optimal rate of return should be equal on average to

\$19 and 0.14, respectively. Both values are somewhat larger than observed ones (see Table 1) with a larger range of values leading to more discrimination among firms. These results suggest that (i) the model provides a reasonable fit to the data, and (ii) the CPUC tends to be cautious in the rents given to firms.

Under complete information, the price would be about 10.7% less on average than the observed one resulting in an increase of 3% in water consumption. The monthly access fee would be also at about \$4. These two factors lead to an increase in consumer surplus by 6.7% while considering a maximum consumers' willingness to pay for water at \$10. The capital stock would be about 16.9% lower, while the firms' profit would be equal to zero by definition. We note the parallel with the Averch-Johnson effect. Overcapitalization follows asymmetric information and allows firms to self-select in the desired way. As such, the cost of asymmetric information is quite substantial. The simulation of a price cap under asymmetric information leads to a uniform water price at about \$2.8. All the firms would be subject to the cap as their monopoly prices would take larger values than \bar{p} . This larger price would result in a reduction of expected consumption by 30% and a reduction in consumer surplus by 23%. The capital stock chosen by firms would be about 35% less as it would be not be subject to any distortion. This would result in a substantial increase in firms' profit. When considering the social welfare defined as the sum of the consumer surplus and the firms' profit, this increase in firms' profit does not, however, counterbalance the loss in consumer surplus. As such, our results suggest that the actual regulation provides a superior outcome.

References

- Baron, David P.** “Design of Regulatory Mechanisms and Institutions,” in Richard Schmalensee and Robert D. Willig, eds., *Handbook of industrial organization, volume II*, Amsterdam: North Holland, 1989, pp. 1347-1447.
- Baron, David P. and Myerson, Roger B.** “Regulating a Monopolist with Unknown Costs.” *Econometrica*, 1982, 50 (4), pp. 911-930.
- Besanko, David.** “On the Use of Revenue Requirements Regulation Under Imperfect Information,” in Michael A. Crew, ed., *Analyzing the impact of regulatory change in public utilities*, Lexington: Lexington Books, 1984, pp. 39-55.
- Gasmi, Farid; Kennet, Mark, Laffont, Jean-Jacques and Sharkey, William.** *Cost proxy models and telecommunications policy: A new empirical approach to regulation*, Cambridge: MIT Press, 2002.
- Joskow, Paul L. and Rose, Nancy L.** “The Effects of Economic Regulation,” in Richard Schmalensee and Robert D. Willig, eds., *Handbook of industrial organization, volume II*, Amsterdam: North Holland, 1989, pp. 1450-1506.
- Perrigne, Isabelle.** “Incentive Regulatory Contracts in Public Transportation: An Empirical Study.” Working paper, Pennsylvania State University, 2002.
- Wolak, Frank A.** “An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction.” *Annales d’Economie et de Statistique*, 1994, 34, pp. 13-69.

*Brocas: Department of Economics, University of Southern California, Los Angeles, CA 90089-0253; Chan: Brattle Group, 1133 Twentieth Street NW, Washington, DC 20036; Perrigne: Department of Economics, Pennsylvania State University, University Park, PA 16802. Financial support from the National Science Foundation, Grant SES-0452154, is gratefully acknowledged by the last author. We thank the California Public Utility Commission, the California Water Service Company and the Southern California Water Company for letting us access to their data.

¹ The lower index refers to the partial derivative of the function.

² Note that (p, q, C, K, T, R) are endogeneously determined by the model, while the econometric model contains only three error terms leading to a singular model. We consider only the three equations determining (q, p, C) .

³ This multi-step estimation procedure does not provide valid standard errors except for the first step. Bootstrap methods can be used to compute standard errors.

TABLE 1 – SUMMARY STATISTICS

Variable	Mean	STD
Population	68,635.87	76,712.70
Income per Capita	19,216.99	7,589.44
Rainfall (in inches)	40.00	23.80
Water delivered (in hundred cubic feet)	5,073,353.52	5,629,467.36
Water Price	1.17	0.62
Meter Price	9.88	5.24
Capital Stock	22,401,770.35	21,957,186.28
Footage of Pipeline	1,111,151.71	1,068,542.94
Variable Cost	5,548,544.59	6,952,976.35
Labor Price	25,472.15	11,922.65
Energy Price	0.51	0.35
Capital Price	0.09	0.006
Rate of Return	0.10	0.006

TABLE 2 – ESTIMATION RESULTS

Variable	Coefficient	<i>t</i> -ratio
Constant	-1.5738	-1.66
Water Price	-0.2946	-4.96
Population	1.0285	37.77
Income	0.5155	5.60
North	-0.2014	-3.49
Population/Pipe	-0.2039	-3.18
Labor Price	0.1975	2.63
Energy Price	0.4776	3.38
Water	1.2590	4.78
Rainfall	-0.0028	-0.01
γ ($r = 1$)	1.0434	0.35