Identification and Estimation of Incentive Contracts under Asymmetric Information: An Application to the French Water Sector *

Christian Bontemps† Max Lesellier ‡ David Martimort §

Incomplete version - October 2022.

Abstract

We develop a Principal-Agent model to represent management contracting for public-service delivery. A firm (the Agent) has private knowledge of its marginal cost of production. The local public authority (the Principal) cares both about the consumers' net surplus from consuming the services and the (weighted) firm's profit. Contractual negotiation is modeled as the choice by the privately informed firm within a menu of options determining both the unit-price charged to consumers and the fixed fee. Our theoretical model characterizes optimal contracting in this environment. We then explicitly study the nonparametric identification of the model and perform a semi-parametric estimation on a dataset coming from the 2004 wave of a survey from the French environment Institute (IFEN, Institut Français de l'Environnement).

Keywords: Principal-Agent, optimal contracts, structural model, nonparametric estimation, Instrumental Variable Quantile Regression.

JEL codes: C12, C15, D82.

*This paper is a complete redraft of an original version that circulated under the title "Estimating Optimal Contracts under Delegated Management, Application to the French Water Industry". We thank for useful comments Steve Berry, Aviv Nevo, Bernard SalaniÃ© and participants at various conferences and seminars.

†ENAC & Toulouse School of Economics. christian.bontemps@tse-fr.eu

‡Toulouse School of Economics. max.lesellier@tse-fr.eu

§Toulouse School of Economics. david.martimort@tse-fr.eu

1 Introduction

The increasing participation of the private sector in public-service delivery is often motivated by the need to expand access to services, increase or update existing delivery networks, and operate public utilities more efficiently. Many services provided by public utilities being associated with health, environmental or household income considerations (gas, electricity, water, transportation), there is a broad consensus on the need for public regulation of these utilities. In particular, industries such as water, gas and electricity usually meet the conditions for a local natural monopoly (large fixed costs and constant or declining marginal cost), so that protecting consumers from large price increases is often advocated as the main reason behind public regulation of utilities in these industries.

Private-sector participation in public utilities may take very different forms: private ownership of networks and facilities, centralized or local regulation by a public or independent authority, or contracting-out utility operation to private companies. In the latter case, typical arrangements are lease or concession contracts which can be renegotiated over time between a local community and a private company in charge of operating the public utility. Although contracting-out seems an interesting way of promoting public-private partnerships for public utilities because it combines flexibility with legal commitment, it can deteriorate consumer welfare if the arrangement concerning utility pricing is not carefully specified. A major reason behind the difficulty to design an optimal pricing rule for the utility is the fact that common information on the operator's ability to manage the utility efficiently is rare. For example, the operator's technical know-how and expertise may not correspond exactly to the actual state of the facilities (based on past maintenance). Such a situation of asymmetric information on the operator's efficiency in a contract-based relationship has been studied extensively in the literature on incentives and regulation (Salanié 2005, Laffont and Martimort 2002, Laffont and Tirole 1994).

In the standard theory of contracts, agents (operators in our case) are indexed by a private-information parameter which ultimately determines their actions, within a contract-based relationship with a "principal" (the local community). Whether this parameter denotes an unobservable action (moral hazard) or an unknown characteristic of the agent (adverse selection), the principal is assumed to have prior information used to design an optimal contract (in most cases, maximizing social welfare). The range of industries in which production, marketing and regulation activities are subject to contract-based relationships between economic agents is sufficiently large to guarantee an increasing number of empirical applications for such a theory.

The first approach in the literature on econometric estimation of delegated management models with asymmetric information was considering reduced-form models, or structural versions with restrictive parametric restrictions on the technology and the distribution of the private parameter. Examples of this first generation of models with asymmetric information include Wolak (1994), Ivaldi and Martimort (1994), Thomas (1995), Lavergne and Thomas (2005), Gagnepain and Ivaldi (2002), Brocas, Chan and Perrigne (2006). See Chiappori and Salanié (2000) for references on reduced-form estimation of models with asymmetric information, and Lavergne and Thomas (2005) for a survey on structural and reduced-form models. Most empirical applications confirm the fact that neglecting asymmetric information in the estimation of structural models yields biased estimates of, e.g., marginal cost or consumer price elasticity. However, it is also true that specification crucially matters for this type of structural models, as misspecification is likely to affect estimates of agents' preferences or production technology as much as neglecting asymmetric information altogether.

The second approach is more recent in the literature and proposes a way round this problem, in a series of articles based on nonparametric approaches. The move from parametric to nonparametric methods for estimating structural models with asymmetric information was following the development of structural models of auctions, with which they share some common features, as well as the literature on nonparametric identification. (d'Haultfoeuille and Fevrier, 2020, Luo et al. (2018) among others).

Structural models of delegated management with asymmetric information share common features with models of auctions or nonlinear pricing. The structural equations are nonlinear in an unobserved heterogeneity component (the private-information parameter) whose distribution is explicitly part of the solution to the underlying economic model. Moreover, endogenous selection of companies may also occur because of individual-rationality or incentive-compatibility constraints, in a way similar to auction participants or consumers in nonlinear pricing problems. For instance, Luo et al. (2018) propose a nonparametric identification method for models of nonlinear pricing, whose estimating equations closely resemble the ones associated with structural models of management delegation.

We propose in this paper a new methodology for estimating optimal contracts under asymmetric information, which relaxes most assumptions considered in the literature. More precisely, we rely extensively on non parametric techniques to derive the distribution of the private information type and functional estimates of the model. The paper discusses the non parametric identification of the structural model under several assumptions regarding unobserved heterogeneity. The estimation method is applied to the case of delegated management of water utilities in France. The advantage of considering such industry for our empirical application lies in the fact that contracts between local communities and private firms for the operation of water utilities in France exist for a long time, under various modes that make the application of the theory of contracts particularly relevant. In the case we consider, the local community manager has incomplete information on the efficiency of the (private) operator of the water utility, before signing the delegation contract. The optimal contract is a second-best solution which depends on marginal social surplus for water, marginal cost of water supply, as well as on the distribution of the private type. A particular aspect of the model is the existence of two sources of unobserved heterogeneity (to the econometrician). Beside the usual private parameter, the econometrician does not observed an heterogeneity term associated with the social surplus and which is community-specific.

The paper is organized as follows. Section 2 presents the French water sector. Section 3 presents a theoretical principal-agent model of contract-based regulation between the local community and the operator of the water utility in a context of information asymmetry. This model is tailored to the specificities of the sector under scrutiny. The system of equations describing the optimal contract are first derived. Those equations summarizes the optimality of behavior of the municipality and the operator. They serve as the basis for the estimation procedure. In Section 4, we discuss the conditions for nonparametric identification of the structural model. Two different cases are considered, depending on assumptions made on unobserved heterogeneity. Section 6 introduces the empirical application to the French water utilities with a particular focus on residential water pricing rules. Some counterfactuals are presented in Section 7. Section 8 discusses alternative modeling of optimal contracting relationships and illustrates some specificities of our approach. Finally, Section 9 concludes.

2 The French Water Sector

2.1 Governance

As in other countries throughout the world, the provision of water services in France is a regulated activity although it has always been fully decentralized at the local level since the 1789 French Revolution. The need to regulation comes from the fact that each stare network is indeed a public monopoly, with high fixed costs and low variable costs, as well as a declining (long-run) average cost curve. In each urban area of significant size equipped with such a water network, a local authority (a single city or a group of cities) has full responsibility to contract with an operator for providing water to the corresponding population.

In terms of size, water utilities represent no less than 1% of French national GDP. Yet, modes of governance significantly differ across networks. Water utilities in charge of supply, distribution and sewage activities may indeed be public (the so-called *"régies municipales"* or private (*"gestation déléguée"*).[1] Delegation of utility operation to a private company is sometimes viewed as improving efficiency although, empirical studies are not really conclusive as to the relative efficiency of a management mode or the other.[2]

When the service is delegated to the private sector, various kinds of contracts (*"concession"*, *"affermage"*, *"gérance"*..) rule details of this public-private partnership related to water production and distribution, maintenance, and quality supervision.[3] For so-called *régies intéressées*, the operator does not own the network and is paid as a fraction of the benefits of the service that accrues to the municipality. For a contract d'*"affermage"* , the operator is again paid directly by consumers but the cost of maintaining and renewing assets remains borne by the municipality. If the contract is a lease contracts *"concession"*, the operator must cover the costs of building and maintaining the infrastructure which are given back to the municipality at the end of the contract. The operator is directly paid through consumers' bills (the so-called *"redevance"*). By 2007, although private arrangements only concerned 39 % of water services, its share presented 72 % of the overall population. By 2008, 28% of networks were under public management. At the same time, the private sector is highly concentrated with three large companies sharing the market: Veolia Eau France, 39%, Lyonnaise des Eaux, 19 %, Saur, 11% and others companies or joint agreements among big ones amounting to less than 3 %. Veolia is the leading firm in the private sector. The smallest company Saur is more present in rural areas. It is also interesting to notice that public management is more frequent in small municipalities.

Given the high concentration in the sector, practitioners, medias and even politicians have sometimes complained that the market may not be so competitive after all, leaving inexpert municipalities in a weak bargaining position in front of big private players and subject to corruptive behavior. As a result, the *Loi Sapin* was enacted in 1993 to improve transparence and competitive biding while the 1995 *Loi Mazeaud* was designed to improve control on the operator. The competitive bidding generally relies on a two-step procedure. First the public authority chooses an operator through tender, whose criteria need not be publicized. Then a winner is selected after negotiation within those operators having made the best offers.

[1]As a matter of comparison, ownership is mostly public in Germany and Italy, while it remains mixed in Spain. In Great Britain, water utilities are always run by private companies under the aegis of an independent national regulator (OFWAT). In the U.S., water utilities operate under rate-of-return regulation. The fact that State commission for public utilities may differ in their mandates gives rise to fairly heterogeneous rates of return.

[2]See Bhattacharyya et al. (1995), Bhattacharyya et al. (1994) and Estache and Rossi (1999) among others.

[3]In France, the public authority itself is in charge of verifying the quality of the service.

2.2 Pricing

Importantly, heterogeneity is a key aspect of the sector both on the demand and the cost sides. Networks are disconnected from one municipality to the other and vary significantly in terms of length, age, and thus leakages. For instance, although leakages amount only to 3% in Paris, they may reach up to 40% in some rural areas to average overall 21,9%. More maintenance helps reducing leakages and save water resource. Even though existing infrastructures are often old, the replacement rate for years 2006-2008 remained rather small with only 0,6% of the networks being renewed over that period.[4]

Given the specificities of each network, the operator managing the service may learn over time the state of the network. Private information over the cost of providing the service is thus pervasive. It introduces a fundamental asymmetry between municipalities and operators. This asymmetry is also partly due to the difficulties in assessing how costs, including labor costs, overhead costs, and maintenance investments, are allocated between water supply, distribution or sewage treatment, especially when more than one of these three operations are shared by the same operator.[5]

Such heterogeneity explains also significant disparities in consumption prices, with an average bill of 183 euros per inhabitant by 2008 which slightly increases over time as a mean to finance an improving quality or new investments.

As far as pricing is concerned, water is billed with two-part tariffs; an usual feature of pricing in network industries.[6] The fixed-fee helps to cover fixed cost while the variable part that depends on consumption aims at paying for variable costs. Fixed-fees significantly vary across networks and but on average are equal to 32 euros which represents around 20% of the bill for an average consumption of 120 m3.

3 Theory

Our theoretical model of the contractual relationship between municipalities (sometimes referred to as “principals” in the sequel) and service providers (the “agent’) fits the actual contractual practices reviewed in Section 2. In particular, informational asymmetries, heterogeneity and the form of pricing are key ingredients of any formal description of the sector.

[4]See Commissariat Général au Developpement Durable (2010).

[5]For a discussion of cost in the French stare sector, see Garcia and Thomas (2001).

[6]Hall (2000).

3.1 Preliminaries

We first set up the stage before entering into more details into the characterization of an optimal contract in the environment under scrutiny.

Demand side. We consider a population of heterogeneous municipalities which differ in terms of the consumers surplus that prevails locally. More precisely, we assume that there exists a shift parameter ε, common knowledge for contracting parties (although not observed by the econometrician), such that the surplus in a municipality characterized by ε writes as $S(q, \varepsilon)$, where $S(q, \varepsilon)$ is increasing and concave in the consumption q and increasing in ε. This parameter ε allows to take into account heterogeneity on the demand side.

In a given municipality, aggregate demand for water at price p is then denoted as $D(p, \varepsilon) = (S'_q)^{-1}(p, \varepsilon)$ with $D'_p(p, \varepsilon) = \frac{1}{S''_{qq}(D(p,\varepsilon),\varepsilon)} < 0)$ for any realization of ε. Of course, the following identity holds

$$p = S'_q(D(p, \varepsilon), \varepsilon).$$

Supply side. The cost function of the service operator is parameterized as $\theta C_0(q)$, where q is the amount produced and and θ is an efficiency parameter that enters multiplicatively. We assume that the function C_0 is strictly increasing and convex. Observe that this cost function satisfies the usual Spence-Mirrlees assumption; an operator with a more efficient technology (θ lower) also produces at a lower marginal cost. As usual in the screening literature, this assumption ensures that different operators can be sorted according to their marginal cost of producing the service and choose accordingly to produce under different contractual terms.

Information. In this paper, we are instead interested in the case where the cost parameter θ is the firm's private information. This parameter is distributed according to a common knowledge atomless distribution F, with a positive density function f on a bounded support $\Theta = [\underline{\theta}, \bar{\theta}]$. In accordance with the screening literature,[7] we impose the familiar *monotone hazard rate property*[8] that ensures fully separation allocations at an optimal contract.

Assumption 1 (MHR).

$$\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0, \quad \forall \theta \in \Theta.$$

[7] Guesnerie and Laffont (1984) and Laffont and Martimort (2002, Chapter 3).

[8] See Bagnoli and Bergstrom (2005).

Contracts. The contract between the municipality and the firm stipulates not only a price p per unit of water produced but also an upfront subsidy A under the form of (per capita) subscription fees paid by consumers to access the service. That subsidy distribute the overall surplus between the consumers and the operator. Implicit in this specification of the contract is the idea that controlling the unit price amounts to controlling demand and thus the production that meets this demand.[9]

Following the incentive regulation literature, we will envision the result of contractual negotiations between the firm and the municipality as the choice of an item by the privately informed party within a menu of options. Two equivalent approaches might be used to model this choice. The first one relies on the so-called *Revelation Principle*[10] which states that there is no loss of generality in looking for contracts which are direct and truthful mechanisms of the kind $\{A(\hat{\theta}, \varepsilon), p(\hat{\theta}, \varepsilon)\}_{\hat{\theta}\in\Theta}$. Note that we index the contract by the demand shock ε which is commonly known by contracting parties. With such direct communication, the operator picks a subsidy/unit price price according to his efficiency parameter by communicating information on its cost parameter. The mechanism is incentive compatible when each operator ends up preferring the option targeted to his own type.

An alternative approach based on the so-called *Taxation Principle*[11] gives up the abstract direct communication process underlying the Revelation Principle and focuses instead on the true economic choice made by the privately informed party. Facing a nonlinear scheme $A(p, \varepsilon)$[12] linking the value of the subscription fee to the actual per unit price chosen by the firm, the firm chooses optimally at which price it stands to produce.[13] Again, this choice can be viewed as a metaphor for more complex negotiation procedures where firms and public authorities negotiate over both the fixed fee and the unit price charged to consumers. We will favor this second approach since it relates to actual observables available to us (price, quantity, fees) which will be used in our econometric analysis.

Objective functions of the contracting parties. With our previous notations at hands, the

[9]See Baron (1989a) for a general formulation of regulatory mechanisms relying on such approach. It is only incidentally different from the standard approach that focuses on the direct control of quantities that is developed in Baron and Myerson (1982) and Laffont and Tirole (1993).

[10]See Myerson (1982) and Laffont and Martimort (2002) for a textbook approach.

[11]See for instance Rochet (1985).

[12]Without fear of confusion and for the sake of simplifying presentation, we slightly abuse notations here by denoting similarly the fixed fee viewed as a function of the agent's type in a direct mechanism and viewed as a function of the price in an indirect scheme. Notice also that the nonlinear scheme $A(\cdot, \varepsilon)$ is again indexed on the commonly observable variable ε that characterizes the relationship under scrutiny.

[13]Of course, any incentive compatible direct mechanism $\{A(\hat{\theta}, \varepsilon), p(\hat{\theta}, \varepsilon)\}_{\hat{\theta}\in\Theta}$ can be transformed into a nonlinear scheme by setting $A(p, \varepsilon) = A(\hat{\theta}, \varepsilon)$ if $p = p(\hat{\theta}, \varepsilon)$ and $A(p, \varepsilon) = -\infty$ otherwise.

expression of the firm's profit becomes:

$$\mathcal{U}(\theta,\varepsilon,p,A) = A + pD(p,\varepsilon) - \theta C_0(D(p,\varepsilon)).$$

Following Baron and Myerson (1982), a municipality maximizes a welfare function which includes not only consumers' net surplus from consuming the service but also the firm's profit weighted by some parameter $\gamma \in [0,1[$. This parameter can be viewed as an index of the firm's bargaining power at the the of tenders or during contract negotiations. It can also be inherited from how local political forces interact as argued in Baron (1989).[14] The corresponding welfare function $W(\cdot)$ can thus be written as:

$$\mathcal{W}(\theta,\varepsilon,p,A) = S(D(p,\varepsilon),\varepsilon) - A - pD(p,\varepsilon) + \gamma\mathcal{U}(\theta,\varepsilon,p,A).$$

Using the expression of the fixed-fee as a function of the firm's profit, the latter definition becomes:

$$\mathcal{W}(\theta,\varepsilon,p,A) = S(D(p,\varepsilon),\varepsilon) - \theta C_0(D(p,\varepsilon)) - (1-\gamma)\mathcal{U}(\theta,\varepsilon,p,A).$$

This latter expression stresses the rent/efficiency trade-off faced by the local government in designing the regulatory contract. On the one hand, the principal would like to charge a price $p(\theta,\varepsilon)$ close to $p^*(\theta,\varepsilon)$ as defined in (1) so that the overall surplus is maximized. On the other hand, the principal would also like to reduce the firm's informational rent which is viewed as socially costly. Under asymmetric information, rents and outputs are linked altogether through incentive compatibility conditions and this leads to an important trade-off between the conflicting objectives of promoting efficiency and extracting rents.

Remark 1. *The expressions of the objective functions above can easily be extended to accounts for some fixed-cost F in the operatopr's cost function. Suppose indeed that the operator's profit can be written as $\mathcal{U}(\theta,\varepsilon,p,A) = A' + pD(p,\varepsilon) - \theta C_0(D(p,\varepsilon)) - F$. for some fixed-fee A'. Setting $A \equiv A' - F$ then amounts to having the principal pays for the fixed-cost in the first place, which is basically an accounting convention.*

Benchmark. Had θ been common knowledge, efficiency would require to produce a quantity $q^*(\theta,\varepsilon) = D(p^*(\theta,\varepsilon),\varepsilon)$ such that the marginal social value of production is equal to marginal cost:

$$S'_q(q^*(\theta,\varepsilon),\varepsilon) = p^*(\theta,\varepsilon) = \theta C'_{0q}(q^*(\theta,\varepsilon)). \tag{1}$$

[14]See also Gagnepain, Ivaldi and Martimort (2013) for a model of the French transportation sector that relies on a similar specification of the preferences of local authorities.

Then, the firm's operates if $A^*(\theta, \varepsilon)$ extracts revenues from the service:

$$A^*(\theta, \varepsilon) = -p^*(\theta, \varepsilon)D(p^*(\theta, \varepsilon), \varepsilon) + \theta C_0(D(p^*(\theta, \varepsilon), \varepsilon)). \tag{2}$$

Remark 2. *In the empirical part of our analysis, the two functions $S(\cdot, \varepsilon)$ and $C_0(\cdot)$ will depend on a set of explanatory variables. For example, the treatment made for making the water drinkable has an impact on the unit cost. This treatment is however observed in the data. For the exposition, we omit the dependance in the explanatory variables without loss of generality for the results derived in the analysis of the theoretical model.*

3.2 Optimal Contract

Incentive compatibility constraints. Let define the firm's information rent $U(\theta, \varepsilon)$ and an optimal price[15] respectively as:

$$U(\theta, \varepsilon) = \max_p A(p, \varepsilon) + pD(p, \varepsilon) - \theta C_0(D(p, \varepsilon)) \tag{3}$$

and

$$p(\theta, \varepsilon) = \arg\max_p A(p, \varepsilon) + pD(p, \varepsilon) - \theta C_0(D(p, \varepsilon)). \tag{4}$$

From (3), $U(\theta, \varepsilon)$ is the maximum of a family of decreasing linear functions in θ. As such it is decreasing, convex in θ and absolutely continuous so that one can write:

$$U(\theta, \varepsilon) = U(\bar{\theta}, \varepsilon) + \int_{\theta}^{\bar{\theta}} C_0\left(D(p(x, \varepsilon), \varepsilon)\right) dx. \tag{5}$$

At any point of differentiability in θ (i.e., almost everywhere), we get:

$$U'_\theta(\theta, \varepsilon) = -C_0\left(D(p(\theta, \varepsilon), \varepsilon)\right). \tag{6}$$

Because $U(\theta, \varepsilon)$ is convex in θ, $C_0\left(D(p(\theta, \varepsilon), \varepsilon)\right)$ is non-decreasing in θ, which in turn implies:

$$p(\theta, \varepsilon) \text{ is non-increasing in } \theta. \tag{7}$$

This condition expresses the fact that firms endowed with less efficient technologies produce lower volumes at higher prices . From this monotonicity, it also follows that $p(\theta, \varepsilon)$ is almost everywhere differentiable in θ.

[15]Or at least, a selection within the best-response correspondence.

Let us turn now to the expression of the nonlinear schedule $A(p, \varepsilon)$ that plays an important role in our empirical analysis. By a standard duality argument of (*generalized*) convex analysis, we may first rewrite:[16]

$$A(p, \varepsilon) + pD(p, \varepsilon) = \min_{\theta} U(\theta, \varepsilon) + \theta C_0(D(p, \varepsilon)).$$

From which, it also follows that $A(p, \varepsilon)$ is absolutely continuous in p and thus such that

$$\begin{aligned} A(p, \varepsilon) + pD(p, \varepsilon) =& A(p(\bar{\theta}, \varepsilon), \varepsilon) + p(\bar{\theta}, \varepsilon) D(p(\bar{\theta}, \varepsilon), \varepsilon) \\ &+ \int_{p(\bar{\theta}, \varepsilon)}^{p} \vartheta(p, \varepsilon) C'_{0q}\left(D(p, \varepsilon)\right) D'_p(p, \varepsilon) dp \end{aligned} \tag{8}$$

where $\vartheta(p, \varepsilon) = \min_\theta U(\theta, \varepsilon) + \theta C_0(D(p, \varepsilon))$ is an assignment function[17] (a selection within the best-response monotonically increasing p (thus almost everywhere differentiable) and such that $\vartheta(p(\theta, \varepsilon), \varepsilon) \equiv \theta$.

At any point of differentiability in $p = p(\theta, \varepsilon)$, we thus have also:

$$A'_p(p(\theta, \varepsilon), \varepsilon) = -\left(p(\theta, \varepsilon) - \theta C'_{0q}\left(D(p(\theta, \varepsilon), \varepsilon)\right)\right) D'_p(p(\theta, \varepsilon), \varepsilon) - D(p(\theta, \varepsilon), \varepsilon). \tag{9}$$

Participation constraints. The firm chooses to always operate the service irrespectively of its costs when it at least breaks even. For a fixed ε, this participation constraint can be written as:

$$U(\theta, \varepsilon) \geq 0 \qquad \forall \theta.$$

As usual in the incentive regulation literature,[18] this constraint is binding for the worst type $\bar{\theta}$ at the optimal contract. Otherwise reducing uniformly the fixed fee by some small amount would improve the principal's expected payoff while maintaining incentive compatibility. From this observation, an immediate manipulation of (5) yields the following expression of the firm's information rent as:

$$U(\theta, \varepsilon) = \int_{\theta}^{\bar{\theta}} C_0\left(D(p(x, \varepsilon), \varepsilon\right) dx. \tag{10}$$

Observe that the rent left to the operator is greater as prices are lower. The intuition is a standard one. By pretending being slightly less efficient, an operator with efficiency parameter θ can produce the same quantity than this slightly less efficient type $\theta + d\theta$ but

[16] See for instance Basov (2005, Chapter 7).
[17] See Noldecke and Samuelson (2007).
[18] See Armstrong and Sappington (2007), Baron and Myerson (1982) and Laffont and Tirole (1993) among others.

at a lower marginal cost. To induce this operator to report truthfully his type he must be given an extra fee that equals the corresponding cost saving $d\theta C_0\left(D(p(\theta+d\theta,\varepsilon),\varepsilon)\right) \approx d\theta C_0\left(D(p(\theta+d\theta,\varepsilon),\varepsilon)\right)$. The right-hand side of (10) expresses how those marginal information rents just pill up over all supra-marginal types.

Optimal contracts. Under asymmetric information, an optimal contract maximizes the expected welfare of the municipality subject to incentive and participation constraints. From our observations above, that incentive feasible set can be summarized by constraints (7) and (10). As usual, the monotonicity condition (7) will be omitted in a first step and checked ex post on the solution to the so relaxed problem. Formally, this relaxed problem can be written as:

$$\max_{\{p(\cdot,\varepsilon),U(\cdot,\varepsilon)\}} \int_{\underline{\theta}}^{\overline{\theta}} \left[S(D(p(\theta,\varepsilon),\varepsilon)) - \theta C_0\left(D(p(\theta,\varepsilon),\varepsilon)\right) - (1-\gamma)U(\theta,\varepsilon)\right] dF(\theta) \text{ subject to (10).}$$

Using (10) and integrating by parts yields the following expression of the expected rent left to the operator:

$$\int_{\underline{\theta}}^{\overline{\theta}} U(\theta,\varepsilon)dF(\theta) = \int_{\underline{\theta}}^{\overline{\theta}} \frac{F(\theta)}{f(\theta)} C_0\left(D(p(\theta,\varepsilon),\varepsilon)\right) dF(\theta).$$

This expression can be incorporated into the maximand above before proceeding to pointwise optimization. This last step leads to the following expression of the price per-unit of consumption $p(\theta,\varepsilon)$.

$$p(\theta,\varepsilon) = \left(\theta + (1-\gamma)\frac{F(\theta)}{f(\theta)}\right) C'_{0q}\left(D(p(\theta,\varepsilon),\varepsilon)\right). \tag{11}$$

The corresponding volume that is supplied $q(\theta,\varepsilon)$ is then defined as:

$$S'_q(q(\theta,\varepsilon),\varepsilon) = p(\theta,\varepsilon). \tag{12}$$

Equation (11) indicates that the price is now above marginal costs and, as a result of (12). Equilibrium quantities are also lower than at the first best. Increasing the unit price above marginal cost reduces the demand addressed to the operator. It thus reduces the latter's information rent. Formally, everything happens as if the cost parameter was now replaced by a *virtual cost parameter* $H(\theta,\gamma)$ which is greater:

$$H(\theta,\gamma) = \theta + (1-\gamma)\frac{F(\theta)}{f(\theta)}.$$

This expression first illustrates the rent/efficiency trade-off that arises under asymmetric information and, second, how this trade-off is modified as parameters of the model change. Indeed, the virtual cost parameter is greater as the public authority is more concerned by rent extraction (i.e., γ lower) and as the types distribution is more front-loaded, in the sense of having a greater hazard rate $F(\theta)/f(\theta)$.

Turning now to sufficient conditions for optimality, observe that Assumption 1 ensures that $p(\theta, \varepsilon)$ so defined by (11) is non-decreasing in θ as requested by condition (7). Hence, the solution of the relaxed problem really characterizes the optimal contract. When we get to our empirical analysis, we will actually check on our estimated distribution that it indeed satisfies Assumption 1.

4 Nonparametric Identification

In this section, we study the nonparametric identification of our model. It is indeed important to figure what are the structural functions that can be fully recovered from the available data. In our dataset, we observe for each local community, the unit price p, the quantity consumed q and the fixed fee A. We also observe some explanatory variables W and Z which are related to respectively the cost function $C_0(q, W)$ and the surplus function $S(q, Z, \varepsilon)$. The vectors W and Z do not have any variable in common in our case but the results are unchanged when the reverse holds as long as there are some exclusion restrictions, i.e., there exists an explanatory variable in Z which is not part of W and an explanatory variable in W which is not part of Z. Our purpose is to identify the production technology $C_0(q, W)$, the distribution $F(\theta)$ of the types θ, the consumers surplus function $S_0(q, Z, \varepsilon)$, and the weight γ of the firm's profit in the principal's objective.

A scale normalization Multiplying θ by a positive scalar λ and dividing the cost function by the same value would give the same equilibrium outcome. θ and the marginal cost are therefore identified up to a scale and we first need to impose a normalization for the distribution of θ, by assuming that a prespecified quantile is equal to a given value. Standard normalizations are $\underline{\theta}$ equal to 1 or the median of θ equal to 1. In the latter case, the function $C'_{0q}(q, W)$ is then interpreted as the marginal cost function for the median type firm, given the observed characteristics W.

Assumption 2. *[Normalization]*

$$Median(\theta) = 1.$$

4.1 The Simple Case Without Heterogeneity

To start with, we do not consider neither the presence of explanatory variables (no W and no Z) nor the presence of heterogeneity, *i.e.* $\varepsilon \equiv 0$, simplifying notations accordingly. For exposition purpose, it is important to know what are the identifying power of the first-order conditions (9), (11) and (12) and what do the additional assumptions help to identify. The system of first-order conditions reduces in this case to[19]

$$p = H(\theta,\gamma)C'_{0q}(q)\,, \tag{13}$$

$$S'(q) = p, \tag{14}$$

$$A'(p) = -q - \left(p - \theta C'_{0q}(q)\right) D'_p(p). \tag{15}$$

Observe that equation (14) directly identifies $S'(q)$ on the support of the equilibrium quantities and hence provides the expression of $D'_p(p) = \frac{\partial}{\partial p}\left(S'^{-1}\right)(p)$ that we insert into (15). Finally, (15) provides information about the operator's price-cost margin

$$\frac{p - \theta C'_{0q}(q)}{p} = \frac{A'(p) + q}{pD'_p(p)}.$$

This is so because the terms on the right-hand side are either observed ($q = D(p)$ and p) or derived directly from the observations ($A'(p)$ and $D'(p)$).

Making the dependence of the price-cost margin on θ explicit, we define a price-cost margin $r(\theta,\gamma)$ as:

$$r(\theta,\gamma) = \frac{p(\theta) - \theta C'_{0q}(q(\theta))}{p(\theta)} = \frac{1}{1 + \frac{\theta f(\theta)}{(1-\gamma)F(\theta)}} \tag{16}$$

where the last equality immediately follows from (13). Thus, the price-cost margin only depends only on the efficiency parameter θ and the bargaining power γ. In the sequel, we shall assume that different operators can be perfectly sorted according to that price-cost margin. Formally, we require that there is a one-to-one mapping between price-cost margins and efficiency parameters, i.e., $r(\theta,\gamma)$ is a monotonically decreasing transformation of θ which always lies between 0 and 1 and is worth 0 at $\underline{\theta}$, i.e., for the most efficient type who produces efficiently:

Assumption 3 (MPCM)**.**

$$\frac{d}{d\theta}\left(\frac{F(\theta)}{\theta f(\theta)}\right) \geq 0, \quad \forall \theta \in \Theta.$$

[19] We omit, in the notations, the dependance of p and q in θ for the clarity of exposition.

This assumption is actually **stonger** than Assumption 1 and is satisfied for all standard parametric distributions.

To give a bit more intuition about the role played by Assumption 3, let us come back to (15) which can be rewritten as:

$$A'(p(\theta)) = -D(p(\theta))\left(1 + r(\theta)\varepsilon_D(p(\theta))\right) \tag{17}$$

where we denote the demand elasticity by $\varepsilon_D(p) = -\frac{pD'(p)}{D(p)}$. Assuming also that demand is more elastic at greater price, i.e., $\varepsilon_D(p)$ is increasing with p, it is straightforward to check by differentiating (17) that Assumption 3 ensures that $-A'(p(\theta))/D(p(\theta))$ is increasing in θ, or alternatively that $A(p)$ is quasi-concave in p.

Equipped with this one-to-one relationship between cost-price margins and efficiency parameters, we now let $G(\cdot)$ (resp. $g(\cdot)$) be the cumulative distribution function (resp. probability density function) of that margin. Of course, we have $G(r(\theta,\gamma)) = F(\theta)$ and by differencing the last equality with respect to θ, we also get $g(r) = f(\theta)\theta'_r(r,\gamma)$ where $\theta(r,\gamma)$ denotes the inverse function of $r(\theta,\gamma)$. Reintroducing the last two equalities into (16), we obtain after some manipulations:

$$(1-\gamma)\frac{\theta_r(r,\gamma)}{\theta(r,\gamma)} = \frac{r}{1-r}\frac{g(r)}{G(r)}.$$

Assume, just for this case that we impose the normalization $\underline{\theta} = 1$.[20] As $r = 0$ when $\theta = \underline{\theta} = 1$, we can solve the latter differential equation and get:

$$\theta(r,\gamma)^{1-\gamma} = \exp\left[\int_0^r \frac{s}{1-s}\frac{g(s)}{G(s)}ds\right]. \tag{18}$$

The density function for θ is then derived from $f(\theta) = \frac{g(r)}{\dot{\theta}(r)}$:

$$f(\theta) = (1-\gamma)\frac{1-r}{r\theta}G(r). \tag{19}$$

Observe that the last expression gives the density of the types as a function of the c.d.f. of the price cost margin. Reintroducing (18) into (13) and observing that $\theta = (1 - r(\theta,\gamma))H(\theta,\gamma)$

[20] If we use the normalization related to the median, Equation (18) below becomes

$$\theta(r,\gamma) = \underline{\theta}\exp\left[\frac{1}{1-\gamma}\int_0^r \frac{s}{1-s}\frac{g(s)}{G(s)}ds\right].$$

$\underline{\theta}$ is determined to ensure that $Median(\theta) = 1$.

lead to

$$C'_{0q}(q) = \frac{p}{H(\theta,\gamma)} = \frac{p(1-r)}{\exp\left[\frac{1}{1-\gamma}\int_0^r \frac{s}{1-s}\frac{g(s)}{G(s)}ds\right]}.$$

Therefore, when γ is known or predetermined from some additional source of information, the model is thus identified since the marginal cost and the distribution of the firms' types are identified up to the standard normalization.

Instead, when γ is not known, only $\theta^{1-\gamma}$ is identified. Consequently, for any β in $[0,1[$, the two sets $(\theta, C'_{0q}(q), \gamma)$ and $(\theta^{\frac{1-\gamma}{1-\beta}}, C'_{0q}(q)\theta^{\frac{\gamma-\beta}{1-\beta}}, \beta)$ are observationally equivalent. We thus need additional assumptions or information to identify the model. It is worth noting that the conclusions of this section does not depend on the existence of observed explanatory variables. Adding an explanatory variable in either the cost function or the surplus one does not change the previous results. Observe also that we can exploit the bounds on β to bound the distributions of interest. In particular

$$C'_{0q}(q) \leq \frac{p}{H(\theta,\gamma)} = \frac{p(1-r)}{\exp\left[\int_0^r \frac{s}{1-s}\frac{g(s)}{G(s)}ds\right]},$$

and

$$f(\theta) \leq \frac{1-r}{r\theta}G(r), \text{ on } [1;+\infty[.$$

4.2 Full Identification of the Model with Explanatory Variables and Heterogeneity

We now consider the full model with explanatory variables, Z and W, that appears respectively in the surplus function and the cost function.

4.2.1 Identification of the Marginal Surplus Function under Completeness Assumption

• In a first step, we assume that the marginal cost function is separable and additive in both the explanatory variables, Z,[21] and the unobserved heterogeneity term, ε, i.e. $S'_q(q,Z,\varepsilon) = S'_0(q) + \beta_Z Z + \varepsilon$, Equation (12) can be written

$$p = S'_0(q) + \beta Z + \varepsilon, \tag{20}$$

[21]The separability in Z does not play any role in the identification result. This is nevertheless the model we estimate in the empirical application.

with $[\varepsilon|Z,W]=0$.

The equilibrium quantity q being endogenous, Equation (20) is a typical nonparametric instrumental regression like the one studied in Florens et al. (2005).[22] Following Theorem 3 in Florens, Johannes and van Bellegem (2005), we recover the identification of the primitives of the model under a condition of completeness of W with respect to q. Observe that here, in the partial linear regression like the one in (20), we need additional regularity conditions than the completeness assumption due to the linear part $\beta_Z Z$ (see Florens et al., 2005)[23]. The completeness assumption is nevertheless the most restrictive assumption though relatively standard in this literature. The set of instruments W is complete for q if, for any measurable function ΔS in L^1,

$$E[\Delta S(q)|W]=0 \ a.s. \ \Rightarrow \Delta S(q)=0 \ a.s.$$

Sufficient conditions for completeness can be found in Newey and Powell (2003), Chernozhukov, Imbens, and Newey (2007), and Andrews (2011). It can be replaced by weaker concepts like the bounded completeness in Chernozhukov and Hansen (2005), Blundell, Chen and Kristensen (2007), d'Haultfoeuille (2010), and Andrews (2011).[24]

• We now consider the non-separable case,

$$p = S'_q(q,Z,\varepsilon),$$

with $[\varepsilon|Z,W]=0$. Let V be indeed a given c.d.f., i.e. an injective function from $\mathbb{R}$ to $[0,1]$. Let $\tilde{\varepsilon}=V^{-1}\circ\Phi(\varepsilon)$, where $\Phi(\cdot)$ denotes the c.d.f. of ε, and $\tilde{S}(q,Z,\tilde{\varepsilon})=S(q,Z,\Phi^{-1}\circ V(\tilde{\varepsilon}))$. The equilibrium quantities A, p and q are invariant to this transformation.

The distribution of the parameter ε is therefore non identifiable as such and we therefore need to normalize the distribution of ε. The standard one is to assume that it is uniformly distributed, i.e. $\varepsilon\sim U[0;1]$. It is worth noting, that this normalization does not restrict the economic interpretation of our model. The value of ε does not have an interpretation in its own and is only an index of the position of the firm in a demand ranking. We therefore identify the marginal surplus (given Z) of any quantile of the distribution of ε, like the marginal surplus of the median local community, given Z.

The identification arises from the monotonicity constraints implied by the economic

[22]See also Newey and Powell (2003), Blundell et al. (2007), Darolles et al. (2011) and Ai and Chen (2003) for additional references.

[23]It is required that the conditional expectations of Z, q and any L^2 function of q given W are in $L^2(\mathbb{R}^{dimW})$ and that the matrix $E\left[E(Z|W)E(Z|W)^\top\right]$ is full rank.

[24]See also Chen et al. for a general discussion. This assumption is however not testable as recently shown by Canay et al. (2012) .

model (see, for example Chesher, 2007, or Chernozhukov and Hansen, 2005). For a given rank $\alpha \in [0,1]$:

$$\begin{aligned}
\mathbb{P}\left(p \leq \left(S'_q(q,\alpha,Z)\right)|Z,W\right) &= \mathbb{E}_{q|Z,W}\mathbb{P}\left(S'_q(q,\varepsilon,Z) \leq \left(S'_q(q,\alpha,Z)\right)|q,Z,W\right) \\
&= \mathbb{E}_{q|Z,W}\mathbb{P}(\varepsilon \leq \alpha|q,Z,W) \text{ due to the monotonicity in } \varepsilon \\
&= \mathbb{P}(\varepsilon \leq \alpha|Z,W) = \alpha
\end{aligned} \tag{21}$$

Equation (21) derived above can be reexpressed as a conditional moment restriction:

$$\mathbb{E}\left(\mathbf{1}\{p \leq S'_q(q,\alpha,Z)\} - \alpha|Z,W\right) = 0. \tag{22}$$

This is a standard quantile IV equation and we need essentially an assumption of completeness of W with respect to q to identify any quantile of the marginal surplus function given Z. The following result summarizes the main conclusion if this section.

Proposition 1. *If W is complete with respect to q the marginal surplus function given Z is identified.*

4.3 Identification of the marginal cost function and the distribution of types

We now prove that the marginal cost function is identified without the knowledge of γ when there is unobserved heterogeneity of the demand function. Assume initially that we do not have explanatory variables for the marginal cost.

We consider the points (q_i, p_i) of the quantities consumed and marginal prices paid for all the contracts (see Figure 1). For the most efficient firms, $\theta = \underline{\theta}$, the marginal price equation (11) reveals that the marginal cost function for the lower type $\underline{\theta}C'_{0q}(q)$ can be estimated using the lower envelope of the points in the space (q,p) between the quantity $q^l_{\underline{\theta}}$ which corresponds to the most efficient firm contracting with the lowest demand city (the contract which has the lowest price) and the quantity $q^u_{\underline{\theta}}$ which corresponds to the most efficient firm contracting with the highest demand city (i.e. the contract with the highest quantity). This is the dashed curve in the figure.

Similarly, for the less efficient firms, $\theta = \overline{\theta}$, the marginal prices are equal to $kC'_{0q}(q)$ where k is equal to the limit of $\theta + (1-\gamma)/f(\theta)$ on the upper bound, when it exists. This function is therefore the upper envelope between the minimum quantity among all contracts $q_{\overline{\theta}}^{l}$ (i.e. the quantity of the contract of the less efficient firm contracting with the local community with the lowest surplus) and the quantity corresponding to the maximum price $q_{\overline{\theta}}^{u}$ (the quantity consumed in the local community with the highest surplus which contracts with the less efficient firm).

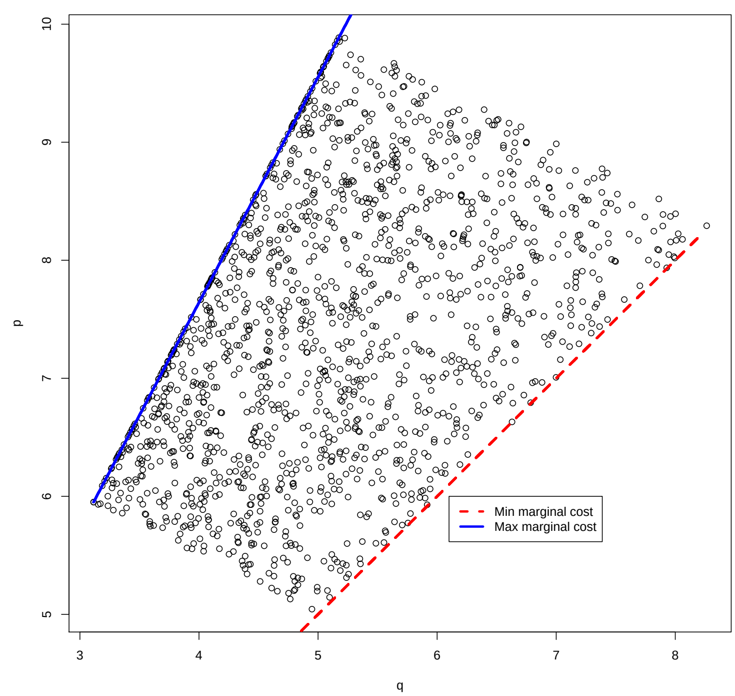


Figure 1: The set of contracts in the $(q \times p)$ space.

If $q_{\overline{\theta}}^{u} \geq q_{\underline{\theta}}^{l}$ and if k is finite, we can estimate the ratio of the two extreme costs with the same quantity q and therefore estimate $C_{0q}'(q)$ up to a scale between the minimum and the maximum quantities of the population of contracts. We need to have sufficient heterogeneity in ε. If there are explanatory variables W, we can use the same argument conditionally on

W.

Alternatively, if k is infinite, or if there is not enough heterogeneity, we can use a moment equation like the one used for estimating the surplus function in (21). Using the monotone property of $H(\theta) = \theta + (1-\gamma)\frac{F(\theta)}{f(\theta)}$ in θ, we can indeed write for any quantile $\alpha \in]0,1[$, θ_α being the corresponding quantile of θ,

$$\mathbb{P}\left(p \leq H(\theta_\alpha, \gamma) C'_{0q}(q, W) | Z, W\right) = \alpha. \tag{23}$$

Again, $C'_{0q}(q, W)$ is identified through (23) up to a scale under an assumption of completeness of Z. We can in a first step normalize $H(\theta_{0.5}, \gamma)$ to 1 before changing the scale to ensure that the median of θ is equal to 1.

Once $C'_{0q}(q, W)$ is known up to a scale, we obtain θ from (9) and estimate the scale to ensure $Median(\theta) = 1$. The marginal cost function is therefore identified. The knowledge of θ provides the identification of $F(\theta)$ and $f(\theta)$.

Proposition 2. *If Z is complete with respect to q the marginal cost function given W, $C'_{0q}(q, W)$ and the distribution of types are identified.*

4.4 Identification of γ

Once, all the functions are identified, $H(\theta, \gamma)$ is known from (11). We can now identify the bargaining weight γ using

$$H(\theta, \gamma) = \theta + (1-\gamma)\frac{F(\theta)}{f(\theta)}.$$

Identification when γ is not unique It is worth noting that the identification strategy does not require the unicity of γ across the cities as long as the distribution of γ is exogeneous. Assume that γ varies across cities, its c.d.f being denoted by $\Gamma(\cdot)$, and that the observed heterogeneity Z and W are both complete with respect to q as required in Proposition 1 and 2. The equilibrium equations are not changed and we identify similarly $S'_q(q, Z)$.

Equation (23) is now modified as we have to condition on γ. For any quantile $\alpha \in]0,1[$,

$$\mathbb{P}\left(p \leq H(\theta_\alpha, \gamma) C'_{0q}(q, W) | Z, W, \gamma\right) = \alpha. \tag{24}$$

For any $\gamma \in [0,1[$, the function $H(\cdot, \gamma)$ is a one-to-one mapping from $[\underline{\theta}, \overline{\theta}[$ to $[\underline{\theta}, \lim_{\theta \to \overline{\theta}} \theta + (1-\gamma)\frac{F(\theta)}{f(\theta)}[$.

Let $\theta(k,\gamma) = \max_{\theta\in\Theta}\{H(\theta,\gamma)\leq k\}$. For any γ, $\theta(\underline{\theta},\gamma) = \underline{\theta}$ and $\theta(\cdot,\cdot)$ is increasing in k and decreasing in γ. Therefore

$$\mathbb{P}\left(p \leq kC'_{0q}(q,W)|Z,W,\gamma\right) = F(\theta(k,\gamma)),$$

and, integrating with respect to gamma,

$$\mathbb{P}\left(p \leq kC'_{0q}(q,W)|Z,W\right) = \int F(\theta(k,\gamma))d\Gamma(\gamma) = r(k). \tag{25}$$

$r(k)$ is increasing in k and maps $[\underline{\theta}, \lim_{\theta\to\overline{\theta}}\theta + \frac{F(\theta)}{f(\theta)}[$ into $[0,1[$. We have therefore derived the same structure of quantiles than in the case where γ is unique. Again, we can identify the marginal cost up to a scale before deriving θ from (9). Once θ is known, we can rescale it to meet our normalization $Median(\theta) = 1$. $H(\theta,\gamma)$ is therefore identified from $H(\theta,\gamma) = p/C'_{0q}(q,W)$.

Finally the distribution of γ is identified from

$$H(\theta,\gamma) = \theta + (1-\gamma)\frac{F(\theta)}{f(\theta)}.$$

5 Simulations and Estimation Procedure

In this section, we expose the different steps of our estimation strategy. Some Monte Carlo simulations are also displayed to assess the sample properties of the proposed estimator.

The results are based on $n_s = 100$ replications of samples of size $n = 1000$. In this scenario the marginal cost function is a linear function $C'_{0q}(q,W) = q + W$, where $W \sim \mathbb{U}_{[0,1]}$. θ the type of the firm follows a Beta distribution up to a location-scale transformation, i.e. $(\theta-1)/5 \sim B(2,5)$ and therefore $\Theta = [1,6]$. $S'_0(q,Z) = 15 - 5q + Z$, where $Z \sim \mathbb{U}_{[0,1]}$ and $\gamma = 0.3$. The unobserved heterogeneity ε is drawn from a centered uniform distribution.

We display for the estimated functions the integrated square bias, integrated variance and integrated mean square error based on 100 simulations. Let $\hat{f}_s(\cdot)$ be the estimated function of $f(\cdot)$ for simulation s. Let $x_1, ..., x_{99}$ be the 99 percentiles of the variable x in the population. We define $\overline{f}(x) = \frac{1}{n_s}\sum_{s=1}^{n_s}\hat{f}_s(x)$.

The integrated square bias is estimated by $ISB = \sum_{i=1}^{99}\left(\overline{f}(x_i) - f(x_i)\right)^2$, the integrated square variance by $ISV = \sum_{i=1}^{99}\frac{1}{n_s}\sum_{s=1}^{n_s}\left(\hat{f}_s(x_i) - \overline{f}(x_i)\right)^2$. Finally the integrated Mean Square Error IMSE is estimated by $IMSE = \sum_{i=1}^{99}\frac{1}{n_s}\sum_{s=1}^{n_s}\left(\hat{f}_s(x_i) - f(x_i)\right)^2$.

5.1 Estimation of the marginal surplus function

We assume that the marginal surplus is separable and additive in both the explanatory variables, Z and the unobserved heterogeneity ε like in Equation (20):

$$p = S_q'(q) + \beta_Z Z + \varepsilon. \tag{26}$$

We apply Ai and Chen (2003) to estimate the marginal surplus using cubic spline approximation. This is attractive as $S_q'(q)$ can be forced to be decreasing by imposing that the sequence of coefficients related to the spline approximation is also decreasing.

1. We choose the cubic splines $B_4(t)$ as basis for our sieve estimation and renormalize q on $[0,1]$ (q^* is the normalized quantity) for the choice of the knots (G denotes the c.d.f. of q):

$$\begin{aligned} S_q' \circ G^{-1}(q^*) = \phi(q^*) &= \sum_{l=-3}^{2^k-1} B_4(2^k q^* - l) \\ &= P_{q^*}\beta^{K_n}, \end{aligned}$$

   where $K_n = 2^{d_n} + 4$ is the number of parameters that are estimated. d_n is a tuning parameter which drives the approximation of the true (normalized) marginal surplus function.

2. We then project the moment equation

$$m(Z, W, \beta^{K_n}, \beta_Z) = \mathbb{E}\left[p - P_{q^*}\beta^{K_n} - \beta_Z Z | Z, W\right]$$

   on a basis of the square-integrable function of Z, W, $p_j(Z,W)$, $j = 1, \ldots, J_n$, where J_n tends slowly to infinity as $n \to \infty$. We denote by $P^{J_n}(Z,W) = (p_1(Z,W), \ldots, p_{J_n}(Z,W))$ and

$$P = (P^{J_n}(Z_1, W_1), \ldots, P^{J_n}(Z_n, W_n).$$

   An empirical estimator of $m(Z, W, \beta^{K_n}, \beta_Z)$ is therefore:

$$\hat{m}(Z, W, \beta^{K_n}, \beta_Z) = \sum_{j=1}^{n} \left(p_j - P_{q_j^*}\beta^{K_n} - \beta_Z Z_j\right) P^{J_n}(Z_j, W_j) \left(P'P\right)^{-} P^{J_n}(Z, W).$$

   We can now compute the (empirical) distance function $\hat{Q}(\beta_Z, \beta^{K_n})$ as:

$$\hat{Q}(\beta_Z, \beta^{K_n}) = \frac{1}{n}\sum_{i=1}^{n} \hat{m}(Z_i, W_i, \beta^{K_n}, \beta_Z)\hat{m}(Z_i, W_i, \beta^{K_n}, \beta_Z).$$

   We could have used another metric but it appears that the result are not very sensitive to the choice of the metric.

3. Finally, we estimate the parameters β_Z and β_{K_n} by minimizing the penalized criterion:

$$\hat{Q}(\beta_Z, \beta^{K_n}) + \lambda_n(d_0 + d_2).$$

The second term in the expression above is a penalization term added to control for the ill-posedness where $d_0 = \int_0^1 \phi(q^*)^2 dq^*$ and $d_2 = \int_0^1 \phi''(q^*)^2 dq^*$. d_2 is controlling for the oscillations of the estimated marginal surplus (if the monotonicity is not imposed in the estimation step). λ_n is a tuning parameter which controls for the strength of the penalisation term.

Table 1 displays the integrated square bias, the integrated variance and the integrated MSE for various choices of d_n, J_n and λ_n.

The results are quite good except for small values of λ_n for which the integrated variance increases a lot. The MSE decreases with the improvement of the sieve approximation (K_n) and the size of the basis used to project the conditional moment on (J_n).

5.2 Estimation of the marginal cost function

We estimate the marginal cost function using the SMD procedure of Chen and Pouzo (2011) which is, among additional contributions, a generalization of Ai and Chen (2003) to the case of nonsmooth moments. It is therefore particularly attractive here as we estimate the marginal cost function by quantile IV methods. The marginal cost function $C'_{0q}(q)$ is estimated by another spline approximation. This is a similar procedure than the one explained above. Here the moment condition is

$$\mathbb{E}\left[\mathbf{1}\{\,\big(p \leq \lambda_\alpha(P_{q^*}\delta^{J_n} + \beta_W W)\big)\,\} - \alpha | Z, W\right] = 0,$$

for a choice of 10 quantiles (from 0.05 to 0.95) plus the median that is imposed to be equal to one (in a first step). λ_α is an increasing sequence of parameter which corresponds to the value $H(\theta_\alpha, \gamma)$ for the α-quantile of the firm type, θ_α.

From the estimation of the marginal cost, we can now estimate θ in (9) as all the other quantities can be derived from the first two functions that have been estimated. We finally renormalize θ and the marginal cost function to ensure that the median is equal to 1.

Table 2 and 3 reports for respectively the estimated marginal cost and distribution of types the integrated square variance, mean bias and MSE for various choices of d_n, J_n and λ_n.

Table 1: Monte Carlo Study - Integrated MSE of Sieve IV estimator of $S'_q(q)$

	λ_n											
	$K_n = 4$						$K_n = 5$					
$J_n - K_n$	0.2	0.1	0.01	0.001	1e-4	1e-5	0.2	0.1	0.01	0.001	1e-4	1e-5
1	0.005	0.002	0.001	0.016	0.016	0.034	0.006	0.002	0.002	0.016	0.015	0.059
	0.230	0.221	0.294	0.974	3.392	4.881	0.256	0.205	0.283	0.955	4.319	9.399
	0.235	0.223	0.295	0.990	3.408	4.915	0.262	0.207	0.284	0.971	4.334	9.457
3	0.008	0.004	0.010	0.028	0.014	0.007	0.010	0.001	0.006	0.020	0.009	0.159
	0.280	0.221	0.327	1.049	1.730	2.356	0.260	0.230	0.277	0.847	2.618	2.939
	0.288	0.225	0.337	1.077	1.745	2.363	0.271	0.231	0.283	0.867	2.628	3.098
6	0.014	0.003	0.019	0.020	0.033	0.030	0.015	0.016	0.024	0.012	0.011	0.051
	0.204	0.171	0.275	0.477	0.602	0.656	0.195	0.179	0.292	0.556	0.884	0.900
	0.219	0.174	0.294	0.497	0.635	0.686	0.210	0.195	0.316	0.568	0.895	0.950
11	0.009	0.018	0.011	0.014	0.036	0.017	0.026	0.015	0.033	0.016	0.022	0.031
	0.132	0.141	0.194	0.269	0.329	0.325	0.121	0.111	0.193	0.294	0.380	0.432
	0.141	0.159	0.205	0.283	0.365	0.342	0.147	0.126	0.226	0.309	0.401	0.463

	λ_n											
	$K_n = 7$						$K_n = 11$					
$J_n - K_n$	0.2	0.1	0.01	0.001	1e-4	1e-5	0.2	0.1	0.01	0.001	1e-4	1e-5
1	0.004	0.003	0.008	0.027	0.010	0.324	0.002	0.004	0.018	0.031	0.100	0.056
	0.194	0.219	0.306	0.809	4.223	8.404	0.033	0.049	0.201	0.988	2.477	6.113
	0.198	0.222	0.313	0.835	4.233	8.729	0.036	0.053	0.219	1.019	2.578	6.169
3	0.009	0.014	0.001	0.003	0.131	0.172	0.002	0.005	0.026	0.078	0.042	0.210
	0.205	0.181	0.306	0.940	2.023	5.660	0.023	0.044	0.198	0.757	2.168	6.379
	0.214	0.195	0.307	0.943	2.153	5.831	0.025	0.049	0.224	0.835	2.210	6.589
6	0.012	0.000	0.012	0.029	0.043	0.032	0.004	0.007	0.038	0.030	0.036	0.063
	0.146	0.157	0.298	0.664	1.117	1.477	0.029	0.038	0.177	0.586	1.385	2.182
	0.158	0.158	0.310	0.693	1.160	1.509	0.033	0.045	0.215	0.616	1.421	2.245
11	0.025	0.015	0.028	0.027	0.039	0.026	0.008	0.015	0.024	0.041	0.046	0.020
	0.120	0.135	0.192	0.356	0.532	0.623	0.031	0.050	0.146	0.325	0.725	0.887
	0.146	0.150	0.219	0.384	0.571	0.649	0.039	0.065	0.170	0.366	0.772	0.907

Note: for each value of (λ_n, K_n, J_n) we report in each cell, the integrated squared bias, integrated variance and integrated MSE (divided by 100).

Table 2: Monte Carlo Study - Integrated MSE of Sieve IV estimator of $C'_{0q}(q)$

	λ_n											
	$K_n = 4$						$K_n = 5$					
$J_n - K_n$	0.2	0.1	0.01	0.001	1e-4	1e-5	0.2	0.1	0.01	0.001	1e-4	1e-5
1	0.19	0.14	0.18	0.12	0.18	0.13	0.14	0.12	0.14	0.17	0.13	0.16
	0.42	0.24	0.36	0.48	0.54	0.29	0.30	0.26	0.28	0.32	0.26	0.32
	0.61	0.39	0.54	0.59	0.72	0.41	0.44	0.39	0.42	0.50	0.39	0.49
3	0.11	0.16	0.12	0.12	0.15	0.17	0.14	0.17	0.13	0.13	0.14	0.16
	0.34	0.41	0.27	0.33	0.34	0.44	0.35	0.34	0.26	0.28	0.18	0.18
	0.45	0.57	0.39	0.45	0.49	0.61	0.48	0.51	0.39	0.41	0.32	0.34
6	0.11	0.12	0.18	0.11	0.11	0.10	0.10	0.10	0.09	0.12	0.07	0.09
	0.21	0.24	0.28	0.21	0.25	0.21	0.19	0.24	0.21	0.23	0.13	0.19
	0.32	0.36	0.46	0.32	0.36	0.31	0.29	0.34	0.30	0.35	0.20	0.28
11	0.16	0.11	0.13	0.11	0.11	0.07	0.08	0.07	0.08	0.07	0.10	0.08
	0.23	0.18	0.23	0.38	0.24	0.24	0.14	0.12	0.21	0.16	0.22	0.20
	0.39	0.30	0.36	0.49	0.35	0.30	0.21	0.20	0.28	0.23	0.32	0.28

	λ_n											
	$K_n = 7$						$K_n = 11$					
$J_n - K_n$	0.2	0.1	0.01	0.001	1e-4	1e-5	0.2	0.1	0.01	0.001	1e-4	1e-5
1	1.02	0.78	1.09	0.84	1.26	0.71	2.06	1.97	1.88	2.15	2.09	2.40
	0.95	0.69	1.20	0.79	1.04	0.82	0.99	0.66	0.96	0.80	0.74	1.08
	1.97	1.47	2.30	1.64	2.30	1.53	3.05	2.63	2.83	2.94	2.83	3.48
3	0.73	0.70	0.74	0.97	0.69	0.83	1.54	1.85	2.03	1.65	1.87	1.81
	0.56	0.76	0.91	0.75	0.57	0.92	0.81	0.97	0.84	0.66	1.24	0.86
	1.29	1.46	1.66	1.72	1.26	1.75	2.34	2.82	2.86	2.30	3.11	2.67
6	0.64	0.50	0.53	0.53	0.51	0.52	1.05	0.82	1.08	1.08	1.37	1.19
	0.55	0.44	0.56	0.32	0.52	0.47	0.64	1.66	1.26	0.82	0.59	0.60
	1.18	0.94	1.09	0.85	1.02	0.98	1.69	2.48	2.34	1.90	1.96	1.79
11	0.40	0.42	0.66	0.40	0.38	0.58	0.67	0.89	0.72	0.75	0.90	0.88
	0.40	0.39	0.91	0.41	0.33	0.42	0.75	0.67	1.40	0.80	1.29	0.53
	0.80	0.80	1.57	0.80	0.71	1.00	1.42	1.57	2.11	1.55	2.18	1.41

Note: for each value of (λ_n, K_n, J_n) we report in each cell, the integrated squared bias, the integrated variance and the integrated MSE (divided by 100).

Table 3: Monte Carlo Study - Integrated MSE of the nonparametric estimator of $F(\theta)$

	λ_n											
	$K_n = 4$						$K_n = 5$					
$J_n - K_n$	0.2	0.1	0.01	0.001	1e-4	1e-5	0.2	0.1	0.01	0.001	1e-4	1e-5
1	0.03	0.03	0.04	0.04	0.05	0.08	0.04	0.05	0.04	0.05	0.08	0.11
	0.03	0.03	0.04	0.05	0.16	0.23	0.03	0.03	0.04	0.06	0.19	0.41
	0.07	0.06	0.08	0.09	0.21	0.31	0.07	0.08	0.08	0.11	0.27	0.51
3	0.03	0.03	0.04	0.04	0.06	0.05	0.04	0.04	0.04	0.06	0.07	0.08
	0.03	0.03	0.04	0.05	0.08	0.12	0.04	0.04	0.03	0.06	0.15	0.15
	0.06	0.06	0.07	0.09	0.14	0.17	0.08	0.07	0.08	0.12	0.22	0.23
6	0.04	0.03	0.03	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.06	0.06
	0.03	0.03	0.03	0.04	0.04	0.05	0.04	0.03	0.03	0.04	0.06	0.08
	0.07	0.06	0.06	0.08	0.08	0.09	0.08	0.07	0.07	0.09	0.12	0.14
11	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.04
	0.02	0.03	0.03	0.03	0.04	0.04	0.03	0.03	0.03	0.04	0.05	0.04
	0.05	0.06	0.06	0.07	0.08	0.07	0.07	0.07	0.07	0.09	0.09	0.09

	λ_n											
	$K_n = 7$						$K_n = 11$					
$J_n - K_n$	0.2	0.1	0.01	0.001	1e-4	1e-5	0.2	0.1	0.01	0.001	1e-4	1e-5
1	0.05	0.05	0.05	0.06	0.10	0.20	0.08	0.08	0.07	0.07	0.12	0.25
	0.04	0.03	0.04	0.05	0.22	0.37	0.03	0.02	0.03	0.06	0.21	0.38
	0.09	0.08	0.09	0.11	0.32	0.57	0.11	0.10	0.11	0.13	0.33	0.63
3	0.06	0.05	0.07	0.07	0.09	0.15	0.07	0.07	0.07	0.07	0.16	0.44
	0.03	0.03	0.04	0.07	0.14	0.33	0.03	0.03	0.03	0.05	0.19	0.42
	0.09	0.08	0.11	0.14	0.23	0.49	0.10	0.10	0.10	0.13	0.36	0.87
6	0.06	0.06	0.05	0.06	0.08	0.11	0.08	0.07	0.07	0.08	0.16	0.37
	0.03	0.03	0.04	0.05	0.10	0.15	0.02	0.03	0.03	0.05	0.16	0.24
	0.09	0.08	0.09	0.11	0.18	0.26	0.10	0.10	0.10	0.13	0.32	0.61
11	0.06	0.05	0.05	0.06	0.09	0.08	0.06	0.06	0.08	0.09	0.14	0.22
	0.03	0.02	0.04	0.04	0.06	0.07	0.03	0.03	0.03	0.05	0.14	0.18
	0.09	0.08	0.09	0.10	0.15	0.15	0.09	0.09	0.11	0.13	0.28	0.39

Note: for each value of (λ_n, K_n, J_n) we report in each cell, the integrated squared bias, integrated variance and integrated MSE (divided by 100).

The integrated squared bias is higher than for the estimation of the marginal surplus as the estimation is now based on quantile IV methods. However, the variance is smaller and not very sensitive to the tuning parameters.

5.3 Estimation of the weight parameter γ and the parametric component

Finally we present in Table 4 the same results for the parameters involved in the simulation. We only report them for $\lambda_n = 0.2$ or 0.1 and $J_n - K_n = 3$ as the MSE is of the same order of magnitude across the different cases.[25] β_Z is the coefficient related to the parametric component in the marginal surplus, β_W to the component in the marginal cost and γ is the bargaining weight (assumed here unique) estimated from the following equation:

$$H(\theta, \gamma) = \theta + (1 - \gamma)\frac{F(\theta)}{f(\theta)}.$$

Table 4: Monte Carlo Study - Estimated bias, variance and MSE of the parameters of the model, $J_n - K_n = 3$.

		λ_n							
		$K_n = 4$		$K_n = 5$		$K_n = 7$		$K_n = 11$	
γ	Bias	0.010	0.059	-0.038	-0.056	-0.074	-0.033	-0.034	-0.034
	Variance	0.035	0.039	0.028	0.028	0.030	0.030	0.028	0.023
	MSE	0.035	0.042	0.001	0.001	0.001	0.001	0.001	0.001
β_Z	Bias	0.008	0.001	0.016	0.003	0.011	0.002	0.003	0.005
	Variance	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000
	MSE	0.001	0.001	0.005	0.006	0.009	0.003	0.002	0.001
β_W	Bias	-0.012	-0.011	-0.008	-0.005	-0.002	0.007	-0.007	-0.011
	Variance	0.009	0.011	0.005	0.004	0.003	0.005	0.001	0.004
	MSE	0.009	0.011	0.005	0.004	0.003	0.005	0.001	0.004

Note: for each value of (λ_n, K_n) we report in each cell, the mean bias, the variance and the MSE.

The estimation of γ is much noisier than for the other parameters but this is still good. In the estimation procedure, we use $K_n = 7$ and $\lambda_n = 0.1$.

[25]The full results are available upon request.

6 Empirical Application

We apply the estimation method detailed above to the case of water utilities in France. We first present the data set, before estimating our structural model of regulation. The counterfactuals are presented in the next section.

6.1 The Data

The production and distribution of water to households in France are decided at a local level. A survey has been conducted by the French environment institute (IFEN, Institut Français de l'Environnement) amongst local municipalities in metropolitan France.[26] All municipalities with a population of more than 10000 are in the survey and the sampling rate is decreasing with the population. The sample is therefore representative from the population of French local communities (36203 local municipalities in 2004).

We select the observations for which the chosen mode is either operating through a public company (*régie*) or through a private operator with a lease contract (this is the majority of the cases). Then we select the typical contracts, i.e. the two-part tariff contracts (around 95% of the observations). We also merge our dataset with an administrative dataset which reports the median income in a given municipality. For anonymity reasons, it is not reported for city size less than 150 inhabitants. Our sample does not contain such small villages. In the end, we have 3959 observations which each represent one municipality. Table 5 presents the management mode of municipalities ranked according to their population. When delegation is the chosen mode, it is in most of the cases conducted by one of the three major firms of the sector: Veolia-Environnement, Lyonnaise des Eaux-Suez Environnement, and Saur-Cise Veolia is the leading firm in the private sector, Saur is more present in rural areas. Observe also that public management is more frequent in small municipalities.

In our dataset, we observe the fix fee of the contract A and the variable part p paid by meter-cube consumed. The overall quantity q is also observed. The characteristics of the network are also given (length and density, quality of the water used as an input, treatment applied before distribution...). Some additional information related to the local municipality (population, median income, size of the city, local weather measures) are also recorded. We present in Table 6 some descriptive statistics of these variables. For continuous variables, we present the median and the 1st and 3rd quartiles (Q_1 and Q_3) which are respectively

[26]Corsica and overseas departments are excluded from our study.

Table 5: Management mode by municipality sizes

	Population N of the different municipalities					
	$N < 400$	$400 \leq N$ $N < 1\,000$	$1\,000 \leq N$ $N < 2\,000$	$2\,000 \leq N$ $N < 3\,500$	$3\,500 \leq N$ $N < 10\,000$	$10\,000 \leq N$
Number n of municipalities						
n	813	652	490	356	1210	636
Public	0.53	0.47	0.42	0.33	0.3	0.27
Delegation	0.47	0.53	0.58	0.67	0.7	0.73
Providers in the delegation mode						
Veolia	0.33	0.32	0.37	0.4	0.44	0.49
Suez	0.15	0.2	0.21	0.2	0.24	0.23
Saur	0.36	0.33	0.29	0.27	0.15	0.07
Other	0.16	0.15	0.13	0.13	0.17	0.2

robust measures of the location and the dispersion of the empirical distribution. For the two qualitative variables related to the quality of the water and the treatment (they are both recoded with three modalities), we report the empirical frequencies. The fix fee represents around one quarter of the transfer paid to the provider. The variable part is around 1.25 euro per m^3. The total bill charged to the consumer is much higher than what is charged by the firm. First, there are taxes paid to the local water agencies and to the state (value added tax). Then, mechanically, the amount given to the entity in charge of the waste collection is charged proportionally to the quantity consumed. This discrepancy between the prices is taken into account in our analysis.

The price paid by the consumer is generally lower in a local municipalities which are publicly managed. On the other hand, private firms have a higher probability to operate in more dense cities and to distribute water which requires heavier treatment to be drinkable.

Finally, we also collect variables which characterize the environment of a given city, i.e. number of houses, temperature, amount of rain, having in mind that water can be used for watering gardens.

6.2 Estimation of the model

We now follow the estimation strategy that is exposed in Section 5.

Table 6: Descriptive statistics of the main variables

	All mode			Public			Delegation		
	median	Q_1	Q_3	median	Q_1	Q_3	median	Q_1	Q_3
q per household	119.05	98.46	143.55	116.28	96.15	142.79	120.99	99.82	144.37
Fixed fee firm	33.62	19.86	53.84	27	15.54	46.74	37.86	22.63	59.26
Fixed fee	45.2	23.42	77.11	37.83	18.64	66.19	49.23	27	86.29
Fixed fee firm, tax incl.	35.01	20.96	56.35	28.3	16.07	48.87	40.12	23.96	62.59
Fixed fee, tax incl.	47.61	24.82	81	39.38	19.45	69	51.88	28.89	91.02
variable price firm	1.25	1.01	1.53	1.11	0.88	1.35	1.35	1.09	1.62
variable price	2.14	1.48	2.68	1.82	1.17	2.36	2.31	1.76	2.8
variable price, tax incl.	2.24	1.54	2.82	1.89	1.21	2.48	2.42	1.85	2.96
Transfert firm	189.73	155.22	231.96	167.2	133.33	198.23	205.28	170.8	248.95
Bill, tax incl.	323.11	246.01	401.19	272.68	190.74	341.86	353.78	284.34	432.92
Population	2412	540	6383	1214	368	4757	3616	776	7509
Median income	23943	20222	28858	23129	19730	27687	24457	20546	29790
Household size	2.61	2.41	2.78	2.57	2.39	2.75	2.63	2.42	2.8
% of Houses	71.29	53.97	82.38	71.3	55.28	82.38	71.27	52.7	82.38
% of Secondary res	4.39	1.67	12.61	5.88	2.04	15.81	3.64	1.51	10.26
% of Pop under 20	25.03	22.13	27.63	24.55	21.39	27	25.31	22.49	27.99
Temperature	25.1	23.6	26.3	25.1	23.6	26.3	25.1	23.6	26.4
Rain (mm)	194	151	223	200.5	151	251	190	149	212
Sunshine (in h)	581	547	665	591	547	669	579	535	665
Population density	138.87	40.88	480.6	80.37	25.93	261.26	184.27	56.89	631.4
Network characteristics									
network length (in km)	35	14	70	26	9	60	40	18	75
Nb connections	1024	264	2507	600	199	1950	1380	347	2805
Total tank volume	0	0	350	0	0	500	0	0	200
Nb of sensors	1	1	3	1	1	3	1	1	3

	All mode Frequency	Public Frequency	Delegation Frequency
Basic treatment	0.52	0.53	0.51
High treatment	0.22	0.13	0.27
Deep water	0.72	0.77	0.68
Superficial water	0.28	0.23	0.32

6.2.1 Fitting with Real-World Practices

In the real world, the observed per-unit prices paid by consumers are in fact affected by various taxes. Indeed, we can decompose the price p paid by consumers as:

$$p = (p_1 + p_2)(1 + \tau)$$

where p_1 is the price received by the producer, p_2 is a price paid to another party to finance waste water and τ is an ad-valorem tax imposed by the State. In the sequel, we will take a partial equilibrium approach, taking p_2 and τ as given (which may vary across municipalities) and we will sometimes denote $p = P(p_1)$.

Taking into account this specification of the prices, we can rewrite the agent's and the principal's objectives respectively as:

$$\mathcal{U}(\theta, \varepsilon, p_1, A) = A + p_1 D(P(p_1), \varepsilon) - \theta C_0(D(P(p_1), \varepsilon))$$

and

$$\mathcal{W}(\theta, \varepsilon, p, p_1, A) = S(D(P(p_1), \varepsilon), \varepsilon) - A - P(p_1)D(P(p_1), \varepsilon) + \gamma\mathcal{U}(\theta, \varepsilon, p_1, A)$$

or

$$\mathcal{W}(\theta, \varepsilon, p, p_1, A) = S(D(P(p_1), \varepsilon), \varepsilon) - \theta C(D(P(p_1), \varepsilon)) + (p_1 - P(p_1))D(P(p_1), \varepsilon) - (1-\gamma)\mathcal{U}(\theta, \varepsilon, p_1, A)$$

The fact that $P(p_1) \neq p_1$ creates a discrepancy between consumers expenditures and producer's revenues. Yet, the optimization of the principal's problem proceeds as above and the system (13)-(14)-(15) should now be replaced by the new set of optimality conditions:

$$S'(q) = p, \tag{27}$$

$$p_1 - \frac{\tau D(p)}{D'_p(p)} = H(\theta, \gamma) C'_{0q}(q), \tag{28}$$

$$A'(p)(1+\tau) = -q - \left(p_1 - \theta C'_{0q}(q)\right) D'_p(p)(1+\tau). \tag{29}$$

6.3 Estimation of the demand function

The individual demand function is estimated with different specification though we keep working with the log-log model in the following.

$$\log(D(p, Z, \varepsilon)) = \beta_p \log p + \beta_Z' Z + \varepsilon.$$

As usual in the case of a system of simultaneous equations, equilibrium quantities and prices are co-determined and, hence, p is endogenous. To circumvent this issue , we instrument p by a vector W of exogenous cost shifters ($\mathbb{E}[\varepsilon | Z, W] = 0$) and we estimate the demand parameters by a simple IV regression. The instruments are the following: treatment of the water and characteristics of the water pumped before treatment. These variables increase the cost of distribution by requiring more advanced techniques to make the water drinkable but are independent from the demand shocks. The results are presented in Table 7, with standard errors in parenthesis.

The price elasticity is estimated at -0.171 (0.073), which is in the usual intervals estimated for residential water in France (see Reynaud, 2003). Observe that it is the total price charged to the consumer, p, which is used here, not the one charged to the firm, p_1.

Except for the percentage of the houses, the control variables have the expected sign. Richer cities consume more water and the climate conditions matter. Finally, everything else equal, cities with more secondary residences consume less. The Sargan test is not rejected with a p-value of 6%.

6.4 Estimation of the cost function

In our data, the size of the different cities vary a lot. Therefore, we make the following assumption about the cost function:

$$C(q_{tot}, W, N) = C_0(q_{tot}/N) N^{\alpha} \exp(\beta_W^{\top} W),$$

in which N is the number of households of the city and W, the other control variables (quality of the water, density of the network, treatment). In the following, $q = q_{tot}/N$ denotes the mean quantity per household consumed. α measures the return to scale to deliver the same quantity of water per household in a bigger city. We should expect α to be lower than one. Fundamentally, we make the assumption that the technology is the same across all providers but that they differ from one city to another through the observed heterogeneity, i.e., quality of the water pumped before treatment, density of the network, level of chemical treatment and through θ which measures the efficiency of the firm in charge of the distribution of the water.

Table 7: Demand estimation

	Dependent variable:	
	q	log(q)
	(1)	(2)
p	-8.837^{*}	
	(4.560)	
log(p)		-0.171^{**}
		(0.073)
income	1.314^{***}	0.012^{***}
	(0.120)	(0.001)
household size	31.682^{***}	0.204^{***}
	(3.729)	(0.031)
% of houses	-0.320^{***}	-0.003^{***}
	(0.050)	(0.0004)
% secondary res	-0.655^{***}	-0.007^{***}
	(0.095)	(0.001)
temperature 2004	1.631^{***}	0.014^{***}
	(0.401)	(0.003)
rain 2004	-0.059^{***}	-0.001^{***}
	(0.010)	(0.0001)
sunshine 2001	0.009^{***}	0.0001^{***}
	(0.003)	(0.00002)
pop 0-20	-0.864^{***}	-0.005^{**}
	(0.241)	(0.002)
constant	40.895^{*}	4.141^{***}
	(23.661)	(0.159)
Observations	3,690	3,690
R^2	0.229	0.249
Adjusted R^2	0.228	0.248
Note:	*p<0.1; **p<0.05; ***p<0.01	

As explained above, Equation 28 rearranged to take into account the price paid to the other stakeholders and the tax implies that for any quantile $\lambda \in]0,1[$, we have:

$$\mathbb{P}\left(p_1 - \tau \frac{D(p,Z,\varepsilon)}{\partial D/\partial p D(p,Z,\varepsilon)} \leq H(\theta_\lambda,\gamma) C_0'(q) N^{\alpha-1} \exp(\beta_W^\top W) | Z, W, N\right) = \lambda, \quad (30)$$

with θ_λ being the corresponding quantile of θ and p_1 and p are the prices charged respectively by the firm and the other stakeholders, before tax. $C_{0q}'(q,W)$ is identified up to a scale under an assumption of completeness of Z. In practice, we first estimate $\log C_{0q}'(q,N,W)$ before changing the scale to ensure that the median of θ is equal to 1 for publicly managed firms. Equation 30 can also be expressed as a conditional moment condition:

$$\mathbb{E}\left(\mathbf{1}\left\{p_1 - \tau \frac{D(p,Z,\varepsilon)}{\partial D/\partial p D(p,Z,\varepsilon)} \leq H(\theta_\lambda,\gamma) C_0'(q) N^{\alpha-1} \exp(\beta_W^\top W)\right\} - \lambda | Z, W, N\right) = 0.$$

This conditional moment equality can be easily be transformed into a moment equality to estimate the marginal cost function and the quantiles $H(\theta_\lambda,\gamma)$ through a GMM procedure. As it has been highlighted in the literature, in practice, the global minimization of the GMM criterion is arduous due to the discontinuous indicator function $\mathbf{1}\{\cdot\}$. Following deCastro et al. (2019), we smooth the indicator function in order to facilitate the minimization. In the estimation, 10 quantiles are used[27]. We run simultaneously the estimation for both the private and the public providers assuming that they face the same marginal cost but that each type of firm has its own distribution of θ. We run the estimation 10 times with different starting values and select the estimates which yield the lowest objective function. The results are presented in Table 8.

Table 8: GMM estimation of $C_{0q}'(q,W)$

	estimate	standard error
constant	-2.9621	4.7e-02
log(q)	0.7806	3.1e-03
log(N)	-0.1035	5.1e-04
basic treatment	0.7625	3.6e-03
high treatment	1.1660	3.6e-03
deep water	-0.6762	2.9e-03
population density	0.0002	3.1e-07

[27] The quantiles are 0.05,0.15,0.25,...

Observe that, as expected, $\hat{\alpha} = 1 - 0.1035 < 1$. Again, the estimates have the expected signs because more treatments increase the cost of the m^3 of water. Pumping deep water is less costly, because on average this water is cleaner.

6.4.1 Estimation of the distribution of types

In this part, we estimate the types θ taking into account the fact that there is no asymmetric information for public firms and that empirical evidence suggest that $\gamma = 0$ when the city designs its menu of contracts. As a matter of fact, it is politically risky to give a higher rent to the private operator. Having at hand, the estimate of the marginal cost function, we do the following:

1. From the estimates of $C'_{0q}(q)$, we estimate the values θ for the public firms using Equation (28) with θ in replacement of $H(\theta, \gamma)$.
2. We readjust our estimates to fit with the normalization that the median of the θ is equal to 1. $C'_{0q}(q)$ represents the variable cost function for a median public firm.
3. We estimate
$$H(\theta) = \theta + \frac{F(\theta)}{f(\theta)},$$
from Equation (28).
4. We invert the expression above to recover θ from the observation of $H(\theta)$. Let $G(\cdot)$ (resp. $g(\cdot)$) the cumulative distribution function (resp. p.d.f) of the $h = H(\theta)$. Standard arguments allow us to back cast θ by expliciting the inverse mapping from $h = H(\theta)$ to θ well-known in the literature on auctions (see, for example Guerre et al. (2000)). $\theta(h) = h - \frac{1}{G(h)} \int_{\underline{h}}^{h} G(x)dx$.

Table 9 compare our estimates of the distribution of the type within each population of publicly managed and private firms. The distribution of private types (see Figure 2) taking into account the observed heterogeneity and the asymetric information is shifted toward the left and, also, more concentrated.

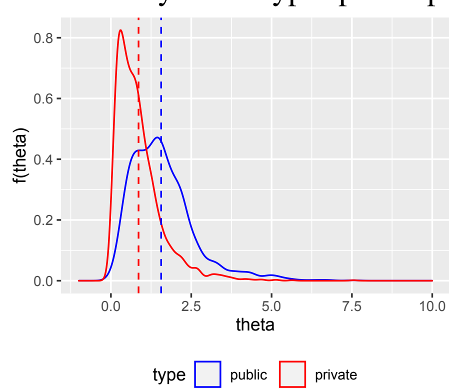
7 Counterfactuals

In this part, we run two counterfactuals of interest.

Table 9: Estimation of the distribution of types

	1	2
1%	0.21	0.01
5%	0.41	0.07
15%	0.63	0.23
25%	0.83	0.35
50%	1.35	0.70
75%	2.00	1.14
85%	2.45	1.45
95%	3.61	2.22
99%	5.20	3.51

Figure 2: Density of the types public/private



7.1 Complete versus Asymmetric Information

7.2 Private versus Public Ownership

Upper bound on the value of investment. Denote now by F_{pr} and F_{pu} the cumulative distributions of cost parameters under the private and the public scenario respectively. By investing an amount I ex ante (i.e., before the realization of the cost parameters), a private firm can shift the types distribution from F_{pu} to F_{pr}. As argued by Riordan (1990), Laffont and Tirole (1993, Chapters 3 and 17) and Schmidt (1996), asymmetric information and the prospects of getting some information rent under private ownership is here the sole engine of investment. Instead, a public firm because it earns no rent under complete information won't invest. The terms of the trade-off are clear. By relying on private firms, the types distribution is shifted to the left while the counterpart of such efficiency gains is that information rent should be given up and prices end up above marginal costs.

Following Riordan (1990), we assume that this investment is non-verifiable, so that although whether it is incurred or not is perfectly anticipated by the local authority at the time of designing the contract, this contract cannot directly influence that decision. Such investment is meant for all the know-how, organizational and technological innovations, and expertise that may be done and that can hardly be explicitly included into a contract.

Our purpose in this section is to get an upper bound on the investment incurred by private firms. Given the contracting scenario so depicted, investment takes place whenever the following incentive constraint holds:

$$E_{\varepsilon}\left(\int_{\underline{\theta}}^{\overline{\theta}}(f_{pr}(\theta)-f_{pu}(\theta))U_{pr}(\theta,\varepsilon)d\theta\right)\geq I.$$

Integrating by parts the left-hand side, we may rewrite this incentive constraint as:

$$E_{\varepsilon}\left(\int_{\underline{\theta}}^{\overline{\theta}}(F_{pr}(\theta)-F_{pu}(\theta))C(D(p_{pr}(\theta),\varepsilon),\varepsilon)d\theta dG(\varepsilon)\right)\geq I.$$

The left-hand side above is the firm's expected gains when shifting the type distribution towards from F_{pu} to F_{pr} when the regulator anticipates that the investment is incurred and, as a result, offers a contract inducing the output profile $D(p_{pr}(\theta),\varepsilon)$ that prevails under a scenario of private ownership. Of course, this term is positive whenever $F_{pr}(\theta)\geq F_{pu}(\theta)\geq 0$ for all θ, a first-order stochastic dominance condition that holds from our previous empirical findings. The right-hand side is just the investment outlay.

Our previous econometric analysis allows us to compute the upper bound on any such investment because all terms on the left-hand side have been previously derived.

8 Alternative Formulations

This section discusses alternative formulation for the contracting environment and how our identification procedure applies or not. While Section 8.1 highlights environment where our basic identification strategy would still be useful, Section 8.3 shows the importance of the kind of contracts observed (relying on fixed fees and per-unit of consumption prices) to get such positive results.

8.1 Non-Separability in the Cost Function

Let us now suppose that the cost function is no longer separable and can be more generally written as $C(q, \theta)$, i.e., costs are not necessarily linear in θ. We assume that $C(q, \theta)$ is increasing and concave in q with on top $C'_\theta > 0$ (operators with lower types produce at lower costs) and the Spence-Mirrlees condition $C''_{q\theta} > 0$ (those operators also produce at lower marginal costs) being satisfied. It is routine to check tha the system (9)-(11)-(12) now becomes:

$$A'_p(p(\theta,\varepsilon),\varepsilon) = -\left(p(\theta,\varepsilon) - C'_q\left(D(p(\theta,\varepsilon),\varepsilon),\theta\right)\right) D'_p(p(\theta,\varepsilon),\varepsilon) - D(p(\theta,\varepsilon),\varepsilon). \quad (31)$$

$$p(\theta,\varepsilon) = C'_q\left(D(p(\theta,\varepsilon),\varepsilon),\theta\right) + (1-\gamma)\frac{F(\theta)}{f(\theta)}C''_{q\theta}\left(\left(D(p(\theta,\varepsilon),\varepsilon)\right),\theta\right), \quad (32)$$

$$S'_q(D(p(\theta,\varepsilon),\varepsilon),\varepsilon) = p(\theta,\varepsilon), \quad (33)$$

8.2 Identification of a More General Specification

We now prove the identification of this more general model with non-separability in both the surplus and cost functions (Equations (31)-(32)-(33)). First, observe that the identification of the marginal surplus is similar in this general model to what we have done above.

For the marginal cost function, we need an additional normalization like for the surplus function. We shall assume that θ is assumed to be uniformly distributed on $[0, 1]$ and we therefore identify the marginal cost function (given W) for the α-th quantile firm in an efficiency ranking (ranked from the most efficient , $\alpha = 0$, to the less efficient, $\alpha = 1$).

Equation (31) identifies $c'_q\,(q, \alpha, W)$ for any quantile α of θ under an assumption of completeness of $\tilde{Z}$ in q. We can indeed derive a moment condition similarly than what have been done in the seprable case earlier. We are indeed able to know, after having derived the marginal surplus function for any quantile of ε the quantity $d = p(\theta, \varepsilon) - \frac{A'_p(p(\theta,\varepsilon),\varepsilon) - D(p(\theta,\varepsilon),\varepsilon)}{D'_p(p(\theta,\varepsilon),\varepsilon)}$, which is equal to the marginal cost (given W) for some (unknown) quantile of θ. Moreover the monotonicity of the marginal cost in θ (given W) ensures that (see above for a similar derivation):

$$\mathbb{P}\left(\left(c'_q(\alpha, q, W) \leq d\right) | Z, W\right) = \alpha. \tag{34}$$

The completeness assumption ensures therefore the identification of any marginal cost function.

Coming back now to Equation (32), the only unknown quantity is now γ and we also recover the identification of the bargaining power in the non-separable case.

8.3 Contracting on Outputs

The fact that the model can be fully identified (at least when $\gamma = 0$) might seem quite surprising at first glance. Indeed, the literature on the structural estimations of incentive contracts has repeatedly failed to obtain such joint identification of cost functions and types distributions because the data available to the econometrician do not provide enough information. The key difference between our model and those of the existing literature, especially d'Haultfoeuille and Février (2020) and Luo et al. (2018), comes actually from the set of contracting variables available under different scenarios that fit with specific institutional environments.

In the regulatory context under scrutiny, for instance, consumers adopt a competitive behavior, expressing individual demand for water at the stipulated unit price while the operator stands ready to supply aggregate demand at that a market clearing price. Instead, had the principal being a single big customer dealing directly with the operator, contracting on the quantity bought would be a feasible option. This is such scenario that is analyzed in d'Haultfoeuille and Février (2020).

To facilitate comparison with Section 3.1, let us thus assume that contracts are based only on the operator's quantity. In this hypothetical scenario where output could be directly contracted upon, a nonlinear contract might now specify a payment $T(q)$ to the operator if he offers a volume q.[28] For the sake of simplifying exposition, we also assume in this section that there is no fluctuations in demand (i.e., $\varepsilon \equiv 0$) and $\gamma = 0$. Under those conditions, we already know from Section 4.1 that our base model is actually identified.

Theoretical results. Mimicking some of the earlier steps of our above analysis, we can rewrite the operator's information rent and optimal supply respectively as:

$$U(\theta) = \max_q T(q) - \theta C_0(q) \tag{35}$$

and

$$q(\theta) = \arg\max_q T(q) - \theta C_0(q). \tag{36}$$

From there, it again follows from incentive compatibility that $U(\cdot)$ is absolutely continuous and admits an integral representation as:

$$U(\theta) \int_{\theta}^{\bar{\theta}} C_0\left(q(x)\right) dx \tag{37}$$

where again, we take into account that the operator's participation constraint is binding at $\bar{\theta}$ for the optimal contract. Observe that this rent is again greater as the operator is requested to produced more.

Under asymmetric information, an optimal contract maximizes the expected welfare of the municipality subject to incentive and participation constraints. From our above observations, that incentive feasible set can be summarized by constraints (37) and a monotonicity condition (namely $q(\theta)$ weakly decreasing in θ) that will be omitted in a first step and checked ex post on the solution to the solution of the so relaxed problem. Formally, this relaxed problem can be written as:

$$\max_{\{q(\cdot),U(\cdot)\}} \int_{\underline{\theta}}^{\bar{\theta}} \left[S(q(\theta)) - \theta C_0\left(q(\theta)\right) - U(\theta)\right] dF(\theta) \text{ subject to (37).}$$

Using (37), integrating by parts as we did above and optimizing pointwise leads to the following expression of the optimal output $q(\theta)$:

$$S'_q(q(\theta)) = \left(\theta + \frac{F(\theta)}{f(\theta}\right) C'_{0q}\left(q(\theta)\right). \tag{38}$$

[28]Observe that, even in this output scenario, the principal still does not observe costs. Otherwise, he could retrieve from the joint observation of costs and outputs information on the agent's type. See Laffont and Tirole (1986) for a theoretical model where such cost observation is available and yet, because of an extra moral hazard variable, asymmetric information still matters. Perrigne and Vuong (2011) for the corresponding empirical study.

The structural model is then defined also by appending to that optimality condition for the principal's problem a second optimality condition related to the agent's problem, namely:

$$T'_q(q(\theta)) = \theta C'_{0q}\left(q(\theta)\right). \tag{39}$$

From an economic viewpoint, whether outputs as here or prices as in our base line model are the relevant contracting variables leads to the same optimal outputs. The principal's optimality condition (38) is similar to that obtained in our base model (11) (for $\gamma = 0$) and downward output distortions also follow from replacing cost parameter by their virtual counterparts. Since optimal outputs in the two models are the same, the agent's information rents, consumer's net surpluses and welfares are also identical. This economic equivalence should come at no surprise. Whether the demand side of the market adopts a "command and control" approach to specify how much should be supplied to satisfy aggregate needs or whether individual consumers are left to express demand is equivalent since that demand side is not the source of any informational asymmetry.

Differences in the scope for identification. Although economically equivalent, the two contracting scenarios differ not only in terms of the set of observables that are available to the econometrician but also in terms of the possibilities left for identification. In other words, whether prices or quantities are controlled gives two alternative implementations of the same allocation which are equivalent from the point of view of the players in the regulatory game which are equally informed on cost and surplus functions and distribution functions.

Yet, those two implementations differ from the perspective of the econometrician. The econometrician must infer from market conduct some information that is common knowledge among players. At a rough level, this outside observer faces a much harder signal extracting problem. In the scenario where quantities are directly controlled, no information on marginal surplus can be learned from the firm's choices. Instead, when only prices are controlled, information on marginal surplus gets communicated to outside observer. This simple fact allows identification of the model.

More precisely, for the output scenario, the only observables are the output q and the overall payment $T(q)$. As in our base model, S'_q and C'_{0q} are not observed in the data so that the principal's optimality condition (38) cannot be used to retrieve information either on types or on their distribution. On top, per-unit consumption prices are by construction not observed, a key difference with our base model.

Defining an implicit per-unit consumption price for this output scenario as $p(\theta) = S'_q(q(\theta))$, we may follow a procedure similar to that developed in our base model and introduce a fictitious price-cost margin. Accordingly, we thus rewrite the principal's optimality condition (38) as a familiar condition:

$$\frac{p(\theta) - \theta C'_{0q}\left(q(\theta)\right)}{p(\theta)} = r(\theta, 0).$$

Even if such price is not observed, one could hope as in our base line model to retrieve the relevant information from the agent's optimality condition (39). However, this step is no longer possible here because the agent's optimality condition (39) now refers only to the marginal nonlinear price $T'_q(q(\theta))$ and not on the price as in our baseline scenario. Indeed, this marginal nonlinear price $T'_q(q(\theta))$ always differs from the per-unit consumption price $p(\theta) = S'_q(q(\theta))$ under asymmetric information, there is always a wedge between $T'_q(q(\theta))$ and $p(\theta) = S'_q(q(\theta))$, as:

$$p(\theta) - T'_q(q(\theta)) = \frac{F(\theta)}{f(\theta)} C'_{0q}\left(q(\theta)\right) > 0.$$

Thus, the agent's optimality condition does not bring new information on marginal surplus at the equilibrium point, making it impossible to identify this function.

In response to such impossibility, different alleys have been pursued in the literature. Luo et al. (2018) consider a model of "false moral hazard"[29] where costs, which results from innate types but also no verifiable effort variables, can be observed. Such observation of course brings extra information. Although positive results are now obtained, the analysis remains complex because the additional effort variable that characterizes such model must also be disentangled from innate types. Developing an approach tailored to the specificities of their data set, d'Haultfoeuille and Février (2020) work with some discrete heterogeneity on the demand side while still assuming that such information is observed by the econometrician.[30] Three points are in this case sufficient to recover identification. Finally, Luo et al. (2018) consider a nonlinear pricing model with fixed quantity which *de facto* determines marginal surplus at the equilibrium point. Let us conclude this section by pointing out an alternative route. Indeed, parametric assumptions on both the surplus function and the cost function, or the cost function and the firms' types could help identifying the model but, of course, at a risk of misspecification as these assumptions cannot be tested.

[29]In the parlance of the incentive literature, see Laffont and Martimort (2002, Chapter 7).

[30]This is in striking contrast with our assumption in the base line model that such heterogeneity is actually not observed.

9 Conclusion

In this paper, we consider a principal-agent model à la Baron-Myerson. We show that the contract negociation can be modeled by a choice within a menu of options of a two-part tariff scheme. Interestingly, we show that the model is non parametrically identified from the observation of the tariff (fixed fee and per unit price) and the quantity consumed.

We apply our methodology to the distribution of water in France, in which we quantify the amount of asymmetric information and its impact on the per-unit price charged.

Some analysis remain to be done. First, we would like to design a more flexible approach than the parametric estimates, borrowing from the recent literature on nonparametric quantile IV. Then, we proved that we the dsitribution of the γs could be identified. However it relies on the estimation of derivatives. We need to find a stable way to estimate it.

References

[1] Ai, and X. Chen (2003), "Moment Restrictions Containing Unknown Functions," *Econometrica*, 71: 1795-1843.

[2] Armstrong, M. and D. Sappington (2007), "Recent Developments in the Theory of Regulation," *Handbook of Industrial Organization, Vol. 3*, M. Armstrong and R. Porter eds. North-Holland.

[3] agnoli, C. and T. Bergstrom (2005), "Log-Concave Probability and its Applications," *Economic Theory*, 26: 445-469.

[4] Baron, D. (1989a), "Design of Regulatory Mechanisms and Institutions" in *Handbook of Industrial Organization*, vol. 2, 1347-1447.

[5] Baron, D. (1989b), "Regulation and Legislative Choice," *The RAND Journal of Economics,* 19: 467-477.

[6] Baron, D. and R. Myerson (1982), "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50: 911-930.

[7] Basov, S. (2005), *Multidimensional Screening,* Springer-Verlag.

[8] Bhattacharyya, A., T.R. Harris, R. Narayan and K. Raffiee (1995), "Specification and Estimation of the Effect of Ownership on the Economic Efficiency of the Water Utilities," *Regional Science and Urban Economics*, 25: 759-784.

[9] Bhattacharyya, A., E. Parker and K. Raffiee (1994), *An Examination of the Effect of Ownership on the Relative Efficiency of Public and Private Water Utilities*, Land Economics, 70; 197-209.

[10] Blundell, R., Chen, X. and D. Kristensen (2007), "Semi-Nonparametric IV Estimation of Shape Invariant Engel Curves," *Econometrica*, 75: 1613-1669.

[11] Brocas, I., K. Chan and I. Perrigne (2006), "Regulation under Asymmetric Information in Water Utilities," *American Economic Review,* 96: 62-66.

[12] Carrasco, M., Florens, J.-P. and E. Renault (2008), "Linear Inverse Problems in Structural Econometrics: Estimation Based on Spectral Decomposition and Regularization," in *Handbook of Econometrics,* Vol 6..

[13] Cazals, C., J.P. Florens and L. Simar (2002), "Nonparametric Frontier Estimation: a Robust Approach," *Journal of Econometrics*, 106: 1-25.

[14] Chen, X. and D. Pouzo (2008), "Efficient Estimation of Semiparametric Conditional Moment Models with Possibly Non-Smooth Moments," Cowles Foundation Discussion Paper 1640.

[15] Chernozhukov, V. and C. Hansen (2005), "An IV Model of Quantile Treatment Effect," *Econometrica,* 73: 245-261.

[16] Chernozhukov, V., Imbens, G. and W. Newey (2007), "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics,* 139: 4-14.

[17] Chesher, A. (2003), "Identification in NonSeparable Models", *Econometrica*, 71: 1405- 1441.

[18] Chesher, A., (2007), "Identification of Nonadditive Structural Functions," in *Advances in Economics and Econometrics, Theory and Applications: 9th World Congress of the Econometric Society,* Blundell, R., Persson, T. and W. Newey, Eds., Vol. 3, Cambridge University Press, Cambridge.

[19] Chong, E., F. Huet, S. Saussier and F. Steiner (2006), "Public-Private Partnerships and Prices: Evidence from Water Distribution in France," *Review of Industrial Organization* 29: 149-169.

[20] Commissariat Général au Developpement Durable (2010), *Le point Sur*, Vol. 67.

[21] Darolles, S., Fan, Y., Florens, J.-P. and E. Renault (2011) "Nonparametric Instrumental Regression", *Econometrica*, 79: 1541-1565.

[22] de Castro, L., Galvao, A. F., Kaplan, D. M., and Liu, X. (2019), "Smoothed gmm for quantile models", *Journal of Econometrics*, 213, 121-144.

[23] Estache, A. and M. Rossi (1999), "Comparing the Performance of Public and Private Water Companies in Asia and Pacific Region: What a Stochastic Costs fFontier Shows," *World Bank working paper* WPS 2152, Washington, DC.

[24] Gagnepain, P. and M. Ivaldi (2002), "Incentive Regulatory Policies: The Case of Public Transit Systems in France," *The RAND Journal of Economics,* 33: 605-629.

[25] Gagnepain, P., M. Ivaldi and D. Martimort (2013), ”The Cost of Contract Renegotiation: Evidence from the Local Public Sector,” *American Economic Review*, 103, 2352-2383.

[26] Garcia, S. and A. Thomas (2001), “The Structure of Municipal Water Supply Costs: Application to a Panel of French Local Communities,” *Journal of Productivity Analysis*, 16: 5-29.

[27] Hall, D. (2000), “Public Choice and Water Rate Design, ” in *The Political Economy of Water Pricing Reforms*, Ariel Dinar (ed.): 189-212. Oxford University Press, Oxford.

[28] d’Haultfoeuille X. and P. Fevrier (2020), ”The provision of wage incentives: A structural estimation using contracts variation”, *Quantitative Economics*, 11, 349-397.

[29] IFEN (Institut Français de l’Environnement) (2006), “Les Services Publics de l’Eau en 2004 - Volet Eau Potable,” Paris.

[30] Koenker, R. (2005), “Quantile Regression,” Econometric Society Monographs, 38, Cambridge University Press, Cambridge.

[31] Koenker, R., Ng., P. and S. Portnoy (1994), “Quantiles Smoothing Splines,” *Biometrika* 81: 673-680.

[32] Laffont, J.J. and D. Martimort (2002), *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press.

[33] Laffont, J.J. and J. Tirole (1993), *A Theory of Incentives in Procurement and Regulation*, MIT Press.

[34] Lavergne P. and A. Thomas (2005), “Semiparametric Estimation and Testing in Models of Adverse Selection”, *Empirical Economics,* 30: 1-22

[35] Lim, C. and A. Yurukoglu (2018), ”Structural analysis of nonlinear pricing”, *Journal of Political Economy*, 126, 2523-2568.

[36] Luo, Y., I. Perrigne, and Q. Vuong (2018), ”Structural analysis of nonlinear pricing”, *Jour- nal of Political Economy*, 126, 2523-2568.

[37] Myerson, R. (1982), “Optimal Coordination Mechanisms in Generalized Principal-Agent Models,”*Journal of Mathematical Economics,* 10: 67-81.

[38] Noldeke, G., and L. Samuelson, (2007), ”Optimal Bunching Without Optimal Control”, *Journal of Economic Theory*, 134: 405-420.

[39] Perrigne, I., and Q. Vuong (2011), “Nonparametric Identification of a Contract Model with Adverse Selection and Moral Hazard”, *Econometrica*, 79: 1499-1539.

[40] Reynaud, A. (2003), ”An Econometric Estimation of Industrial Water Demand in France”, *Environmental and Resource Economics*, 25, 213-232.

[41] Robinson, P. (1988), “Root-n Consistent Semiparametric Regression,” *Econometrica*, 56: 931-954.

[42] Wolak, F. (1994), “An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction,” *Annales d’Economie et de Statistique,* 34: 13-69.