Nonparametric Estimation for Regulation Models *

Andreea ENACHE† and Jean-Pierre FLORENS‡

September 2017

Forthcoming in Annals of Economics and Statistics

Abstract

This paper presents a nonparametric structural analysis of a class of contract models à la Baron et al. (1982). Our analysis is based on a well-posed inverse problem linking the quantile function of the observations and the functional parameter of interest. The resolution of this problem gives the identification properties of the model and leads to an estimation procedure. We provide implementation and asymptotic properties of this type of L-estimator. We extend our analysis by introducing an instrumental variable estimator of the cost function.

Keywords: L-functionals, regulation models, principal-agent model, adverse selection, nonparametric statistics, structural econometrics

JEL Classification: C40, D86, L51

*We are indebted to Stéphane Bonhomme, Christian Bontemps, Xavier D’Haultfoeuille, David Martimort and Quang Vuong for many helpful discussions and constructive comments. We also benefitted from participants’ comments at the CRESSE Conference 2014, the International Industrial Organization Conference 2015, the London School of Economics, the Paris School of Economics European Doctoral Program Jamboree 2014, and the seminars at University of Chicago, European University Institute, Georgetown University, Université de Montréal and Northwestern University.

†Carlo Giannini Fellow at IGER-Bocconi University, email: andreea.enache@unibocconi.it (corresponding author).

‡Toulouse School of Economics, email: jean-pierre.florens@tse-fr.eu.
1 Introduction

Contracts are part of our daily life, present in any economic activity we might think about: financial markets (the relationship between a lender and a borrower), labor market (employer-employee), selling mechanisms (owner-buyer), public utilities (regulatory agency-supplier), insurance market (brokerage company-customers), corporate organisation (shareholders-managers) and the list is far from being exhaustive. All these instances have in common one main feature: they can all be represented as a principal-agent relation with incomplete information.

In practice, we very often encounter the situation where a local municipality delegates the task of providing a so-called "public service" (water supply, water management, waste management, public transportation, energy) to another entity, which is usually a natural monopoly.\footnote{Usually in these sectors of activity, the production features make optimal the provision of the good by one firm only.} We face here a principal-agent relationship\footnote{Throughout this paper, we will refer to the principal as "she" and to the agent as "he".} characterized by the presence of informational asymmetries under the form of adverse selection or moral hazard or very often a mix of the two (for an extensive description of these issues see Laffont and Martimort (2001)). The adverse selection arises from the fact that the local municipality does not have perfect information about the demand or the cost conditions faced by the firm. Hence, in order to increase its revenues, the natural monopolist may try to distort reality when reporting to the principal. A low cost firm may try to pass itself off as a high cost supplier in order to be able to fix a higher price for its product or service. Another informational disadvantage faced by the principal comes from the fact she cannot verify if the managerial team made enough effort to reduce the costs of operating in the industry. This creates the so-called moral hazard problem or "hidden action". For example, the firm may find it in its own interest not to invest in a new technology, if the revenues received from the principal depend on the observed costs.

We focus our attention on the public-utilities problem, but our econometric approach can be extended to other examples. In this paper we take into account only cases of pure adverse selection (i.e. we assume that the principal doesn’t face an effort shirking problem). Moreover she is not able to observe the ex-post realisations of the cost
(no possibility of auditing). This could be the case of development countries where the regulatory agency has a limited regulatory power that leads to a weak, and hence useless, auditing system (for an in-depth discussion about the consequences of institutional failures on the design of regulatory policies in development countries, see Estache and Wren-Lewis (2009)). In order to account for the informational asymmetries concerning cost opportunities, the local municipality must design a regulation policy that will limit the rents the monopolistic firm can extract as a result of its informational advantage. The main instruments at the regulator’s disposal for inducing truth telling is the price for the regulated product and the monetary transfers (taxes or subsidies). Thus, the principal will offer a menu of contracts that specifies the prices for the good or service at different cost realisations and a payment scheme. These options are designed in such a manner that the firm will have incentives to report its actual cost. For a survey of the early literature on regulation see Baron (1989), Armstrong and Sappington (2007), and for more recent developments in the field see Joskow (2007).

Although the economic theory behind different forms of informational asymmetries is very well developed, the econometric literature that treats these models is still scarce. The early empirical applications in the field of economic regulation were conducted using a reduced-form approach. Moreover, these papers were mainly concerned with the implications of regulatory policies on the level and structure of prices or on innovation and production growth, but they did not incorporate the effects of informational asymmetries on the agent’s behaviour. The first paper that used a structural approach to account for the presence of private information in a regulator-utility relation is Wolak et al. (1994). The author used a parametric setting to estimate the parameters of the agent’s production and cost function for the water supplier in California. Other empirical applications are in particular conducted by Ivaldi and Martimort (1994) or Chiappori and Salanié (2002). D’Haultfoeuille and Février (2007) and Bontemps and Martimort (2014) provide point identification results in an adverse selection problem. Perrigne and Vuong (2011) analyse nonparametric point identification issues in a static mixed model of asymmetric information, while D’Haultfoeuille and Février (2015) provide a partial identification re-

3The authors show that one cannot jointly point identify the distribution of types, the marginal surplus function and the marginal cost function. Nevertheless, if one of these three functions is known, the other two are point identified.
result for the distribution of agent’s types and the cost function. Although the identification issue in the static problem of contracting has been studied previously in Perrigne and Vuong (2011), our contribution comes from the fact we also present the estimation issues and compute the speed of convergence for our estimator. Another contribution comes from the use of the quantile approach to study the identification and estimation of the model. The quantile approach is also used by Bontemps and Martimort (2014) in the derivation of their results.

Our paper is essentially devoted to the analysis of the adverse selection problem (the "hidden information") in a static interaction between the regulated supplier and his regulator. For this purpose, we will use the economic settings exposed in Baron et al. (1982) and Laffont and Tirole (1993) as follows.

## 2 Economic model

Let us first present the features of the static model of hidden information à la Baron et al. (1982).

We consider the situation where the private information of the agent consists of the knowledge of his cost inefficiency parameter denoted by \( \theta \) and in order to simplify notations suppose that \( \theta \in [0, 1] \). The principal cannot observe the \( \theta \) but knows that its distribution is given by \( F : [0, 1] \rightarrow [0, 1] \) and \( f(\theta) \equiv F'(\theta) \). The principal also knows the demand function denoted by \( X(p) \), where \( p \) is the transaction price. The supplier’s production cost is a function of his cost inefficiency parameter and the quantity provided and is denoted by \( C(x, \theta) \).

The following assumptions are standard and come from the economic literature (see Baron et al. (1982)):

**Assumption 1.** \( h(\theta) = \frac{F(\theta)}{f(\theta)} \) is a well-defined function on \([0, 1]\) \((f \text{ is continuously differentiable with strictly positive support on } [0, 1]) \) and is nondecreasing as a function of \( \theta \).

**Assumption 2.** \( C_\theta > 0, C_{x, \theta} > 0, C_{xx} \geq 0, C_{xx\theta} \geq 0, C_{x\theta \theta} \geq 0, \forall x, \theta \).

---

4The quantile approach has been previously used by Marmer and Shneyerov (2012) in the setting of the first-price auction model.
Using the Revelation Principle, the regulator offers the firm the possibility to self-select from a menu of contracts specifying pairs of price and transfers \( \{p(\theta), T(\theta)\} \). One can think about the transfer \( T(\theta) \) as being a subsidy (in this case, the pricing decision belongs to the regulator, but she has to pay a subsidy to the firm when the price doesn’t fully cover the cost) or as tax (the agent decides the price, but he has to pay a tax corresponding to different price levels). The benevolent regulator solves a maximisation problem of the weighted sum of consumer surplus and the profit obtained by the firm. If we denote by \( S \) the surplus and by \( \Pi \) the profit of the firm, the principal will maximise \( S + \delta \Pi \), where \( \delta \) is the importance that the principal gives to the rent left to the agent. The usual assumption is that \( \delta \in [0, 1) \) as the principal is more concerned with the taxpayers’ welfare. This hypothesis is quite important as without it the solution to the principal’s problem will simply be to allow the firm to be the decision maker and to maximise its profit.

We write the decision maker’s problem as follows:

\[
\max_{p,T} \int_0^1 \left\{ S(p(\theta)) - T(\theta) + \delta [\Pi(p(\theta), T(\theta); \theta)] \right\} f(\theta) \, d\theta
\]

subject to the participation constraint and, respectively, the incentive compatibility constraint:

\[
\Pi(p(\theta), T(\theta); \theta) \geq 0, \quad \Pi(p(\theta), T(\theta); \theta) \geq \Pi(p(\hat{\theta}), T(\hat{\theta}); \theta), \quad \forall \theta, \hat{\theta} \in [0, 1]
\]

After some computations (which can be found in Baron et al. (1982)) and denoting by \( V(x) = \int_0^x P(u) \, du \) and \( P(x) \) the inverse demand function, one can write the previous problem in an equivalent way:

\[
\int_0^1 \left\{ V(x(\theta)) - C(x(\theta), \theta) - (1 - \delta) \frac{F'(\theta)}{f(\theta)} C_\theta(x(\theta), \theta) \right\} f(\theta) \, d\theta
\]

\(^5\)The Revelation Principle is a powerful tool that tells us the principal can focus without loss of generality on a direct revelation mechanism that is individually rational and incentive compatible.

\(^6\)We set the agent’s reservation profit to 0.

\(^7\)\( S(p) = V(x) - P(x)x \).
s.t.:
\[ T(\theta) = P(x(\theta))x(\theta) - C(x(\theta), \theta) - \int_0^1 C_\theta(x(u), u) \, du, \]
\[ p(\theta) = P(x(\theta)) \]
and
\[ x(\theta) = X(p(\theta)) \] is weakly decreasing.\(^8\)

The optimal contract under the presence of asymmetric information is characterized by the following First Order Condition:

\[ P(x(\theta)) = C_x(x(\theta), \theta) + (1 - \delta) \frac{F(\theta)}{f(\theta)} C_{x \theta}(x(\theta), \theta). \] (1)

One can see from equation (1) that a higher inefficiency is associated with a higher price (the benchmark being price equals marginal cost).

3 Econometric model: Identification

The econometric model is mainly based on an analysis of the First Order Condition presented in (1). We consider several observation schemes and look at their implications for the identification issue. We consider two cases. In the first, only prices are observed, while in the second case, we observe prices, quantities and, for reasons explained below, some instrumental variables. In both situations, restrictions are necessary to obtain identification.

3.1 Identification without instruments

In this case, we assume that \( C(x, \theta) \) is decomposed into:

\[ C(x, \theta) = c_1(\theta)x. \]

\(^8\)The monotonicity of the hazard ratio imposed in Assumption 1 is crucial for the monotonicity of the quantity with respect to the cost parameter (and thus to avoid pooling).
Equivalently, we may assume \( C(x, \theta) = c_1(\theta)c_2(x) \), where \( c_2 \) is given. Then, without any restrictions, we redefine \( x \) such that \( c_2(x) = x \). Then \( c_1(\theta) \) is the constant marginal cost and increasing in \( \theta \). \( \theta \) will be interpreted as the rank of agents in terms of their cost inefficiencies. The weight on the firm’s profit, \( \delta \), is different from zero. At the same time we suppose that \( \delta \) is constant and known (otherwise, \( \delta \) would not be identified). Thus we observe an i.i.d. sample of the random element \( p \) defined by:

\[
p = c_1(\theta) + (1 - \delta) \frac{F(\theta)}{f(\theta)} c'_1(\theta).
\]

Our goal is to estimate the structural parameters \( c_1(\theta) \) and \( F(\theta) \) from the observations \((p_i)_{i=1,...,n}\). Remember that the principal knows the distribution from which the types are drawn, but this distribution is unobservable to the econometrician.

We concentrate our analysis on equation (2) which is a particular case of a general class of game-theoretic models characterized by a relation between an observable variable \( p \) and an inobservable variable \( \theta \) of the form \( p = \sigma_F(\theta) \). The \( \sigma_F \) is the structural element of the model and is assumed to be a given equilibrium (for example, Bayesian Nash Equilibrium). The strategic component of the game is formalized by the fact that \( \sigma \) depends on \( \theta \) and on its distribution, \( F \). As remarked by Florens et al. (1998) and Florens and Sbaï (2010), the dependence of \( \sigma \) on \( F \) modifies the structural properties of the model (identification and estimation). This paper illustrates the importance of this dependence relation.

The distribution of primitives and the distribution of observables in our simple model of adverse selection are related by the following functional relation:

\[
G(p) = \Pr(P \leq p) = \Pr(\sigma_F(\theta) \leq p) = F(\sigma_F^{-1}(p)),
\]

where \( G \) denotes the cdf of \( p \) and \( \sigma_F(\cdot)\) is an increasing function of its argument. One can note that \( G(\cdot) \) is identified by the data. From (3) we obtain the implicit relationship between \( F \) and \( G \):

\[
F \circ \sigma_F^{-1} = G \text{ or } F = G \circ \sigma_F.
\]
The previous equation can be written in a quantile version:

\[ G^{-1} = \sigma F \circ F^{-1}. \]  

(4)

This paper is based on the analysis of this quantile equation. \( G^{-1} \) is estimated by the quantile function of the observables and \( F^{-1} \) is the solution of an inverse problem.

The identification of the model can easily be assessed by solving the quantile equation \( G^{-1} = \sigma F \circ F^{-1} \) (as we are in the identification stage, the population distribution of the observable is supposed to be known and therefore there is no noise). In our case, the \( \sigma \) function has the following shape: \( \sigma_F(\theta) = c_1(\theta) + (1 - \delta) \frac{F(\theta)}{f(\theta)} c_1'(\theta) \). If we do the composition in (4) we get that:

\[
G^{-1}(\alpha) = c_1(F^{-1}(\alpha)) + (1 - \delta) \frac{\alpha}{f(F^{-1}(\alpha))} c_1'(F^{-1}(\alpha)).
\]

This can be rewritten moreover as:

\[
G^{-1}(\alpha) = c_1(F^{-1}(\alpha)) + (1 - \delta) \alpha c_1'(F^{-1}(\alpha)) F^{-1}'(\alpha).
\]

If we denote \( c_1(F^{-1}(\alpha)) = \varphi(\alpha) \) and rearrange the previous equation we obtain that:

\[
G^{-1}(\alpha) = \varphi(\alpha) + (1 - \delta) \alpha \times \varphi'(\alpha).
\]

(5)

This relation shows that we cannot identify separately \( c_1 \) and \( F^{-1} \) (or \( F \)) and that only \( c_1 \circ F^{-1} \) may be derived from \( G^{-1} \) that contains all the information coined by the data. Two possible restrictions imply identification: \( c_1 \) may be assumed to be given or \( \theta \) may be normalized to be uniformly distributed on \([0, 1]\). In this latter case, \( F^{-1} \) is the identity and \( \varphi \) is equal to the cost function. We will focus now on the global identification of \( \varphi \). Equation (5) is a first order differential equation in \( \varphi(\alpha) \) with non-constant coefficients. Our economic model is identified if the equation has a unique solution in \( \varphi \). The solution to this equation is:
\[ \varphi(\alpha) = \alpha - \frac{1}{1-\delta} \left[ \int_0^\alpha \frac{G^{-1}(t)}{1-\delta} \times t^{1-\delta} \, dt + \text{constant} \right]. \]  

(6)

If \( \theta \in [0, 1] \), \( F^{-1}(0) = 0 \) and assuming \( c_1(0) = 0 \), we have that \( \varphi(0) = 0 \). We rewrite equation (6) as:

\[ \varphi(\alpha) \times \alpha^{1-\delta} = \int_0^\alpha \frac{G^{-1}(t)}{1-\delta} \times t^{1-\delta} \, dt + \text{constant}, \]

and the constant is identified and equal to zero. Moreover if we apply the l'Hôpital rule we obtain that: \( \lim_{\alpha \to 0^+} G^{-1}(\alpha) = 0 \).

Therefore the solution to our problem is given by the next proposition:

**Proposition 1.** In a pure classical model of adverse selection, \( \varphi(\alpha) \) is described by the following relation:

\[ \varphi(\alpha) = \alpha - \frac{1}{1-\delta} \left[ \int_0^\alpha \frac{G^{-1}(t)}{1-\delta} \times t^{1-\delta} \, dt \right], \]

where \( \alpha \in (0, 1] \) and \( G^{-1}(\alpha) \) is the quantile function of prices.

**Remark 1.** By construction, \( \varphi \) is increasing for any cdf \( G \). This means that we cannot use the restriction that \( \varphi \) is increasing to improve its estimation or to test the model.

### 3.2 Identification with instruments

Let us now consider a new complete observation scheme where \( p \) and \( x \) are observed, and as well as some instrumental variables. In this case, the cost function is decomposed into

\[ C(x, \theta) = c_1(\theta)c_2(x), \]

but we do not assume \( c_2(x) \) to be known. One can always modify the marginal cost \( C'_x(x, \theta) = c_1(\theta)c'_2(x) \) by multiplying \( c_1(\theta) \) by a constant et dividing
\( c'_2(x) \) by the same constant and thus the model remains unchanged. We may, for example, normalize the cost function by \( E(\ln c'_2(x)) = E(\ln p) \). Equation (1) can be written as:

\[
\ln p = \ln c'_2(x) + \ln \lambda,
\]

(7)

where \( \lambda = c_1(\theta) + (1 - \delta) \frac{F}{f}(\theta) c_1'(\theta) \). If we denote by \( m \) the mean of \( \ln \lambda \) we can rewrite (7) as:

\[
\ln p = \psi(x) + u,
\]

(8)

where \( \psi = \ln c'_2(x) + m \) and \( u = \ln \lambda - m \) is a zero mean residual. Equation (8) is not a usual nonparametric regression because \( u \) and \( x \) are related (\( E(u|x) \neq 0 \)). Indeed, both \( p \) and \( x \) depend on the type of the agent, \( \theta \). To solve this problem we use an instrumental variables approach. Therefore, one has to choose an instrument, \( w \),\(^9\) that is independent to the type of the agent (\( \theta \perp w \)) but which explains the quantity very well. \( w \) could be a measure of the population or other descriptive variable of the place where the public service is provided. One can note that we use in our model only the restrictions on the prices and not on the quantities as well. Given the instruments, the observed quantities will vary with \( w \).

Let \( w \) be a vector of instruments that satisfies the following two assumptions:

**Assumption 3.** The instrument, \( w \), and the type of the agent, \( \theta \), are independent.

**Assumption 4.** The \( L^2 \) completeness condition applies: \( E(\psi(x)|w) = 0 \Rightarrow \psi = 0, \forall \psi \) square integrable. This property has been studied intensively in the recent literature.

**Proposition 2.** Under Assumption 3 and Assumption 4, \( \psi(x) \) is identified using standard arguments of nonparametric instrumental variables methods. Therefore, \( \psi \) is the solution of the following equation:

\[
E(\ln p|w) = E(\psi(x)|w).
\]

(9)

The solution of (9) is unique under the completeness condition (see Darolles et al. (2011)). Then \( c'_2(x) \) and \( (\ln c_1(\theta))' \) are identified as well.

---

\(^9\) We do not study here the choice and the validity of the instruments. Basically, in order to justify the validity of the instruments, one should write the complete structural model and then decide where and which instruments to put in the model. For example, one could imagine an instrument that affects the demand function faced by the firm, but doesn’t affect the cost function.
Remark 2. If we assume \( w \) and \( \theta \) independent then we have \( u \) and \( w \) independent. The model becomes an IV model with independence that leads to a nonlinear integral equation which may be solved with respect to \( \psi \). The computations are more tedious, but the identification conditions are weaker (see Dunker et al. (2014), Fève and Florens (2014) for more details, as this approach is not developed in the present paper).

Remark 3. In our model we use only one observable, the price, and we have only one source of unobserved heterogeneity, the type of the agent, i.e. \( \theta \). The quantity is random, endogenous and not explained in this model. If one would like to use more than one observable (for example, the transfers), then additional sources of unobserved heterogeneity are necessary, otherwise the model will clearly be rejected (not enough variation in the data). The model with more than one observable (where \( (p,q) \) are on a curb and thus the distribution of \( q \) in function of \( p \) is degenerated) has been treated in extenso in Bontemps and Martimort (2014)).

4 Estimation

4.1 Estimation without instruments

We consider in this section the estimation issues for the case where only the prices are observed. In the previous section we showed the identification of our model and the estimator can be computed in a straightforward way:

\[
\hat{\phi}(\alpha) = \frac{1}{1-\delta} \int_0^1 J(t) \hat{G}^{-1}(t) dt, \tag{10}
\]

with \( J(t) = 1 (t \leq \alpha) t^{1-\delta} \).

The estimation of the quantile function of types is conducted by using the order statistics of the data, in our case the prices. Thus if \( P_{(i:n)} \) denotes the \( i^{th} \) order statistic of the sample of size \( n \), we know that \( \hat{G}^{-1}(t) = \sum_{j=1}^n P_{(j:n)} 1 \left( t \in \left( \frac{j-1}{n}, \frac{j}{n} \right) \right) \). If we plug this
expression into (10), we get that:

\[ \hat{\varphi}(\alpha) = \alpha \frac{1}{1 - \delta} \sum_{i=1}^{n} \int_{\frac{i}{n}}^{\frac{i+1}{n}} J(t) G^{-1}(t) dt \]

\[ = \alpha \frac{1}{1 - \delta} \times \sum_{i=1}^{n} 1 \left( \alpha \in \left( \frac{i-1}{n}, \frac{i}{n} \right) \right) \times \left\{ 1 (i \geq 2) \sum_{j=1}^{i-1} P_{jn} \left[ \left( \frac{j}{n} \right) \frac{1}{1 - \delta} - \left( \frac{j-1}{n} \right) \frac{1}{1 - \delta} \right] + P_{i,n} \left[ \alpha \frac{1}{1 - \delta} - \left( \frac{i-1}{n} \right) \frac{1}{1 - \delta} \right] \right\} . \]

4.1.1 Asymptotics properties

Our main tool is a general theorem on the asymptotic behaviour of linear functions of order statistics (see Moore (1968), Stigler et al. (1969), Stigler et al. (1974)).\(^{10}\) If \( P_{(i:n)} \) denotes the \( i \)th order statistic of an iid sample of prices of size \( n \) and drawn from the cdf \( G \), with \( P_{(1:n)} \leq \ldots \leq P_{(n:n)} \), then an L-estimator is a linear function of order statistics:

\[ T_n = \frac{1}{n} \sum_{i=1}^{n} J \left( \frac{i}{n} \right) P_{(i:n)} , \]

where \( J \) is a function defined on \([0, 1]\).

In what follows all the asymptotic results will be derived on \([\alpha_0, \alpha_1]\) in order to avoid boundary problems, where \( \alpha_0 > 0 \) and \( \alpha_1 < 1 \). Using Moore (1968), Stigler et al. (1969), Stigler et al. (1974) (see also for regularity conditions on the \( J \) function) we have the following result:

**Proposition 3.** For any \( \alpha \in [\alpha_0, \alpha_1] \),

\[ \sqrt{n} \left[ \hat{\varphi}(\alpha) - \varphi(\alpha) \right] \sim N(0, \sigma^2(\alpha)) , \]

\(^{10}\)The so-called "L-estimators".
where
\[
\sigma^2(\alpha) = 2\alpha - \frac{2}{1-\delta} \int \int J(G(p))J(G(r))(G(p \land r) - G(p)G(r)) \, dp \, dr.
\]

**Remark 4.** Let us assume for simplicity’s sake that \(c(\theta) = \theta\) and then \(\varphi(\alpha) = F^{-1}(\alpha)\).

From the estimation of \(F^{-1}\) we can estimate \(F\) by taking its usual or generalized inverse.

Proposition 3 gives pointwise asymptotic normality of \(\hat{\varphi}(\alpha)\), but a more sophisticated result would show that the function \(\sqrt{n} (\hat{\varphi} - \varphi)\), converges as a stochastic process in \(L^2\) to a gaussian process on \([\alpha_0, \alpha_1]\).

Next if we apply the delta method, we can easily see that the empirical process of the cdf of the types converges to a Gaussian process on strict compact subsets of the support of \(\theta\). Before checking the convergence of the pdf, we need to compute the derivative of the quantile function:

\[
F^{-1'}(\alpha) = -\frac{1}{1-\delta} \frac{\delta - 2}{\alpha^{1-\delta}} \left[ \int_0^\alpha \frac{G^{-1}(t)}{1-\delta} \times t^{1-\delta} \, dt \right] + \frac{1}{1-\delta} \frac{G^{-1}(\alpha)}{\alpha^{1-\delta}} \frac{\delta}{1-\delta}
\]

One can easily show that \(\sqrt{n} (\hat{F}^{-1'} - F^{-1'})\) also converges to a Gaussian process (as \(F^{-1'}\) is a function of \(G^{-1}\) and of an integral of \(G^{-1}\)). Moreover, the probability distribution function can be written as:

\[
f(\theta) = \frac{1}{F^{-1'}(F(\theta))} \Rightarrow \hat{f}(\theta) = \frac{1}{F^{-1'}(\hat{F}(\theta))}.
\]

Using again the delta method, we get that \(\sqrt{n}(\hat{f} - f) \rightsquigarrow\) to a Gaussian process. Note that the estimation of \(f\) at a \(\sqrt{n}\) speed is not in contradiction with the minimax speed of convergence of a density. Actually this result applies if \(f\) is the density of the observed data, which is not the case in our example. The speed of convergence of \(\hat{f}\) is due to the strategic feature of the game model which may improve or decrease the rate of convergence.
4.1.2 Simulations

We conduct Monte Carlo simulations (100 simulations) to check that our estimator performs well in small samples (n=20 and the types are drawn from a Beta distribution, $δ = 0.3$ and $c(θ) = θ^2$). The results are presented in Figure 1.

4.2 Estimation using instruments

We observe an iid sample of $(p, x, w)$ denoted by $(p_i, x_i, w_i)$, $i = 1, ..., n$ and the model is summarized by:

\[
\ln p = \psi(x) + u, \quad E(u|w) = 0 \\
G^{-1}(\alpha) = \varphi(\alpha) + \alpha(1 - \delta)\varphi'(\alpha),
\]

(11)

where $G^{-1}(\alpha)$ is the quantile function of $e^u$ and the parameters of interest are the two functions $\psi$ and $\varphi$. We have assumed that $u$ verifies $E(u) = 0$. Then $\lambda$ is defined up to a multiplication constant and $G^{-1}$ and $\varphi$ are also identified up to a multiplication constant. Equivalently, $G^{-1}(\alpha) = e^{H^{-1}(\alpha)}$, where $H^{-1}$ is the quantile function of $u$. The estimation is performed in two steps:
1. We estimate the $\psi$ function using a usual nonparametric instrumental variables method. Several computer programs are available (see e.g. Centorrino et al. (2015)). For example, we may use a Tikhonov/Kernel method. The function $\psi$ verifies:

$$
\hat{\psi}_\rho(x) = \frac{1}{\rho} \sum_{j=1}^n \left\{ \sum_{i=1}^n (\ln p_i - \psi_\rho(x_i)) K\left(\frac{w_j - w_i}{h}\right) \right\} \frac{K\left(\frac{x - x_j}{h}\right)}{\sum_{j=1}^n K\left(\frac{x - x_j}{h}\right)},
$$

where $\Psi$ denotes the vector of $\psi_\rho(x_i), i = 1,...n$ and verifies:

$$
\Psi = (\rho I_n + K_xK_w)^{-1}K_xKWy.
$$

2. From $\hat{\psi}_\rho$ we derive the vector of residuals $\hat{u}_i$, next we estimate the empirical quantile of $\hat{u}_i$ denoted by $\hat{H}^{-1}(\alpha)$, and finally $\hat{G}^{-1}(\alpha) = e^{\hat{H}^{-1}(\alpha)}$. Then we solve the second equation of the model (11) as in section 3.1 and we obtain an estimation of $\hat{\phi}$.

We illustrate this method by a simulation example, but we don’t develop the asymptotic properties of this estimation. The properties of $\hat{\psi}_\rho(x)$ are now well-established (see Carrasco et al. (2007)) and the properties of the empirical cdf of $\hat{u}_i$ are considered in Babii and Florens (2016). Extension of this method would consider the full independence between $u$ and $x$ and it would consider the joint estimation of $\phi$ and $\psi$. 

15
4.2.1 Simulations

The estimation procedure consists of two steps. In a first stage, we estimate \( c'_2(x) \). The next step is to compute \( \frac{p}{c'_2(x)} \), the quantile function of it, denoted \( D^{-1} \) and finally to estimate \( F^{-1} \) from the integral equation:

\[
F^{-1}(\alpha) = \alpha - \frac{1}{1-\delta} \left[ \int_0^{\alpha} \frac{D^{-1}(t)}{1-\delta} \times \frac{\delta}{1-\delta} dt \right].
\]

Let us next solve the regression in (9) using nonparametric instrumental variables (see Darolles et al. (2011)).\(^{11}\) For the purpose of the simulations, we are going to use the full independence between \( \ln \lambda \) and \( w \), i.e., \( \mathbb{E}(\ln \lambda \mid w) = 0 \). We conduct simulations using the following configuration:

- The type of the agent (its productivity) is drawn from a Beta distribution on the interval \([0, 1]\), i.e: \( F(\theta) = \theta^2 \).
- The instrument \( w \) is drawn from a normal distribution \( w \sim N(0, \sigma^2) \).
- The quantity, \( q \mid \theta, w \) has the following expression: \( x = \exp(a\theta + bw + c\varepsilon) \), where \( \varepsilon \sim N(0, \tau^2) \).
- We consider the following expression for the cost function: \( C(x, \theta) = \theta c_2(x) \) (here \( c_1(\theta) = \theta \)). Next we choose the marginal cost function: \( c'_2(x) = \omega x \) and we compute the prices as \( p = c'_2(x) \left( \theta + (1-\delta) \frac{F(\theta)}{f(\theta)} \right) \). This way, we obtain our data: \( w, x, p \).
- If the parameter \( a = 0 \) we are in the exogenous case and if \( a \neq 0 \) we are in the endogenous case. Therefore, \( a \) is the parameter that controls for the endogeneity of our problem.
- We set the following values for our parameters in the simulations: \( \sigma^2 = 1, \tau^2 = 0.05, a = 1, b = 0.8, c = 0.3, \omega = 3 \) and \( \delta = 0.3 \).
- Our structural parameters are \( c'_2(x) \) and \( F \).

\(^{11}\)For practical implementation, see Centorrino et al. (2015). We use the R "npregiv" command (package "np") and its default regularization method, Landweber-Fridman.
In Figure 2 we present the estimation of the logarithm of the cost function using 50 observations generated as described above. Figure 3 describes the quantile function of the types constructed using the estimation of the cost function. Remember that the $\theta$s are identified only up to a multiplication constant. In the simulations, this constant has been identified using the true mean of $\ln \lambda$ in order to facilitate the comparison with the true quantile function of $\theta$. 

Figure 2: Monte Carlo simulations for the estimation of the log of the cost function, $S=20, n=50$
Figure 3: The estimation of the quantile function of the types, S=20, n=50

5 Conclusions

The presence of asymmetric information between the regulator and the agency leads to the implementation of a less efficient outcome (the so-called "second-best") than the benchmark outcome (the "first-best") achieved in the case of symmetric information. Providing estimates for the primitives of the supplier’s production process is crucial as it guides the regulator in the rate-setting process. Besides this, one can test the optimality of the contracts being implemented and also conduct counterfactuals to see under which payment schemes the consumer surplus would be increased. Once one has estimated the distribution of the supplier’s costs and the cost function, another policy simulation can be envisioned: to check the effects on the level of prices and quantities of letting another supplier enter the market for the public service.

In this paper we provide a framework for the identification and estimation of the cost function composed with the quantile function of the cost inefficiency parameter in a static model of pure adverse selection. Next, we provide a quantile estimation and using the order statistic theory we show that the empirical process of our estimator converges to a
gaussian process at a root-n speed of convergence. Finally, we present an extension of
the model and we relax the linearity of the cost function. We estimate a semi-parametric
specification of the agent’s cost function using the nonparametric instrumental variables
method.
References


