Nonparametric Identification of Incentive Regulation Models*

Isabelle Perrigne
Pennsylvania State University

Quang Vuong
Pennsylvania State University

February 2008

* In memory of our friend Jean-Jacques Laffont. He initiated our research of the past fifteen years. We dedicate this paper to Jean-Jacques, Colette and their children. Preliminary versions were presented at Johns Hopkins University, Penn State University, SITE Meeting, European University Institute in Firenze, Université de Toulouse, Columbia University, New York University, Princeton University and Northwestern University. We thank P. Courty, J. Harrington, A. Pakes, R. Porter, M. Whinston and the participants at seminars and conferences for helpful discussions. We are grateful to the Co-Editor and two referees for detailed comments that improve significantly our paper. Financial support from the National Science Foundation under Grant SES–0452154 is gratefully acknowledged.
Correspondence to: I. Perrigne, Department of Economics, Pennsylvania State University, University Park, PA 16802. E-mail address : perrigne@psu.edu.

Abstract

This paper studies the nonparametric identification of contract models under asymmetric information. Specifically, we consider the incentive regulation model with ex post observed cost developed in Laffont and Tirole (1986), which includes adverse selection, moral hazard and a payment/transfer. We first extend this model to allow for general random demand and cost functions. We then map the resulting model into a structural econometric model with observed and unobserved heterogeneity. We establish the nonparametric identification of the cost of public funds, the demand, cost and effort disutility, as well as the joint distribution of the random elements of the structural model, which are the firm's type, the demand and cost shocks, and the residual transfer.

Fields: Contracts, Asymmetric Information, Regulation, Nonparametric Identification, Cost Efficiency.

Address: Pennsylvania State University, Department of Economics, University Park, PA 16802.

Nonparametric Identification of Incentive Regulation Models

I. Perrigne & Q. Vuong

1 Introduction

Over the past thirty years, economists have emphasized the fundamental role played by asymmetric information in economic relationships. The imperfect knowledge of key economic variables induces strategic behavior among economic agents. Contracts provide an important example of how information asymmetry governs relationships between a principal and an agent. Two types of imperfect information can affect contractual relationships, namely some agent's hidden characteristics or type and some agent's hidden action or effort leading to the socalled adverse selection and moral hazard problems, respectively. See Laffont and Martimort (2001) and Bolton and Dewatripont (2005) for monographs on contract theory. Contracts are widely used in the economic world. To name a few, agriculture, corporate law, finance, insurance, management and retailing provide many examples of contractual relationships.

In this paper, we are interested in the incentive contract model developed by Laffont and Tirole (1986) between a regulatory authority or regulator and a monopolistic firm. Several reasons motivate our interest. First, the Laffont and Tirole's (1986) model contains the main ingredients of a contract model, namely adverse selection, moral hazard and a payment/transfer function. Second, their model is a control problem under incomplete information whose results have a broader impact than regulation itself. See Laffont (1994). Third, from an empirical point of view, data on regulatory contracts are readily available from public regulatory commissions while contract terms are in general well defined in terms of objectives assigned to the firm and compensation arrangements. This contrasts with private contracts that can be under an implicit form or subject to unspecified terms as in incomplete contracts.

Despite these fundamental theoretical developments and the economic importance of contracts, very few empirical studies analyzing contract data rely on a structural modeling. As a matter of fact, the empirical literature is mostly limited to a reduced form approach as surveyed by Chiappori and Salanié (2003) for the analysis of contracts in general and Joskow and Rose (1989) for regulation. Some notable exceptions are Paarsch and Shearer (2000) for labor contracts, Ivaldi and Martimort (1994) and Miravete (2002) for nonlinear pricing and Wolak (1994), Thomas (1995), Gagnepain and Ivaldi (2002) and Brocas, Chan and Perrigne (2006) for regulation. This did not meet the expectation of theorists. For instance, Laffont (1994, p. 532) writes that the paper by Wolak (1994) "is the first in a long series of applied works which will renew the econometrics of regulation with the help of the new theory of regulation." Similarly, Laffont and Tirole (1993, p. 669) conclude that "econometric analyses are badly needed in the area," while they "do wish that such a core of empirical analysis will develop in the years to come." Such high expectations have not been met because asymmetric information models lead to complex econometric models whose estimation requires suitable econometric tools. Moreover, the issue of identification needs to be addressed. Though the analyst may end up specifying a parametric model as in the above papers, studying nonparametric identification is valuable for thinking carefully about which information in the data allows one to identify each unknown function.[1] Another important related question is to derive the restrictions imposed by the model on observables to test the model validity. Without such restrictions, the model could rationalize any data.

In this paper, we adopt a nonparametric approach in the spirit of Guerre, Perrigne and Vuong (2000) to address the identification of the incentive regulation model developed by Laffont and Tirole (1986). Our results are general in the sense that other contract models can be identified using a similar approach. We first adapt their model. In particular, they

[1]The econometrics of auction models provides interesting examples. For instance, the common value auction model can be estimated parametrically, while this model is not identified in general. See Paarsch (1992) and Février, Preget and Visser (2004) for parametric estimation of common value models. See Laffont and Vuong (1996) and Li, Perrigne and Vuong (2000) for the nonidentification of these models.

consider the regulation of a public good with a fixed demand, while a private good with a random demand seems more prevalent. The contract design then considers expected demand, while the firm has to fullfill the realized demand. Additional difficulties lie in the contract implementation and in checking whether the local second-order conditions are globally satisfied. We then derive the corresponding structural econometric model. The structural elements are the demand subject to some random shock, the cost depending on the unobserved firm's type and effort subject to some random shock, the cost of public funds, the effort disutility and the firms' type distribution. The observables are the (ex post) demand, the (ex post) cost, the price and the payment to the firm. A first difficulty arises from the fact that the firm's effort and type, which can be viewed as firm's unobserved heterogeneity, are both unobserved. A second difficulty relates to the singularity of the model. Three unobserved random variables (demand and cost random shocks and firm's type) determine four endogenous variables (demand, cost, price and payment). Thus, the econometric model is singular. We introduce an additional term to the payment function due to the regulator's unobserved heterogeneity. This error term, called the residual transfer, is unrelated to firm's cost performance and can capture possible deviations from the optimal payment due to constraints faced by regulators arising from politics, legal systems and external environments. The econometric model is then based on five equations, which give the demand, the cost, the *generalized* Ramsey pricing, the optimal effort level and the transfer. It also allows for exogenous variables capturing observed firm, regulator, and/or market heterogeneity.

We first show that the model is nonparametrically identified given the cost of public funds. The firm's type is assumed to be conditionally independent of the demand and cost random shocks, while a natural location-scale normalization is imposed. Nonparametric identification then relies on the bijective mapping between the price and the firm's type. The analogy with auction models becomes clear. In particular, the bidder's private value and his bid are equivalent to the firm's type and the price in a contract. Guerre, Perrigne and Vuong (2000) recover the bidder's private value from the equilibrium first-order condition, which relates monotonically the private value to the bid and the bid distribution. A similar idea is exploited here. Next, we address the nonparametric identification of the cost of public funds which relies on the mean independence of the residual transfer from the firm's cost given the firm's type. Our nonparametric identification result represents an important step towards the development of the econometrics of contract models. In addition, our results offer a new approach to the estimation of cost efficiency that incorporates explicitly the effects of information asymmetry. See Park, Sickles and Simar (1998) for a survey of the classical literature on production frontier models with recent developments in a semiparametric setting.

The paper is organized as follows. Section 2 presents a generalization of the Laffont and Tirole (1986) model with a stochastic demand for a private good as well as its implementation through linear contracts and the verification of the second-order conditions. Section 3 addresses its nonparametric identification. It includes the derivation of the econometric model, while considering the identification of the structure for a given cost of public funds. Section 4 addresses the nonparametric identification of the cost of public funds. Section 5 concludes. Two appendices contain the proofs of our results.

2 The Model

In this section we extend the Laffont and Tirole's (1986) model of incentive regulation of a monopolist producing a private good by allowing for general random demand and cost functions.[2] The demand for the private good and the cost for producing it are

$$\begin{aligned} y &= y(p, \epsilon_d) \\ c &= (\theta - e)c_o(y, \epsilon_c), \end{aligned}$$

where y is the quantity of private good, c is the corresponding cost, p is the price per unit of private good, θ represents the firm's (inefficiency) type, e is the level of effort exerted by the firm, and (ϵ_d, ϵ_c) are the random demand and cost shocks, respectively. As usual, θ and e are private information to the firm, where θ is the (scalar) adverse selection parameter known to be distributed as $F(\cdot)$. The random shocks (ϵ_d, ϵ_c) are known to be jointly distributed as $G(\cdot, \cdot)$. Note that ϵ_c is assumed independent of θ in Laffont and Tirole (1986), while ϵ_d is void because the demand y is fixed. Laffont and Tirole (1986) consider a constant marginal cost function, namely $c = (\theta - e)y + \epsilon_c$, while Laffont and

[2]This section owes much to Jean-Jacques Laffont's comments and arose from a term paper we wrote for a course he taught in Spring 2003.

Tirole (1993, p.171) consider the cost function $c = H(\theta - e)c_o(y) + \epsilon_c$, which is, except for the additive separability of ϵ_c, the one we consider with $H(\cdot)$ the identity function. See Perrigne and Vuong (2008a) for the nonidentification of $H(\cdot)$. The function $c_o(\cdot, \cdot)$ can be viewed as the *baseline cost* function, while $\theta - e$ can be interpreted as the firm's *cost inefficiency.* The following assumption is made.

Assumption A1: *The random demand and cost functions satisfy $y(\cdot, \cdot) \geq 0$, $c_o(\cdot, \cdot) \geq 0$ and $\theta - e > 0$. Moreover, the random shocks (ϵ_d, ϵ_c) are independent of θ, which is distributed as $F(\cdot)$ with density $f(\cdot) > 0$ on its support $[\underline{\theta}, \overline{\theta}]$, $\underline{\theta} < \overline{\theta}$.*

The regulator offers a price schedule $p(\tilde{\theta}) \geq 0$ based on the firms' announcement $\tilde{\theta}$ about its true type θ as well as a net transfer $t = t(\tilde{\theta}, c)$ based on $\tilde{\theta}$ and the observed *realized* firm's cost c. The realized cost is paid by the regulator so that t is the net transfer. In the Laffont and Tirole's (1986) model, ex post cost observability is used in the contract design to improve its efficiency by increasing the social welfare relative to the Baron and Myerson's (1982) model in which no ex post information is used. As noted by Joskow and Schmalensee (1986) and Baron (1989), assuming cost observability by the regulator is a reasonable assumption. As a matter of fact, firms are submitted to annual audit of their financial results and costs by regulatory commissions. The random shocks (ϵ_d, ϵ_c) are realized *ex post*, i.e. after contractual arrangements are made between the regulator and the monopolist. Consequently, the contract is designed *ex ante* based on expected values with respect to (ϵ_d, ϵ_c). Upon accepting the contract, the firm must satisfy the *realized* demand $y = y[p(\tilde{\theta}), \epsilon_d]$ at the price $p(\tilde{\theta})$ corresponding to its announcement $\tilde{\theta}$. The regulator and the firm are both risk neutral.

We assume that all functions are at least twice continuously differentiable and that integration and differentiation can be interchanged. Whenever $a(\cdot)$ is a function of more than one variable, we denote its derivative with respect to the kth argument by $a_k(\cdot)$.

2.1. The Firm's Problem

Given the price $p(\cdot)$ and transfer $t(\cdot, \cdot)$ chosen by the regulator, the realized utility for the firm with type θ when it announces $\tilde{\theta}$ and exerts effort e is

$$U(\tilde{\theta}, \theta, e, \epsilon_d, \epsilon_c) = t\Big(\tilde{\theta}, (\theta - e)c_o(y[p(\tilde{\theta}), \epsilon_d], \epsilon_c)\Big) - \psi(e), \tag{1}$$

where $\psi(e)$ is the firm's cost for exerting effort e. Because (ϵ_d, ϵ_c) is *ex ante* unknown and the firm is risk neutral, the firm's optimization problem is

$$(F) \qquad \max_{\tilde{\theta},e} \mathrm{E}[U(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c) \mid \theta] = \int U(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c),$$

where the independence between θ and (ϵ_d,ϵ_c) is used and $\mathrm{E}[\cdot]$ denotes the expectation with respect to (ϵ_d,ϵ_c).

The firm's optimization problem can be solved in two steps. In the first step, the effort level e is chosen optimally given the announcement $\tilde{\theta}$ and the true type θ

$$(FE) \qquad \max_{e} \mathrm{E}[U(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c) \mid \theta] = \int U(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c).$$

This gives $e = e(\tilde{\theta},\theta)$, which solves the first-order condition (FOC)

$$0 = \mathrm{E}[U_3(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c) \mid \theta] = \int U_3(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c), \tag{2}$$

i.e. using (1), $e = e(\tilde{\theta},\theta)$ solves

$$\int t_2\Big(\tilde{\theta},(\theta-e)c_o(y[p(\tilde{\theta}),\epsilon_d],\epsilon_c)\Big)c_o(y[p(\tilde{\theta}),\epsilon_d],\epsilon_c)\ dG(\epsilon_d,\epsilon_c) = -\psi'(e). \tag{3}$$

Denote the corresponding expected utility by

$$U(\tilde{\theta},\theta) \equiv \mathrm{E}[U(\tilde{\theta},\theta,e(\tilde{\theta},\theta),\epsilon_d,\epsilon_c) \mid \theta] = \int U(\tilde{\theta},\theta,e(\tilde{\theta},\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c). \tag{4}$$

In the second step, the firm solves $\max_{\tilde{\theta}} U(\tilde{\theta},\theta)$ giving $\tilde{\theta} = \tilde{\theta}(\theta)$, which solves the FOC $U_1(\tilde{\theta},\theta) = 0$.

2.2. Incentive Constraint

We now consider the Incentive Constraint (IC) arising from the firm telling the truth θ, i.e. $\theta = \tilde{\theta}(\theta)$ for any $\theta \in [\underline{\theta},\overline{\theta}]$. Thus, $U_1(\theta,\theta) = 0$ for any θ. Equivalently, denoting $U(\theta) \equiv U(\theta,\theta)$ and $e(\theta) \equiv e(\theta,\theta)$, and using $U'(\theta) = U_1(\theta,\theta) + U_2(\theta,\theta)$, we obtain

$$\begin{aligned} U'(\theta) &= U_2(\theta,\theta) \\ &= \int U_2(\theta,\theta,e(\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) + e_2(\theta,\theta)\int U_3(\theta,\theta,e(\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) \\ &= \int U_2(\theta,\theta,e(\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) \\ &= \int t_2\Big(\theta,(\theta-e(\theta))c_o(y[p(\theta),\epsilon_d],\epsilon_c)\Big)c_o(y[p(\theta),\epsilon_d],\epsilon_c)\ dG(\epsilon_d,\epsilon_c), \end{aligned}$$

where the third equality follows from (2) since $e(\theta) = e(\theta, \theta)$, and the fourth equality from (1). Hence, using (3) at $\tilde{\theta} = \theta$ and $e = e(\theta) = e(\theta, \theta)$ gives the incentive constraint

$$U'(\theta) = -\psi'(e), \tag{5}$$

where

$$\begin{aligned} U(\theta) &= \int t\Big(\theta, (\theta - e(\theta))c_o(y(p(\theta), \epsilon_d), \epsilon_c)\Big)dG(\epsilon_d, \epsilon_c) - \psi(e(\theta)) & (6) \\ e(\theta) &= \arg\max_e \int t\Big(\theta, (\theta - e)c_o(y(p(\theta), \epsilon_d), \epsilon_c)\Big)dG(\epsilon_d, \epsilon_c) - \psi(e). & (7) \end{aligned}$$

The Laffont and Tirole's (1986) model combines both adverse selection and moral hazard by reducing the latter to a "false" moral hazard problem.

2.3. The Regulator's Problem

The regulator chooses the price schedule and the transfer $[p(\cdot), t(\cdot, \cdot)]$. Suppose that $[p(\cdot), t(\cdot, \cdot)]$ is such that (i) it is truth telling, i.e. the incentive constraint (5) is satisfied and the firm exerts the optimal level of effort $e = e(\theta)$, and (ii) the firm participates for any level of its type θ. Given that the regulated good is private, the *ex post* social welfare when θ is the firm's true type is

$$\begin{aligned} SW(\theta, \epsilon_d, \epsilon_c) &= \int_{p(\theta)}^{\infty} y(v, \epsilon_d)dv + (1 + \lambda)\Big\{p(\theta)y(p(\theta), \epsilon_d) \\ &\quad -t(\theta, (\theta - e(\theta))c_o(y(p(\theta), \epsilon_d), \epsilon_c)) - (\theta - e(\theta))c_o(y(p(\theta), \epsilon_d), \epsilon_c)\Big\} \\ &\quad +t(\theta, (\theta - e(\theta))c_o(y(p(\theta), \epsilon_d), \epsilon_c)) - \psi(e(\theta)), \end{aligned}$$

where $\lambda > 0$ is the shadow cost of public funds as the payment requires raising taxes, which are costly to society. This social welfare corresponds to a benevolent utilitarian maximizer.[3] Thus, using the independence of θ and (ϵ_d, ϵ_c), the expected social welfare is

$$\begin{aligned} &\int SW(\theta, \epsilon_d, \epsilon_c)dG(\epsilon_d, \epsilon_c)dF(\theta) = \\ &\quad \int_{\underline{\theta}}^{\overline{\theta}} \Big\{ \int \Big[\int_{p(\theta)}^{\infty} y(v, \epsilon_d)dv + (1 + \lambda)\Big(p(\theta)y(p(\theta), \epsilon_d) \\ &\quad -\psi(e(\theta)) - (\theta - e(\theta))c_o(y(p(\theta), \epsilon_d), \epsilon_c)\Big)\Big]dG(\epsilon_d, \epsilon_c) - \lambda U(\theta)\Big\}dF(\theta), \end{aligned} \tag{8}$$

[3]Alternatively, following Baron and Myerson (1982), we could consider a weighted average of the consumer surplus and the firm's profit. Perrigne and Surana (2004) consider a weight that depends on political factors. See Section 4.3 for the identification of the weight α.

where we have used (6). Therefore, the regulator's optimization problem is

$$(P) \qquad \max_{[p(\cdot),t(\cdot,\cdot),e(\cdot),U(\cdot)]} \int SW(\theta,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c)dF(\theta),$$

subject to the incentive and the participation constraints

$$U'(\theta) = -\psi'(e(\theta)) \tag{9}$$

$$U(\theta) \geq 0, \tag{10}$$

for all $\theta \in [\underline{\theta},\overline{\theta}]$, where $U(\cdot)$ and $e(\cdot)$ are given by (6) and (7). Note that, without loss of generality, the control functions in the optimization problem (P) include $e(\cdot)$ and $U(\cdot)$ since these functions are determined by $p(\cdot)$ and $t(\cdot,\cdot)$ through (6) and (7). In view of (9), note also that $U'(\theta) < 0$ because $\psi'(\cdot) > 0$ (see Assumption A2 below). Hence, the participation constraint (10) can be written equivalently as $U(\overline{\theta}) \geq 0$ or

$$U(\overline{\theta}) = 0 \tag{11}$$

because the expected social welfare (8) decreases with $U(\cdot)$.

We now solve this optimization problem. First, we note that the objective function (8) depends on the transfer $t(\cdot,\cdot)$ only indirectly through $U(\theta)$ and $e(\theta)$, which are given by (6) and (7). This suggests to consider the simpler optimization problem

$$(P') \qquad \max_{[p(\cdot),e(\cdot),U(\cdot)]} \int SW(\theta,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c)dF(\theta),$$

subject to (9) and (11) only. In Section 2.4 on implementation, we will verify that there exists a transfer $t^*(\cdot,\cdot)$ satisfying (6) and (7) for the solution $[p^*(\cdot),e^*(\cdot),U^*(\cdot)]$ of the optimization problem (P').

We denote the *expected demand at price p* by $\overline{y}(p) \equiv \mathrm{E}[y(p,\epsilon_d)] = \int y(p,\epsilon_d)dG(\epsilon_d)$ where $G(\cdot)$ is the marginal distribution of ϵ_d. Let $\mathrm{E}[a(\epsilon_d,\epsilon_c,\theta)] = \int a(\epsilon_d,\epsilon_c,\theta)\ dG(\epsilon_d,\epsilon_c)$ denote the expectation of a function $a(\cdot,\cdot,\cdot)$ with respect to (ϵ_d,ϵ_c) for fixed θ, or conditional upon θ in view of the independence of θ and (ϵ_d,ϵ_c).

Proposition 1: *Given A1, the price $p^*(\cdot)$ and effort $e^*(\cdot)$ that solve the FOC of the optimization problem (P') satisfy*

$$\frac{p-\widetilde{mc}(p)}{p} = \mu\frac{1}{\tilde{\eta}(p)} \tag{12}$$

$$\psi'(e) = \overline{cs}_e(p) - \mu\frac{F(\theta)}{f(\theta)}\psi''(e), \tag{13}$$

where $p = p^*(\theta)$, $e = e^*(\theta)$, $\mu = \lambda/(1+\lambda)$ *and*

$$\begin{aligned}
\widetilde{mc}(p) &= \frac{\mathrm{E}\Big[(\theta - e)c_{o1}(y(p,\epsilon_d),\epsilon_c)y_1(p,\epsilon_d)\Big]}{\mathrm{E}[y_1(p,\epsilon_d)]} \\
\tilde{\eta}(p) &= -\frac{p\overline{y}'(p)}{\overline{y}(p)} \\
\overline{cs}_e(p) &= \mathrm{E}\Big[c_o(y(p,\epsilon_d),\epsilon_c)\Big].
\end{aligned}$$

Note that $\widetilde{mc}(p)$ differs from the expected marginal cost $\overline{mc}(p) \equiv \mathrm{E}[(\theta - e)c_{o1}(y(p,\epsilon_d), \epsilon_c)]$ for producing one additional unit to satisfy the random demand $y(p,\epsilon_d)$ at price p. Moreover, $\tilde{\eta}_p$ is the elasticity of the expected demand $\overline{y}(p)$, which differs from the expected elasticity of demand $\overline{\eta}(p) \equiv \mathrm{E}[-py_1(p\epsilon_d)/y(p,\epsilon_d)]$. On the other hand, $\overline{cs}_e(p)$ is the expected cost saving for one additional unit of effort at the random demand $y(p,\epsilon_d)$. Thus, (12) can be viewed as a *generalized Ramsey pricing*, while (13) is interpreted as usual with a downward distortion in effort due to the second term arising from asymmetric information because $\psi''(\cdot) > 0$ (see Assumption A2 below). In particular, when $\theta = \underline{\theta}$ so that $F(\theta) = 0$, (13) gives $\psi'(e) = \overline{cs}_e(p)$ and the first-best is achieved for the most "efficient" firm $\underline{\theta}$ as usual. Moreover, when the demand is not random, i.e. ϵ_d has a degenerate distribution so that $y(p,\epsilon_d) = \overline{y}(p)$, we have $\widetilde{mc}(p) = \overline{mc}(p)$ and $\tilde{\eta}(p) = \overline{\eta}(p)$ so that (12) and (13) reduce to the FOC in Laffont and Tirole (1986) with a constant marginal cost function and an additive random shock, namely $c = (\theta - e)y + \epsilon_c$.

Lastly, the optimization problem (P') is complete by determining the optimal firm's rent $U^*(\theta)$. The latter is obtained by integrating out the incentive constraint (9) subject to the participation constraint (11). This gives

$$U^*(\theta) = \int_{\theta}^{\overline{\theta}} \psi'[e^*(\beta)]\, d\beta, \tag{14}$$

which is strictly positive whenever $\theta < \overline{\theta}$ since $\psi'(\cdot) > 0$.

2.4. Implementation

Let the *expected baseline cost* for satisfying the random demand $y(p,\epsilon_d)$ at price p be $\overline{c}_o(p) \equiv E[c_o(y(p,\epsilon_d),\epsilon_c)\Big]$.

Proposition 2: *Given A1, consider the following transfer*

$$t^*(\tilde{\theta}, c) \quad = \quad A(\tilde{\theta}) - \psi'[e^*(\tilde{\theta})] \left\{ \frac{c}{\overline{c}_o[p^*(\tilde{\theta})]} - (\tilde{\theta} - e^*(\tilde{\theta})) \right\}, \tag{15}$$

where $e^*(\cdot)$ and $p^*(\cdot)$ are the optimal price and effort obtained from (P'), $\tilde{\theta}$ is the firm's announcement, c is the firm's realized cost, and

$$A(\tilde{\theta}) = \psi[e^*(\tilde{\theta})] + \int_{\tilde{\theta}}^{\overline{\theta}} \psi'[e^*(\beta)] \; d\beta. \tag{16}$$

Thus, given the price schedule $p^*(\cdot)$ and the transfer $t^*(\cdot, \cdot)$, announcing its true type θ and exerting the optimal effort $e^*(\theta)$ satisfy the FOC of the firm's problem (F). Moreover, $[p^*(\cdot), t^*(\cdot, \cdot), e^*(\cdot), U^*(\cdot)]$ solves the FOC of problem (P).

In view of (14), $A(\tilde{\theta}) = \psi[e^*(\tilde{\theta})] + U^*(\tilde{\theta})$. Hence, (16) shows that the transfer is equal to the cost of effort plus the firm's (expected) rent minus a fraction of the cost overrun, where the latter is the discrepancy between the realized cost and the expected cost. In particular, (16) can be viewed as a menu of linear cost-sharing contracts with slopes and intercepts depending on the firm's announcement $\tilde{\theta}$. Moreover, when $\theta = \underline{\theta}$, (13) and (15) imply that $\psi'[e^*(\underline{\theta})] = \overline{c}_o[p^*(\underline{\theta})]$ so that the slope coefficient in (16) equals -1 when $\tilde{\theta} = \underline{\theta}$. That is, recalling that t is the net transfer, the most efficient firm with type $\underline{\theta}$ chooses a fixed-price contract.

2.5. Second-Order Conditions

We verify that our solution corresponds to a global maximum. Announcing the true type θ must hold not only locally but globally. We verify ex post that our solution satisfies the second-order conditions (SOC) for a local maximum, and that these SOC extend globally.

We follow Laffont and Tirole (1986).[4] Let $V(p, \epsilon_d) = \int_p^\infty y(v, \epsilon_d) dv + (1 + \lambda) p y(p, \epsilon_d)$ be the social value of producing the quantity demanded at price p given demand shock ϵ_d The expected social value is then $\overline{V}(p) \equiv \int V(p, \epsilon_d) dG(\epsilon_d) = \int_p^\infty \overline{y}(v) dv + (1 + \lambda) p \overline{y}(p)$.

Assumption A2: *The demand, cost and effort functions satisfy:*
(i) $\overline{V}(p) > 0$, $\overline{V}'(\cdot) < 0$, $\overline{V}''(\cdot) < 0$,
(ii) $\overline{c}_o(\cdot) > 0$, $\overline{c}_o'(\cdot) < 0$, $\overline{c}_o''(\cdot) \geq 0$,
(iii) $\psi(\cdot) \geq 0$, $\psi'(\cdot) > 0$, $\psi''(\cdot) > 0$, $\psi'''(\cdot) \geq 0$.

[4] The following assumptions are sufficient but not necessary. In Section 4.1, we provide necessary and sufficient assumptions for the second-order conditions and implementation to hold.

Assumption A2-(i) is satisfied if the expected demand is not too inelastic, i.e. $\tilde{\eta}(p) > \lambda/(1+\lambda)$, and if the expected demand is not too convex, i.e. $-p\overline{y}''(p)/\overline{y}'(p) < (1+2\lambda)/(1+\lambda)$ whenever $\overline{y}(p) > 0$ and $\overline{y}'(p) < 0$. Similarly, A2-(ii) is satisfied if $c_{o1}(\cdot,\cdot) > 0$, $c_{o11}(\cdot,\cdot) \geq 0$, $y_1(\cdot,\cdot) < 0$ and $y_{11}(\cdot,\cdot) \geq 0$, i.e. the baseline cost function is strictly increasing and convex in quantity and demand is strictly decreasing and convex in price. In A2-(iii), the effort disutility is strictly increasing and strictly convex in e.

We begin with the firm's optimization problem (F). In particular, for any $(\tilde{\theta}, \theta)$ we consider the optimization problem (FE) with respect to e.

Lemma 1: *Suppose that the transfer $t(\cdot,\cdot)$ is weakly decreasing and concave in realized cost c. Given A1–A2, the effort $e(\tilde{\theta},\theta)$, which solves the FOC (3), is uniquely defined and corresponds to a global maximum of the problem (FE). Moreover, $0 \leq e_2(\theta,\theta) < 1$.*

Because $t^*(\cdot,\cdot)$ is weakly decreasing and concave in realized cost c, Lemma 1 applies.

We turn to the incentive constraint (5). The local SOC for $\tilde{\theta} = \theta$ to be a local maximum is $U_{11}(\theta,\theta) \leq 0$. Using the FOC $U_1(\theta,\theta) = 0$ for any θ, this SOC is equivalent to $U_{12}(\theta,\theta) \geq 0$. But differentiating (4) and using (1) give $U_2(\tilde{\theta},\theta) = \mathrm{E}\Big[t_2(\cdot)c_o(\cdot)[1 - e_2(\tilde{\theta},\theta)]\Big] - \psi'[e(\tilde{\theta},\theta)]e_2(\tilde{\theta},\theta) = -\psi'[e(\tilde{\theta},\theta)]$ using (3) and $e = e(\tilde{\theta},\theta)$. Hence $U_{12}(\tilde{\theta},\theta) = -\psi''[e(\tilde{\theta},\theta)]e_1(\tilde{\theta},\theta)$. Because $\psi''(\cdot) > 0$, the local SOC $U_{12}(\theta,\theta) \geq 0$ is equivalent to $e_1(\theta,\theta) \leq 0$, i.e. $e'(\theta) \leq e_2(\theta,\theta)$, since $e(\theta) = e(\theta,\theta)$ implies $e'(\theta) = e_1(\theta,\theta) + e_2(\theta,\theta)$.

When the transfer $t(\cdot,\cdot)$ is weakly decreasing and concave in c as in (15), Lemma 1 implies that a *sufficient* condition for the local SOC $e'(\theta) \leq e_2(\theta,\theta)$ to hold is that $e'(\cdot) \leq 0$. The next lemma shows that $e^{*\prime}(\cdot) < 0$ under the following assumption.

Assumption A3: *For any $\theta \in [\underline{\theta}, \overline{\theta}]$*
(i) $\psi''[e^*(\theta)]\overline{V}''[p^*(\theta)] + (1+\lambda)\Big\{\overline{c}_o'[p^*(\theta)]\Big\}^2 < 0$
(ii) $d[F(\theta)/f(\theta)]/d\theta \geq 0$.

Assumption A3-(i) is reminiscent of assumption 1-(iii) in Laffont and Tirole (1986), while A3-(ii) is the usual condition that $F(\cdot)$ is log-concave.

Lemma 2: *Given A1–A3, the transfer $t^*(\cdot,\cdot)$ and the price $p^*(\cdot)$, the local SOC (18) for truth telling is satisfied as $e^{*\prime}(\cdot) < 0$. Moreover, $p^{*\prime}(\cdot) > 0$.*

Effort $e^*(\cdot)$ and price $p^*(\cdot)$ are strictly decreasing and increasing in firm's type θ, respectively. It remains to show that $\tilde{\theta} = \theta$ provides a *global* maximum of the firm's utility (4) under the optimal transfer (15).

Proposition 3: *Given A1–A3, the transfer $t^*(\cdot,\cdot)$ and the price $p^*(\cdot)$, truth telling provides the global maximum of the expected utility $U(\tilde{\theta},\theta)$ given in (4). Moreover, the expected transfer $\bar{t} \equiv E\Big[t^*\Big(\theta,(\theta - e^*(\theta))c_o[y(p^*(\theta),\epsilon_d),\epsilon_c]\Big)\Big]$ is strictly decreasing in the firm's cost inefficiency $\theta - e^*(\theta)$.*

Proposition 3 ensures that the regulator can use the menu (15) of cost-sharing contracts. Because $\theta - e^*(\theta)$ is increasing in θ, the expected payment is decreasing in firm's type.

3 Identification Given λ

The structural approach relies on the assumption that the contract offered $[p(\cdot),t(\cdot)]$ is optimal. Upon the firm revealing its type, the model of Section 2 determines the price $p = p^*(\theta)$ per unit of good, the effort $e = e^*(\theta)$ exerted by the firm, the quantity $y = y(p,\epsilon_d)$ of produced good given the realized demand shock ϵ_d, the cost $c = (\theta - e)c_o(y,\epsilon_c)$ for producing y given the realized cost shock ϵ_c, as well as the (net) payment $t = t^*(\theta, c)$ to the firm. Thus, the structural approach leads to a closely related econometric model explaining (y,c,p,e,t) from the random variables $(\theta,\epsilon_d,\epsilon_c)$.

In this section we detail the specification of the econometric model for the observables taking into account possible observed and unobserved heterogeneity. We then study the identification of the structural elements of the model $[y(\cdot,\cdot),c_o(\cdot,\cdot),\psi(\cdot),F(\cdot),G(\cdot,\cdot),\lambda]$, i.e. the demand, baseline cost and effort disutility functions, the distribution of the firms' type, the distribution of the demand and cost shocks and the shadow cost of public funds from the distribution of the observables.

3.1. The Structural Econometric Model

The observables are (Y,C,P,T), i.e. the quantity of good, the production cost, the price per unit of product and the transfer/payment. Hereafter, we use capital letters to distinguish random variables from their realizations. The data correspond to that collected by regulatory commissions in their firms' audit process. As examples, we consider two industries: Water utilities studied by Wolak (1994) and Brocas, Chan and Perrigne (2006) and public transit studied by Gagnepain and Ivaldi (2002), Perrigne (2002) and Perrigne and Surana (2004). In the case of water utilities, Y represents the total amount of water (measured in hundred cubit feet HCF) delivered to residual consumers by a privately owned Class A Utility in California, P is the price per HCF charged to consumers including a fixed component called the access fee, C is the production cost for operating the water utility including the energy and labor costs as well as other cost components such as chemicals, pipeline repair, etc and T is obtained from a profit equation where the rate of return on firm's capital plays a key role in rewarding the firm for producing efficiently since water utilities are subject to rate of return regulation.[5] In the case of public transit, Y represents the number of passengers taking public transit operated by a firm in a medium-sized city in France, P is the average price paid by a passenger when taking the bus, C is the operating cost including energy and labor costs as well as various bus maintenance costs and T takes the form of a monetary transfer or subsidy paid to the firm by the transportation authority.[6] For other industries such as electricity, Joskow and Schmalensee (1986) note that utilities are required to make frequent reports on their cost, price, financial and production variables.

The demand, cost and effort disutility functions may depend on a vector of exogenous variables $Z \in I\!R^d$, where Z includes some characteristics of the firm and/or market. To allow for such dependencies, the demand, baseline cost and effort disutility functions are defined hereafter as $y(p, z, \epsilon_d)$, $c_o(y, z, \epsilon_c)$ and $\psi(e, z)$ when $Z = z$. For instance, Z may contain exogenous variables such as the community average income in the demand and input prices in the baseline cost. Similarly, the firm's type θ, the demand shock ϵ_d and the cost shock ϵ_c may depend on some firms' and market characteristics included in Z. This is accomplished by introducing the conditional distributions $F(\cdot|z)$ and $G(\cdot, \cdot|z)$ for θ and (ϵ_d, ϵ_c) given $Z = z$. Hereafter, we let $[\underline{\theta}(z), \overline{\theta}(z)]$ denote the support of $F(\cdot|z)$. Moreover, the cost of public funds λ may also depend on z, i.e. $\lambda = \lambda(z)$ for some positive function $\lambda(\cdot)$. For instance, the cost of public funds may depend on local economic conditions such as in Perrigne (2002). From such dependencies on z, it follows that price, transfer and

[5] Wolak (1994) and Brocas, Chan and Perrigne (2006) rely on Besanko's (1984) model in which the capital is used as a screening variable. In addition to the variables mentioned above, detailed information is provided on capital and its depreciation by the regulatory commission.

[6] Gagnepain and Ivaldi (2002), Perrigne (2002) and Perrigne and Surana (2004) rely on the Laffont and Tirole's (1986) model. In these data, public transit is heavily subsidized as revenues do not cover operating costs. Moreover, capital plays a minor role in this contract as capital (buses) is provided to the firm by the transportation authority.

effort functions are of the form $p^*(\cdot, z)$, $t^*(\cdot, \cdot, z)$ and $e^*(\cdot, z)$. The correspondingly revised assumptions A1–A3 are then assumed to hold for every value of Z. Hereafter, we let $\mathcal{Z}$ denote the support of the distribution of Z.

Two complications arise in the formulation of the econometric model. First, the effort exerted by the monopolist is unobserved as are the firm's type θ and the demand and cost shocks (ϵ_d, ϵ_c). Second, for every value of Z, the four observed endogenous variables (Y, C, P, T) are determined by the three unobserved random variables $(\theta, \epsilon_d, \epsilon_c)$. Thus, the econometric model is *singular*. In particular, the net transfer T is a deterministic function of (P, C, Z). Because $P = p^*(\theta, Z)$, which is strictly increasing in its first argument by Lemma 2, then $T = t^*(\theta, C, Z) = t^*[\theta^*(P, Z), C, Z]$, where $\theta^*(\cdot, Z)$ is the inverse of $p^*(\cdot, Z)$. Hence, if Z is observed together with (Y, C, P, T), in which case Z represents the observed heterogeneity, the structural model will be immediately rejected as soon as the observed values of (C, P, T, Z) do not lie perfectly on the surface $T = t^*[\theta^*(P, Z), C, Z]$.

To address such a difficulty, it is necessary to introduce another source of randomness. From an empirical point of view, several error terms could be introduced. For instance, the demand and baseline cost could be subject to some unobserved heterogeneity. The first-order conditions defining the price and effort level could also be considered as approximations and additive error terms in (12) and (13) can take into account (say) optimization errors as in Wolak (1994). In the case of public transit, the omitted benefit for pollution and traffic congestion reduction in the regulator's objective function can lead to such error terms on price and transfer. The joint distribution of these additional error terms is unidentified in a nonparametric setting, in which case the empirical analyst will have to consider a parameterization of the econometric model. Because we are interested in nonparametric identification, we adopt a parsimonious strategy by adding a single error term to the econometric model.

Specifically, we assume that the observed transfer T differs from the optimal transfer $T^* = t^*(\theta, C, Z)$ by an additive random term ϵ_t, which is unknown ex ante to the firm. There are several economic justifications to do so. First, such a random term ϵ_t may arise from measuring T^* with error, as data on transfers are likely to be imprecise. Second, ϵ_t may represent discretionary transfers from the regulator to the firm that do not rely on cost efficiency considerations. Third, ϵ_t may capture several factors that can influence the regulatory outcome in practice. As noted by Joskow (2005), political economy and legal considerations can affect regulation. For instance, interest groups, the commissioners' background and selection process among others are potentially affecting the regulatory contract and in particular the payment to the firm. The Laffont–Tirole (1986) model can then be viewed as providing a benchmark for the observed transfer, while ϵ_t can be used to assess deviations from the optimal transfer.

Rearranging (12) and (13), where $\widetilde{mc}(P)$ and $\overline{cs}_e(P)$ are replaced by $[(\theta-e)\overline{c}'_o(P,Z)]/\overline{y}'(P,Z)$ and $\overline{c}_o(P,Z)$, respectively, and combining (15) and (16) evaluated at $\tilde{\theta}=\theta$, the structural econometric model for the endogenous variables (Y,C,P,T) and the unobserved effort e given the exogenous variables Z is defined by the nonlinear nonparametric simultaneous equation model implicit in P with nonadditive error terms $(\theta,\epsilon_d,\epsilon_c)$

$$Y = y(P,Z,\epsilon_d) \tag{17}$$

$$C = (\theta-e)c_o(Y,Z,\epsilon_c) \tag{18}$$

$$P\overline{y}'(P,Z)+\mu\overline{y}(P,Z) = (\theta-e)\overline{c}'_o(P,Z) \tag{19}$$

$$\psi'(e,Z)+\mu\frac{F(\theta|Z)}{f(\theta|Z)}\psi''(e,Z) = \overline{c}_o(P,Z) \tag{20}$$

$$T = \psi(e,Z)+\int_{\theta}^{\overline{\theta}(Z)}\psi'[e^*(\tilde{\theta},Z),Z]d\tilde{\theta} - \psi'(e,Z)\left\{\frac{C}{\overline{c}_o(P,Z)}-(\theta-e)\right\}+\epsilon_t, \tag{21}$$

where $\mu=\mu(Z)$, $P=p^*(\theta,Z)$ and $e=e^*(\theta,Z)$ solve (19)-(20), and a prime denotes derivation with respect to the first argument of a function. Following Section 2, $\overline{y}(p,z)$ and $\overline{c}_o(p,z)$ in (19)–(21) are, conditional upon $Z=z$, the expected demand at price p and the expected baseline cost for producing the random quantity $y(p,\epsilon_d)$ at price p, i.e.

$$\overline{y}(p,z) = \int y(p,z,\epsilon_d)dG(\epsilon_d|z) \tag{22}$$

$$\overline{c}_o(p,z) = \int c_o[y(p,z,\epsilon_d),z,\epsilon_c]dG(\epsilon_d,\epsilon_c|z). \tag{23}$$

To complete the specification of the econometric model, we make the following assumption on the random elements $(\theta,\epsilon_d,\epsilon_c,\epsilon_t)$.

Assumption B1: *θ is independent of (ϵ_d,ϵ_c) conditional upon Z and $E[\epsilon_t|\theta,Z]=0$.*

The first part of B1 follows the theoretical model of Section 2. The condition $E[\epsilon_t|\theta,Z]=0$ holds under the normalization $\mathrm{E}[\epsilon_t|Z]=0$ and under the independence of ϵ_t from θ given

Z. Thus the deviation ϵ_t from the optimal transfer due to possible political and legal factors and/or the regulator's discretion is independent of the firm's efficiency θ.[7]

To summarize, the observables are (Y, C, P, T, Z), where the endogenous variables (Y, C, P, T) are determined by (17)–(21), while $(e, \theta, \epsilon_d, \epsilon_c, \epsilon_t)$ are unobserved. The structural elements of the model are the cost of public funds $\lambda(\cdot)$, the demand $y(\cdot,\cdot,\cdot)$, the baseline cost $c_o(\cdot,\cdot,\cdot)$, the effort disutility $\psi(\cdot,\cdot)$, the conditional type distribution $F(\cdot|\cdot)$ and the joint distribution $G(\cdot,\cdot,\cdot|\cdot)$ of the random terms $(\epsilon_d, \epsilon_c, \epsilon_t)$ given Z. In short, the structure of the model is given by the vector of six functions $[y, c_o, \psi, F, G, \lambda]$. The identification problem is to assess whether these structural elements can be recovered uniquely from the conditional distribution of (Y, C, P, T) given Z. For definitions of identification in nonparametric contexts, see e.g. Roehrig (1988) and Prakasa Rao (1992).

3.2. Identification of $\psi(\cdot,\cdot)$ and $F(\cdot|\cdot)$ Given $\lambda(\cdot)$

In this subsection we study the nonparametric identification of the effort disutility $\psi(\cdot,\cdot)$ and the conditional distribution of firms' type $F(\cdot|\cdot)$ assuming that the cost of public fund $\lambda(\cdot)$ is known. Identification of $\lambda(\cdot)$ is addressed in Section 4.

To begin, a *location-scale normalization* is necessary.

Lemma 3: *Let $\alpha = \alpha(\cdot) \geq 0$ and $\beta = \beta(\cdot) > 0$ be some functions of Z. Consider two structures $S \equiv [y, c_o, \psi, F, G, \lambda]$ and $\tilde{S} \equiv [\tilde{y}, \tilde{c}_o, \tilde{\psi}, \tilde{F}, \tilde{G}, \tilde{\lambda}]$ satisfying A1–A3 and B1, where $\tilde{y}(\cdot,\cdot,\cdot) = y(\cdot,\cdot,\cdot)$, $\tilde{c}_o(\cdot,\cdot,\cdot) = c_o(\cdot,\cdot,\cdot)/\beta$, $\tilde{\psi}(\cdot,\cdot) = \psi[(\cdot - \alpha)/\beta, \cdot]$, $\tilde{F}(\cdot|\cdot) = F[(\cdot - \alpha)/\beta|\cdot]$, $\tilde{G}(\cdot,\cdot,\cdot|\cdot) = G(\cdot,\cdot,\cdot|\cdot)$ and $\tilde{\lambda}(\cdot) = \lambda(\cdot)$. Thus, the structures S and $\tilde{S}$ lead to the same conditional distribution of (Y, C, P, T) given Z, i.e. they are observationally equivalent.*

As the proof of Lemma 3 indicates, the observational equivalence between S and $\tilde{S}$ arises because the unknown firm's type θ can be linearly transformed into a new type $\tilde{\theta} = \alpha(z) + \beta(z)\theta$ for each value z of Z. Several location-scale normalizations can be entertained. For instance, one can fix how two quantiles of θ vary with z. A natural choice for these quantiles are $\underline{\theta}(z)$ and $\overline{\theta}(z)$, which correspond to the most and least efficient firms,

[7] Though the residual transfer ϵ_t enters (additively) in the firm's utility (1), it disappears from the firm's optimization problem (F) given B1 and hence from the FOC of Proposition 1. Alternatively, if some characteristics of the regulator are observed, ϵ_t could be replaced by $m(Z)+\epsilon_t$, where $m(Z)$ captures the deterministic difference from the optimal transfer due to these exogenous regulator's characteristics. The identification of the function $m(\cdot)$ is discussed in Section 4.3.

respectively, when $Z = z$. In this case, a location-scale normalization would be to set $\underline{\theta}(z) = \underline{\theta}_o(z)$ and $\overline{\theta}(z) = \overline{\theta}_o(z)$, where $\underline{\theta}_o(\cdot)$ and $\overline{\theta}_o(\cdot)$ are known functions such as the zero and one functions respectively. Such a normalization, however, is not very convenient as we must have $\theta - e^*(\theta, z) \geq 0$ for all (θ, z) to ensure that $C = [\theta - e^*(\theta, Z)]c_o(Y, Z, \epsilon_c) \geq 0$.

A location-scale normalization with more economic content is obtained by imposing that the cost inefficiency of the most efficient firm is one and that the optimal effort of the least efficient firm is zero, irrespective of the value of Z. Formally, we impose

Assumption B2: For every value z of Z

$$\underline{\theta}(z) - e^*[\underline{\theta}(z), z] = 1 \qquad \text{and} \qquad e^*[\overline{\theta}(z), z] = 0. \tag{24}$$

Because the optimal effort $e^*(\theta, z)$ is strictly decreasing in θ, which implies that the cost inefficiency $\theta - e^*(\theta, z)$ is strictly increasing in θ, the normalization (24) actually determines $\underline{\theta}(z)$ and $\overline{\theta}(z)$ as in the preceding direct location-scale normalization, though (24) fixes those boundaries endogenously through the optimal effort function $e^*(\cdot, z)$. Moreover, the normalization (24) ensures that $\theta - e^*(\theta, z) \geq 1$ so that the cost frontier is defined by the most efficient firm, while $e^*(\theta, z) \geq 0$ for all firms, as desired. In other words, it follows that $c_o(y, z, \epsilon_c)$ can be interpreted as the *cost frontier* for producing y given (z, ϵ_c), while $\theta - e = [\theta - e^*(\theta, z)]/[\underline{\theta}(z) - e^*[\underline{\theta}(z), z]]$ can be viewed as the *relative cost inefficiency* of a firm with type θ relative to the efficient firm with type $\underline{\theta}(z)$.

We now turn to the nonparametric identification of the effort disutility $\psi(\cdot, \cdot)$ and the conditional type distribution $F(\cdot|\cdot)$. We need a preliminary result, which establishes that the expected demand $\overline{y}(\cdot, \cdot)$ and the expected baseline cost $\overline{c}_o(\cdot, \cdot)$ are identified nonparametrically from observations on quantity, price and costs given $\lambda(\cdot)$. Moreover, the relative cost inefficiency of the firm can be recovered uniquely from these observables. Let $[\underline{p}(z), \overline{p}(z)]$ denote the support of the conditional distribution $G_{P|Z}(\cdot|\cdot)$ of P given Z.

Lemma 4: *Suppose that A1–A3 and B1-B2 hold and $\lambda(\cdot)$ is known. Thus the expected demand $\overline{y}(\cdot, \cdot)$ and the expected baseline cost $\overline{c}_o(\cdot, \cdot)$ are uniquely determined by $\underline{p}(\cdot)$ and the conditional means of (Y, C) given (P, Z) as*

$$\overline{y}(p, z) = \mathrm{E}[Y|P = p, Z = z] \tag{25}$$

$$\overline{c}_o(p, z) = \mathrm{E}[C|P = \underline{p}(z), Z = z] \exp\left\{\int_{\underline{p}(z)}^{p} \frac{\tilde{p}\overline{y}'(\tilde{p}, z) + \mu\overline{y}(\tilde{p}, z)}{\mathrm{E}[C|P = \tilde{p}, Z = z]} d\tilde{p}\right\}, \tag{26}$$

where $\mu = \mu(z)$. *Moreover, the relative cost inefficiency is identified as*

$$\theta - e^*(\theta, z) = \Delta(p, z) \equiv \mathrm{E}[C|P\!=\!p, Z\!=\!z]/\overline{c}_o(p, z), \tag{27}$$

where $p = p^*(\theta, z)$, *and the function* $\Delta(\cdot, \cdot)$ *satisfies* $\Delta(\cdot, \cdot) \geq 1$ *and* $\partial\Delta(\cdot, \cdot)/\partial p > 0$.

It is interesting to note that the expected demand (25) can be obtained by a simple regression of Y on (P, Z) despite the possible correlation between the demand shock ϵ_d and P through Z in the demand (17) as B1 only ensures that ϵ_d is independent of P given Z. On the other hand, a simple regression of C (or $\log C$) on (P, Z), as used in the estimation of production/cost frontier (see, e.g. Gagnepain and Ivaldi (2002)), does *not* estimate the expected baseline cost (26). For instance, consider the typical cost specification $\log C = s(Y, Z) + \epsilon_c + \log(\theta - e)$, where $\log(\theta - e) \geq 0$ in view of (24). The composite error term $\epsilon_c + \log(\theta - e)$ is typically correlated with both $Y = y(P, Z, \epsilon_d)$ and Z under B1. IV or ML methods estimate the cost frontier $s(y, z)$, which is different from the expected baseline cost $\overline{c}_o(p, z)$ that is relevant in the FOC (19)-(20) determining price and effort. See Perrigne and Vuong (2008b). Nevertheless, by exploiting the generalized Ramsey pricing rule (19), Lemma 4 indicates that the expected baseline cost $\overline{c}_o(\cdot, \cdot)$ can be estimated from (26) by combining appropriately the regressions of Y and C on (P, Z) with the knowledge of $\underline{p}(\cdot)$.[8] Moreover, (27) shows that a firm's relative cost inefficiency $\theta - e^*(\theta, z)$ can be recovered from its observed pair (p, z) as $\Delta(\cdot, \cdot)$ is known from the regression of C given (P, Z) and the expected baseline cost $\overline{c}_o(\cdot, \cdot)$.

Using Lemma 4, the next result establishes the nonparametric identification of the effort disutility $\psi(\cdot, \cdot)$ and the conditional type distribution $F(\cdot|\cdot)$ from observations on quantity, price, cost and transfer given $\lambda(\cdot)$. To this end, we use an identification strategy in the spirit of Guerre, Perrigne and Vuong (2000). Specifically, we exploit the bijective mapping between the price P and the firm's type θ from Lemma 2. The parallel with auction models becomes clear. In auction models, the bijective mapping between the bidder's (unobserved) private value and his optimal (observed) bid is used to rewrite the FOC of the bidder's optimization problem in terms of observables. Such an equation

[8]This result is reminiscent to recovering the marginal cost in mark-up models. See Bresnahan (1989). The generalized Ramsey pricing rule plays a similar role as the mark-up equation with the exception that recovering $\overline{c}_o(\cdot, \cdot)$ is somewhat more complicated here because of the cost inefficiency term $\theta - e$. See the proof of Lemma 4.

expresses the unobserved private value in terms of the corresponding optimal bid, the bid distribution and density, from which one can identify the private value distribution. A similar strategy is used here. In particular, because $\theta^*(P,Z)=\theta$, we replace in (20) the ratio $F(\theta|Z)/f(\theta|Z)$ by $[G_{P|Z}(P|Z)/g_{P|Z}(P|Z)] \times (\partial\theta^*(P,Z)/\partial p)$, where $G_{P|Z}(\cdot|\cdot)$ is the conditional distribution of P given Z and $g_{P|Z}(\cdot|\cdot)$ its corresponding density. Using (20) and the identification of $\psi(\cdot,\cdot)$ from (21), we derive an expression for θ as a function of the observed price, its distribution and density from which we identify the type distribution $F(\cdot|\cdot)$ as shown in the next proposition. In particular, the unobserved firm's type θ can be recovered uniquely from the observed pair (p,z) once the various unknown functions have been recovered from data on (Y,C,P,T,Z).

We define the functions

$$\Gamma(p,z) = -\frac{\partial \mathrm{E}[T|P=p,Z=z]/\partial p}{\partial\Delta(p,z)/\partial p} \tag{28}$$

$$R(p,z) = \frac{\mu[G_{P|Z}(p|z)/g_{P|Z}(p|z)] \times \partial\Gamma(p,z)/\partial p \times \partial\Delta(p,z)/\partial p}{\Gamma(p,z) - \overline{c}_0(p,z) + \mu[G_{P|Z}(p|z)/g_{P|Z}(p|z)] \times \partial\Gamma(p,z)/\partial p}, \tag{29}$$

for an arbitrary value (p,z). The functions $\Gamma(\cdot,\cdot)$ and $R(\cdot,\cdot)$ are known from the knowledge of the joint distribution of (Y,C,P,T) conditional upon Z in view of Lemma 4. In particular, $\Gamma(\cdot,\cdot)$ can be interpreted as the marginal decrease in the expected transfer T due to a one-unit increase in relative cost inefficiency Δ. As (30) and (31) below show, $\Gamma(p,z)$ is also the marginal cost of effort, while $R(p,z)$ is the marginal decrease in effort due to a one-unit increase in price.

Proposition 4: *Suppose that A1–A3 and B1-B2 hold and $\lambda(\cdot)$ is known. Thus the effort disutility $\psi(\cdot,\cdot)$ is uniquely determined by $\underline{p}(\cdot)$, $\overline{p}(\cdot)$ and the conditional means of (Y,C,T) given (P,Z) as*

$$\psi(e,z) = \mathrm{E}[T|P=\overline{p}(z),Z=z] + \int_0^e \Gamma[p^*(\tilde{e},z),z]\ d\tilde{e}, \tag{30}$$

where $\Gamma(\cdot,\cdot)>0$, $\partial\Gamma(\cdot,\cdot)/\partial p<0$, and $p^*(\cdot,z)$ is the inverse of the optimal effort function $e^*(\cdot,z)$, which satisfies

$$e^*(p,z) = \int_p^{\overline{p}(z)} R(\tilde{p},z)\ d\tilde{p}, \tag{31}$$

with $\partial\Delta(\cdot,\cdot)/\partial p > R(\cdot,\cdot) > 0$. Moreover, the conditional means of (Y,C,T) given (P,Z) and the conditional distribution of P given Z uniquely determine the conditional type *distribution $F(\cdot|z)$ given $Z=z$ as the distribution of*

$$\theta = \theta^*(P,z) \equiv \Delta(P,z) + \int_P^{\overline{p}(z)} R(\tilde{p},z)\ d\tilde{p}, \tag{32}$$

where P is distributed as $G_{P|Z}(\cdot|z)$ for every value z of Z.

The key of Proposition 4 is that the observed price P is in bijection with the unobserved type θ given $Z=z$. Thus, conditioning on (P,Z) is actually conditioning on (θ, Z). That is, (32) can be viewed as *the inverse of the optimal price schedule* $p^*(\theta,z)$. A similar remark applies to the firm's effort e, which can be recovered similarly through (31) from the observed pair (p,z). In particular, while the minimal effort $e^*[\overline{\theta}(z),z]=0$ by the normalization (24), (31) implies that the maximal effort (exerted by the efficient firm with type $\underline{\theta}(z)$) is

$$e^*[\underline{\theta}(z),z] = \int_{\underline{p}(z)}^{\overline{p}(z)} R(\tilde{p},z)\ d\tilde{p} > 0. \tag{33}$$

Similarly, the lower and upper bounds of the conditional distribution $F(\cdot|z)$ of type are

$$\underline{\theta}(z) \;=\; 1 + \int_{\underline{p}(z)}^{\overline{p}(z)} R(\tilde{p},z)\ d\tilde{p} > 1 \tag{34}$$

$$\overline{\theta}(z) \;=\; \frac{\mathrm{E}[C|P=\overline{p}(z),Z=z]}{\mathrm{E}[C|P=\underline{p}(z),Z=z]} \exp\left\{ -\int_{\underline{p}(z)}^{\overline{p}(z)} \frac{\tilde{p}\overline{y}'(\tilde{p},z)+\mu\overline{y}(\tilde{p},z)}{\mathrm{E}[C|P=\tilde{p},Z=z]} d\tilde{p} \right\} > \underline{\theta}(z), \tag{35}$$

from (32) in view of (26)–(27) and $\Delta[\underline{p}(z),z]=1$ from (24). The inequality in (35) follows from $\theta=\theta(p,z)$, where $\partial\theta^*(p,z)/\partial p = \partial\Delta(p,z)/\partial p - R(p,z) > 0$.

3.3. Identification of $y(\cdot,\cdot,\cdot)$, $c_o(\cdot,\cdot,\cdot)$ and $G(\cdot,\cdot,\cdot|\cdot)$ Given $\lambda(\cdot)$

Lemma 4 establishes the identification of the expected demand and expected baseline cost $\overline{y}(\cdot,z)$ and $\overline{c}_o(\cdot,z)$ for every price $p \in [\underline{p}(z),\overline{p}(z)]$ and $z \in \mathcal{Z}$. For counterfactual exercises or policy evaluations, one may need to identify the remaining structural elements of the model, namely the demand $y(\cdot,\cdot,\cdot)$, the baseline cost $c_o(\cdot,\cdot,\cdot)$ and the conditional distribution $G(\cdot,\cdot,\cdot|\cdot)$ of $(\epsilon_d,\epsilon_c,\epsilon_t)$ given Z. This is the purpose of this subsection still assuming that the cost of public fund $\lambda(\cdot)$ is known.

Unlike the random term ϵ_t, which enters additively in the transfer (21), the demand shock ϵ_d and cost shock ϵ_c do not enter additively in the demand (17) and cost (18). If one is willing to consider a specification of the demand and baseline cost in which the random shocks ϵ_d and ϵ_c enter additively so that $y(p, z, \epsilon_d) = \delta(p, y) + \epsilon_d$ and $c_o(y, z, \epsilon_c) = \gamma(y, z) + \epsilon_c$, identification becomes standard under the usual normalizations $\mathrm{E}[\epsilon_d|Z] = 0$ and $\mathrm{E}[\epsilon_c|Z] = 0$. Specifically, under B1 we have ϵ_d independent of P given Z implying $\mathrm{E}[\epsilon_d|P, Z] = \mathrm{E}[\epsilon_d|Z] = 0$. Hence $\mathrm{E}[Y|P = p, Z = z] = \delta(p, z)$ showing that $y(\cdot, \cdot, \cdot)$ and $G_{\epsilon_d|Z}(\cdot|\cdot)$ are identified. Moreover, let the (random) *baseline cost* be $C_o = c_o(Y, Z, \epsilon_c) = C/(\theta - e^*(\theta, Z)) = C/\Delta(P, Z)$. Thus C_o can be recovered from (C, P, Z) as $\Delta(\cdot, \cdot)$ is identified by Lemma 4. Now, using the control function approach to endogeneity (see Blundell and Powell (2003)), we have $\mathrm{E}[C_o|Y, P, Z] = \gamma(Y, Z) + \mathrm{E}[\epsilon_c|\epsilon_d, P, Z] = \gamma(Y, Z) + \rho(\epsilon_d, Z)$, where the second equality follows from the independence of ϵ_c and P given (ϵ_d, Z) from B1. But $\mathrm{E}[\rho(\epsilon_d, Z)|Z] = \mathrm{E}[\mathrm{E}(\epsilon_c|\epsilon_d, Z)|Z] = \mathrm{E}[\epsilon_c|Z] = 0$. Hence, conditional on $Z = z$ the nonparametric regression $\mathrm{E}[C_o|Y, P, Z = z] = \gamma(Y, z) + \rho(\epsilon_d, z)$ is separately additive in Y and ϵ_d with the normalization $\mathrm{E}[\rho(\epsilon_d, z)|Z = z] = 0$ thereby showing that $\gamma(\cdot, \cdot)$ and hence $c_o(\cdot, \cdot, \cdot)$ are identified. See Hastie and Tibshirani (1990). The identification of $G_{\epsilon_c|\epsilon_d, Z}(\cdot|\cdot, \cdot)$ follows.

As Matzkin (2003) argues, however, the structural specification of a random demand or cost function seldom leads to an additive error term. When the random term does not enter additively into the relationship between the endogenous variable and the exogenous variables, Matzkin (2003) shows that this relationship is nonidentified and proposes several normalizations to identify nonparametrically the function and the distribution of the error term. For instance, consider the demand $Y = y(P, Z, \epsilon_d)$, where $P = p^*(\theta, Z)$ is independent of ϵ_d given Z in view of B1. Clearly, $y(\cdot, \cdot, \cdot)$ is nonidentified as a monotonic transformation of ϵ_d can be compensated by an appropriate transformation of $y(\cdot, \cdot, \cdot)$. Our problem differs from Matzkin's framework in two aspects. First, Section 3.2 shows that the expected demand $\overline{y}(\cdot, \cdot)$ and the expected baseline cost $\overline{c}_o(\cdot, \cdot)$ are identified. A natural question is whether the knowledge of such functions can help to identify $y(\cdot, \cdot, \cdot)$ and $c_o(\cdot, \cdot, \cdot)$ and the joint distribution of error terms $G_{\epsilon_d, \epsilon_c|Z}(\cdot, \cdot|\cdot)$. Second, we have a simultaneous equation model, which creates potential endogeneity as Y in the baseline cost may be correlated with the error term ϵ_c through ϵ_d. For a recent contribution to the nonparametric identification of nonlinear simultaneous equation model with nonadditive error terms, see Matzkin (2005).

The next lemma establishes that the knowledge of the expected demand and baseline cost does not help in identifying the desired functions and distributions. We need first to introduce some notations and to make some assumptions. Let $[\underline{\epsilon}_d(z), \overline{\epsilon}_d(z)] \times [\underline{\epsilon}_c(z), \overline{\epsilon}_c(z)]$ be the support of the conditional distribution $G_{\epsilon_d,\epsilon_c|Z}(\cdot,\cdot|\ z)$ of (ϵ_d, ϵ_c) given $Z = z$. Similarly, let $[\underline{y}(z), \overline{y}(z)]$ and $[\underline{y}(p,z), \overline{y}(p,z)]$ denote the supports of the conditional distributions $G_{Y|Z}(\cdot|z)$ and $G_{Y|P,Z}(\cdot|p,z)$ of $Y = y(P, Z, \epsilon_d)$ given $Z = z$ and $(P, Z) = (p, z)$, respectively. We make the following normalizations, while imposing usual strict monotonicity conditions on the demand and cost shocks (ϵ_d, ϵ_c).

Assumption B3:
(i) For all $z \in \mathcal{Z}$, and all $(\epsilon_d, \epsilon_c) \in [\underline{\epsilon}_d(z), \overline{\epsilon}_d(z)] \times [\underline{\epsilon}_c(z), \overline{\epsilon}_c(z)]$, there exist $p_o(z) \in [\underline{p}(z), \overline{p}(z)]$ and $y_o(z) \in [\underline{y}(z), \overline{y}(z)]$ such that

$$y[p_o(z), z, \epsilon_d] = \epsilon_d \qquad \text{and} \qquad c_o[y_o(z), z, \epsilon_c] = \epsilon_c, \tag{36}$$

where $p_o(\cdot)$ and $y_o(\cdot)$ are known. Moreover, there exists $p_\dagger(z, \epsilon_d) \in [\underline{p}(z), \overline{p}(z)]$ such that $y_o(z) = y[p_\dagger(z, \epsilon_d), z, \epsilon_d]$.
(ii) The demand and baseline cost $y(p, z, \cdot)$ and $c_o(y, z, \cdot)$ are strictly increasing in ϵ_d and ϵ_c, respectively, for all values (y, p, z), while the conditional distributions $G_{\epsilon_d|Z}(\cdot|\cdot)$ and $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\cdot,\cdot)$ of (ϵ_d, ϵ_c) are nondegenerated and strictly increasing in their first arguments.

Except for the second part of B3-(i), B3 follows Matzkin's (2003) first normalization though it is slightly more general by allowing p_o and y_o to depend on z. If one considers constant values p_o and y_o, as in Matzkin (2003), these values may not satisfy $p_o \in [\underline{p}(z), \overline{p}(z)]$ and $y_o \in [\underline{y}(z), \overline{y}(z)]$ for any $z \in \mathcal{Z}$.[9] The second part of B3-(i) is used only to establish Proposition 5-(ii) and is discussed later.

Lemma 5: *Let $Y = y(P, Z, \epsilon_d)$ and $C_o = c_o(Y, Z, \epsilon_c)$ satisfy B3-(ii) with P conditionally independent of (ϵ_d, ϵ_c) given Z. There exists an observationally equivalent system $Y = \tilde{y}(P, Z, \tilde{\epsilon}_d)$ and $C_o = \tilde{C}_o(Y, Z, \tilde{\epsilon}_c)$ with P conditionally independent of $(\tilde{\epsilon}_d, \tilde{\epsilon}_c)$ given Z satisfying B3-(ii) and*

$$\overline{y}(p, z) = \overline{\tilde{y}}(p, z), \quad \overline{c}_o(y, z) = \overline{\tilde{c}}_o(y, z)$$

[9] The second normalization in Matzkin (2003) relates to the homogeneity of degree one in z and ϵ. Such a restriction seems natural for the baseline cost, which is homogenous of degree one in input prices. See also Matzkin (1994). This restriction would require to choose some values or functions z_o, ϵ_{co}, C_o and a factor of homogeneity γ. Moreover, because Matzkin's proof relies on using $\gamma = \epsilon_c/\epsilon_{co}$, the conditional distribution of ϵ_c can be recovered along a specific value of z, i.e. $(\epsilon_c z)/\epsilon_{co}$. Thus $G_{\epsilon_d,\epsilon_c|z}(\cdot,\cdot|\cdot)$ cannot be identified everywhere.

$$\tilde{y}(p_o(z), z, \tilde{\epsilon}_d) = \tilde{\epsilon}_d, \quad \tilde{c}_o(y_o(z), z, \tilde{\epsilon}_c) = \tilde{\epsilon}_c$$

for any (p, z) *and* $(\tilde{\epsilon}_d, \tilde{\epsilon}_c)$*, where* $p_o(z) \in [\underline{p}(z), \overline{p}(z)]$ *and* $y_o(z) \in [\underline{y}(z), \overline{y}(z)]$ *are arbitrary.*

Lemma 5 shows that the normalization B3 does not entail any loss of generality since $p_o(\cdot)$ and $y_o(\cdot)$ can be chosen arbitrarily. This argument can be easily seen in the demand case when the error term is additive. For instance, let $\tilde{y}(p, z, \tilde{\epsilon}_d) = \tilde{y}(p, z) + \tilde{\epsilon}_d$, where $\tilde{y}(p, z) = y(p, z) - y(p_o(z), z)$ and $\tilde{\epsilon}_d = \epsilon_d + y(p_o(z), z)$ for an arbitrary $p_o(z) \in [\underline{p}(z), \overline{p}(z)]$. Thus, we have $Y = y(p, z, \epsilon_d) = \tilde{y}(p, z, \tilde{\epsilon}_d)$ and $\tilde{y}(p_o(z), z) = 0$ leading to $\tilde{y}(p_o(z), z, \tilde{\epsilon}_d) = \tilde{\epsilon}_d$ thereby satisfying B3. In this sense, the term normalization is appropriate. Though quite intuitive as the distribution of $y(\cdot, \cdot, \cdot)$ at $p_o(\cdot)$ reduces to that of the error term given $Z = z$, the choice of $p_o(\cdot)$ may be difficult to the analyst. Lemma 5 shows that one should not be concerned by the choice of $p_o(\cdot)$ as long as it satisfies $p_o(z) \in [\underline{p}(Z), \overline{p}(Z)]$. A similar remark applies to $c_o(\cdot, \cdot, \cdot)$ and $y_o(\cdot)$. Moreover, Lemma 5 shows that the knowledge of the expected demand and expected baseline cost does not help in identifying the demand and baseline cost as well as the joint distribution of error terms.

The next result establishes the nonparametric identification of the demand $y(\cdot, \cdot, \cdot)$, the baseline cost $c_o(\cdot, \cdot, \cdot)$ and the conditional distribution $G(\cdot, \cdot|\cdot)$ of (ϵ_d, ϵ_c) given Z from the quantiles of quantity and baseline cost given (P, Z). Recall that the baseline cost $C_o = c_o(Y, Z, \epsilon_c)$ can be recovered. Thus its conditional distribution $G_{C_o|Y,P,Z}(\cdot|\cdot, \cdot, \cdot)$ given (Y, P, Z) is identified. Formally, because $\Delta(P, Z) \geq 1$, we have $G_{C_o|Y,P,Z}(c_o|y, p, z) = G_{C|Y,P,Z}[c_o \Delta(p, z)|y, p, z]$ for any c_o, where $G_{C|Y,P,Z}(\cdot|\cdot, \cdot, \cdot)$ is the conditional distribution of C given (Y, P, Z).

Proposition 5: *Suppose that A1–A3 and B1–B3 hold and* $\lambda(\cdot)$ *is known.*
(i) The demand $y(\cdot, \cdot, \cdot)$ *and the conditional distribution* $G_{\epsilon_d|Z}(\cdot|\cdot)$ *of* ϵ_d *given* Z *are uniquely determined by the conditional distribution* $G_{Y|P,Z}(\cdot|\cdot, \cdot)$ *as*

$$y(p, z, \epsilon_d) = G^{-1}_{Y|P,Z} \left\{ G_{Y|P,Z}[\epsilon_d|p_o(z), z]|p, z \right\} \quad (37)$$

$$G_{\epsilon_d|Z}(\cdot|z) = G_{Y|P,Z}[\,\cdot\,|p_o(z), z]. \quad (38)$$

(ii) The baseline cost $c_o(\cdot, \cdot, \cdot)$ *and the conditional distribution* $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\cdot, \cdot)$ *of* ϵ_c *given* (ϵ_d, Z) *are uniquely determined by the conditional distribution* $G_{C_o|Y,P,Z}(\cdot|\cdot, \cdot, \cdot)$ *as*

$$c_o(y, z, \epsilon_c) = G^{-1}_{C_o|Y,P,Z} \left\{ G_{C_o|Y,P,Z}[\epsilon_c|y_o(z), p_\dagger(z, \epsilon_d), z]|y, p, z \right\} \quad (39)$$

$$G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d, z) = G_{C_o|Y,P,Z}[\,\cdot\,|y_o(z), p_\dagger(z, \epsilon_d), z], \quad (40)$$

where $p_\dagger(\cdot,\cdot)$ *is identified and* $y = y(p, z, \epsilon_d)$.

The proof of (i) follows Matzkin (2003) as θ and hence P are independent of ϵ_d given Z by B1. On the other hand, though $Y = y(P, Z, \epsilon_d)$ is not independent of ϵ_c given Z, the proof of (ii) is only slightly more involved as it exploits the second part of B3-(i). This part says that for any (z, ϵ_d) there exists a price $p_\dagger(z, \epsilon_d)$ for which the output $y[p_\dagger(z, \epsilon_d), z, \epsilon_d]$ is equal to the reference output $y_o(z)$ of B3. As $y_o(z)$ can be chosen arbitrarily by Lemma 5, this condition actually requires that $\bigcap_{\epsilon_d \in [\underline{\epsilon}_d(z), \overline{\epsilon}_d(z)]} \{y = y(p, z, \epsilon_d), p \in [\underline{p}(z), \overline{p}(z)]\}$ is nonempty for every $z \in \mathcal{Z}$.[10]

Lastly, the conditional distribution $G_{\epsilon_t|\epsilon_d,\epsilon_c,Z}(\cdot|\cdot,\cdot,\cdot)$ of ϵ_t given $(\epsilon_d, \epsilon_c, Z)$ is identified from observations on (Y, C, P, T, Z). This follows immediately from (21), since ϵ_t can be expressed as a function of (C, P, T, Z) that is identified by Lemma 4 and Proposition 4. For a simpler expression than (21), see Lemma 8 below. Moreover, because $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\cdot,\cdot)$ and $G_{\epsilon_d|Z}(\cdot|\cdot)$ are identified by Proposition 5, then the joint distribution $G(\cdot,\cdot,\cdot|\cdot)$ of $(\epsilon_d, \epsilon_c, \epsilon_t)$ given Z is identified.

4 The Cost of Public Funds

The previous identification results can be used when the cost of public funds $\lambda(\cdot)$ is known. Several studies provide estimates for λ based on general equilibrium models. For instance, Ballard, Shoven and Whalley (1985) find that the cost of public funds in the US ranges from 0.17 to 0.56 per dollar, while Jorgenson and Yun (1991) find 0.46 per dollar. See also Warlters and Auriol (2007) for a recent survey on estimates of λ in a wide range of countries. While a value around 0.3 is well accepted in western economies, it is generally agreed that it is much higher (above 1) in developing countries. See Laffont (2005). Microeconomic data can provide a new perspective to these estimates. In particular, one can expect that local economic conditions may affect the cost of public funds. See Perrigne (2002) for empirical evidence with the local unemployment rate. In this case,

[10]The latter can be relaxed in which case we obtain only partial identification of $c_o(\cdot,\cdot,\cdot)$ and $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\cdot,\cdot)$. Specifically, from the proof of (ii), one obtains the identification of $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d, z)$ for those values of (ϵ_d, z) for which there exists a price $p_\dagger(z, \epsilon_d)$ satisfying $y_o(z) = y[p_\dagger(z, \epsilon_d), z, \epsilon_d]$. Similarly, one obtains the identification of $c_o(y, z, \epsilon_c)$ for those values of (y, z), where $y = y(p, z, \epsilon_d)$ and (z, ϵ_d) satisfies $y_o(z) = y[p_\dagger(z, \epsilon_d), z, \epsilon_d]$.

identification of the cost of public fund as a function $\lambda(\cdot)$ of some characteristics Z is of interest. This is the purpose of this section. We first characterize the restrictions of the model on observables given $\lambda(\cdot)$. We show that $\lambda(\cdot)$ is not identified since the restrictions of the model are relatively weak. We then discuss several identifying conditions for $\lambda(\cdot)$ and derive a simple expression for it. This section concludes with a discussion of how the regulator's objective can be modified to take into account other factors that may affect price and payment.

4.1. Model Restrictions and Nonidentification of $\lambda(\cdot)$

To address the nonidentification of $\lambda(\cdot)$, we first define the error terms as *identified* functions of the observables given $\lambda(\cdot)$. Specifically, from (37) and (39), the error terms ϵ_d and ϵ_c can be expressed as identified functions of (Y, P, Z) and (Y, C, P, Z), namely $\phi_d(Y, P, Z)$ and $\phi_c(Y, C, P, Z)$, respectively, where we use $C_o = C/(\theta - e^*(\theta, Z))$ and $\theta - e^*(\theta, Z)$ given by (27). Using (21), (26), (27), and (30), the transfer error term ϵ_t can be expressed as an identified function of (Y, C, P, T, Z), namely $\phi_t(Y, C, P, T, Z)$. Their expressions are given in the next lemma.

Lemma 6: *Given A1–A3 and B1–B3, the error terms $(\epsilon_d, \epsilon_c, \epsilon_t)$ are given by*

$$\begin{aligned}
\epsilon_d &= \phi_d(Y, P, Z) \equiv G^{-1}_{Y|P,Z}\Big(G_{Y|P,Z}(Y|P, Z)|p_o(Z), Z\Big) &(41)\\
\epsilon_c &= \phi_c(Y, C, P, Z)\\
&\equiv \frac{1}{\Delta(p_\dagger(Z, \epsilon_d), Z)} G^{-1}_{C|Y,P,Z}\Big(G_{C|Y,P,Z}(C|Y, P, Z)|y_o(Z), p_\dagger(Z, \epsilon_d), Z\Big) &(42)\\
\epsilon_t &= \phi_t(Y, C, P, T, Z)\\
&\equiv T - \mathrm{E}[T|P, Z] + \frac{\partial \mathrm{E}[T|P, Z]/\partial p}{\partial \mathrm{E}[C|P, Z]/\partial p - [P\overline{y}'(P, Z) + \mu \overline{y}(P, Z)]}\Big(C - \mathrm{E}[C|P, Z]\Big), &(43)
\end{aligned}$$

where $\Delta(P, Z)$ is given by (27).

Next we extend A2-A3. Our new assumptions are expressed in terms of the observables (Y, C, P, T, Z). We note that $\underline{\epsilon}_d(z) = G^{-1}_{Y|P,Z}(0|p_o(z), z) \equiv \underline{y}(p_o(z), z)$ and $\overline{\epsilon}_d(z) = G^{-1}_{Y|P,Z}(1|p_o(z), z) \equiv \overline{y}(p_o(z), z)$.

Assumption C1: *Let $\bigcap_{\epsilon_d \in [\underline{y}(p_o(z),z), \overline{y}(p_o(z),z)]}\{y = y(p, z, \epsilon_d), p \in [\underline{p}(z), \overline{p}(z)]\}$ be nonempty for every $z \in \mathcal{Z}$, where $y(\cdot, \cdot, \cdot)$ is defined in (37). The cost of public funds $\lambda(\cdot)$ and the joint distribution of (Y, C, P, T) given Z satisfy*

(i) $\mathrm{E}[Y|p=p,Z=z]>0$, $\mathrm{E}[C|P=p,Z=z]>0$,
(ii) $\Gamma(p,z)>0$, $\partial\Gamma(p,z)/\partial p<0$,
(iii) $\partial\Delta(p,z)/\partial p>0$ *and* $\Gamma(p,z)<\overline{c}_o(p,z)$, *where* $\overline{c}_o(p,z)$ *is given by (25)-(26),*
for any $p\in[\underline{p}(z),\overline{p}(z)]$ *and* $z\in\mathcal{Z}$. *Moreover,*
(iv) the random variables $\phi_d(Y,P,Z)$ *and* $\phi_c(Y,C,P,Z)$ *defined in (41) and (42), respectively, are conditionally independent of* P *given* Z, *while* $\mathrm{E}[\phi_t(Y,C,P,T,Z)|P,Z]=0$,
(v) the conditional distribution $G_{P|Z}(\cdot|\cdot)$ *has a strictly positive density on its support* $\{(p,z):p\in[\underline{p}(z),\overline{p}(z)],z\in\mathcal{Z}\}$ *with* $\underline{p}(\cdot)<\overline{p}(\cdot)$, *while the conditional distributions* $G_{Y|P,Z}(\cdot|\cdot,\cdot)$ *and* $G_{C|Y,P,Z}(\cdot|\cdot,\cdot,\cdot)$ *are nondegenerated and strictly increasing in their first arguments.*

Note that $\mu(\cdot)=\lambda(\cdot)/(1+\lambda(\cdot))$, while $\Delta(p,z)$ and $\Gamma(p,z)$ are defined in (27) and (28), respectively. Assumption C1-(i) is implied by A1. Assumption C1-(ii) ensures that the expected transfer is strictly decreasing in the firm's cost inefficiency $\theta-e^*(\theta,z)$ as required by Proposition 3. Assumption C1-(iii) is actually equivalent to $\partial\Delta(p,z)/\partial p>R(p,z)>0$ in Proposition 4 which ensures a strictly decreasing effort function and a strictly increasing price schedule, respectively as required by Lemma 2. Specifically, from (29), $R(p,z)>0$ is equivalent to having its denominator negative, i.e. $\Gamma(p,z)<\overline{c}_o(p,z)-\mu(z)(G_{P|Z}(p|z)/g_{P|Z}(p|z))\times(\partial\Gamma(p,z)/\partial p)$ in view of $\partial\Gamma(p,z)/\partial p<0$ by C1-(ii) and $\partial\Delta(p,z)/\partial p>0$. On the other hand, $\partial\Delta(p,z)/\partial p>R(p)$ is equivalent to $\Gamma(p,z)<\overline{c}_o(p,z)$ after some elementary algebra. Hence the combination of the two inequalities holds if and only if $\partial\Delta(p,z)/\partial p>0$ and $\Gamma(p,z)<\overline{c}_o(p,z)$. Assumption C1-(iv) is a direct consequence of B1 combined with $p(\theta,z)$ strictly increasing in θ. Lastly, the first part of C1-(v) follows those on $F(\cdot|\cdot)$ in A1, while the second part of C1-(v) parallels the second part of B3-(ii).

We define the set of structures $\mathcal{S}\equiv\{S=[y,c_o,\psi,F,G,\lambda]:\text{A1, B1}-\text{B3, hold}\}$. Thus A2-A3 need not be satisfied by structures in $\mathcal{S}$. The next lemma shows that C1 provides necessary and sufficient conditions for the properties of Lemmas 1-2 and Proposition 3 to hold. Namely,

P1: The firm's objective function is strictly concave in effort so that there is a unique solution to the firm's effort maximization problem,

P2: The optimal effort is strictly decreasing in θ so that the local SOC $e'(\theta)\leq e_2(\theta,\theta)$ is satisfied,

P3: The optimal price schedule is strictly increasing in θ,

P4: Truth telling provides the global maximum of the firm's problem,

P5: The expected transfer is strictly decreasing in the firm's cost inefficiency.

Hence, A2-A3 are sufficient but not necessary for such properties. A conditional distribution for (Y, C, P, T) given Z is *induced* by a structure $S \in \mathcal{S}$ if (Y, C, P, T, Z) satisfies (17)–(21) for some effort $e(\theta, Z)$.

Lemma 7: *If $S \in \mathcal{S}$ satisfies P1–P5, then the conditional distribution of (Y, C, P, T) given Z induced by S satisfies C1. Conversely, if the cost of public fund $\lambda(\cdot)$ and the conditional distribution of (Y, C, P, T) given Z satisfy C1, then there exists a structure in $\mathcal{S}$ satisfying P1–P5 and rationalizing the observations (Y, C, P, T) given Z.*

The first part of Lemma 7 shows that A2-A3 are stronger than C1 because any structure in $\mathcal{S}$ satisfying A2-A3 must also satisfy P1–P5 by Lemmas 1-2 and Proposition 3. This result is not surprising as C1 provide parsimonious conditions for the observables relying on the first-order and second-order conditions and the implementation. The two parts of Lemma 7 show that C1 characterizes all the restrictions imposed by the model on observables given $\lambda(\cdot)$. Such a characterization is crucial in the structural approach as it allows the analyst to test the validity of the model. In particular, violation of any of these restrictions would reject the model for explaining the data. The restrictions in C1 are relatively weak as they do not allow one to identify $\lambda(\cdot)$ as shown next.

The next proposition shows that the cost of public funds is not identified. The proof relies on constructing a structure that is observationally equivalent to the original one generating the observations (Y, C, P, T).

Proposition 6: *In the model consisting of structures $S \in \mathcal{S}$ inducing conditional probability distributions for (Y, C, P, T) given Z that satisfy C1, the cost of public funds $\lambda(\cdot)$ is not identified.*

This result is a consequence of Lemma 7. As shown in the proof, it suffices to find another cost of public funds $\tilde{\lambda}(\cdot)$ in a structure $\tilde{S}$ that is observationally equivalent to the original structure S generating the observables (Y, C, P, T) and that satisfies C1. In particular, it is shown that C1-(iv) is satisfied by construction.

Given that the cost of public funds is not identified, we consider some identifying conditions or assumptions in the next subsection. Another identification strategy would be to consider set identification, i.e. the restrictions on $\lambda(\cdot)$ embodied in C1, and to derive some bounds for $\lambda(\cdot)$ given the distribution of (Y, C, P, T, Z) in the spirit of Manski and Tamer (2002) and Chernozhukov, Hong and Tamer (2007).

4.2 Identification of $\lambda(\cdot)$

Proposition 6 indicates that we need to exploit additional information to identify $\lambda(\cdot)$ as all the restrictions imposed by the model summarized in C1 have been exploited. On the other hand, up to now we have imposed a weak assumption on ϵ_t, namely $\mathrm{E}[\epsilon_t|P, Z] = 0$. Imposing additional assumptions on ϵ_t seems a natural way to achieve identification of $\lambda(\cdot)$. This subsection discusses some possibilities and retains an assumption that will lead to a simple expression of $\lambda(\cdot)$ in terms of observables.

In view of B1, which assumes that the other error terms (ϵ_d, ϵ_c) are conditionally independent of θ, a first strategy is to strengthen the mean independence of ϵ_t from θ given Z by assuming that ϵ_t is independent of θ given Z, i.e. the residual transfer is independent of the firm's efficiency given Z. In this case, we have $\epsilon_t = \phi_t(Y, C, P, T, Z; \mu(\cdot))$ and $\theta = \theta(P, Z; \mu(\cdot))$, where $\phi_t(\cdot, \cdot, \cdot, \cdot, \cdot; \mu(\cdot))$ and $\theta(\cdot, \cdot; \mu(\cdot))$ are identified functions from (43) and (32), respectively. Thus, the conditional independence of $\phi_t(Y, C, P, T, Z; \mu(\cdot))$ and $\theta(P, Z; \mu(\cdot))$ given Z leads to some conditional moment restrictions such as their conditional covariance equal to zero. These restrictions can then be used to identify $\mu(\cdot)$ and hence $\lambda(\cdot)$. Given the nonlinearity of (43) and (32) in $\mu(\cdot)$, it is difficult to establish the identification of $\mu(\cdot)$ from such conditional moment restrictions. Moreover, if it was identified, there would be no closed form solution for $\mu(\cdot)$.

A second strategy, which is more standard in econometrics, is to achieve identification through the use of an additional variable (instrument) W that is conditionally independent of ϵ_t given (θ, Z) leading to the conditional moment restriction $\mathrm{E}[W\epsilon_t|P, Z] = 0$ since $\mathrm{E}[\epsilon_t|P, Z] = 0$ by B1. In this case, we have

$$\mathrm{E}[W(T - \mathrm{E}(T|P, Z))|P, Z] + K(P, Z; \mu)\mathrm{E}[W(C - \mathrm{E}(C|P, Z))|P, Z] = 0,$$

where $K(P, Z; \mu)$ is the coefficient of $C - \mathrm{E}(C|P, Z)$ in (43). To obtain a solution for μ, we must have $\mathrm{E}[W(C - \mathrm{E}(C|P, Z))|P, Z] \neq 0$, i.e. the instrument W must be correlated with C given (P, Z). Solving for μ then gives

$$\mu(z) = \frac{1}{\overline{y}(p,z)} \left[\frac{\partial \mathrm{E}[T|P=p,Z=z]}{\partial p} \frac{\mathrm{Cov}(W,C|P=p,Z=z)}{\mathrm{Cov}(W,T|P=p,Z=z)} + \frac{\partial \mathrm{E}[T|P=p,Z=z]}{\partial p} - p\overline{y}'(p) \right], \tag{44}$$

for any $p \in [\underline{p}(z), \overline{p}(z)]$ and $z \in \mathcal{Z}$, where $\overline{y}(p) = \mathrm{E}[Y|P = p, Z = z]$ provided $\mathrm{Cov}(W, T|P, Z) \neq 0$. Thus $\mu(\cdot)$ is identified. To summarize, a valid instrument W should be uncorrelated with ϵ_t but correlated with the transfer T and the cost C given (P, Z). In particular, this excludes any political variable which is likely to be correlated with ϵ_t. It also excludes P or any variable in Z as these variables are uncorrelated with T and C given (P, Z).

A natural candidate for a valid instrument can be the cost as C is correlated with itself and T provided that it is independent of ϵ_t. In Section 3.1, ϵ_t is interpreted as the residual transfer due to the regulator's discretion and political/legal constraints faced by the regulator. It seems reasonable to assume that the residual transfer is unrelated to cost. Thus we make the following identifying assumption.

Assumption C2: $\mathrm{E}[C\epsilon_t|\theta, Z] = 0$.

Because $\mathrm{E}[\epsilon_t|\theta, Z] = 0$ by B1, C2 requires that C and ϵ_t are uncorrelated given (θ, Z) or equivalently given (P, Z). In particular, because $P = p^*(\theta, Z)$ and $e = e^*(\theta, Z)$, it follows from (17)-(18) that C2 is satisfied if (ϵ_d, ϵ_c) is independent of ϵ_t given (θ, Z), i.e. the demand and cost shocks are independent of the residual transfer given (P, Z).

The next proposition establishes the nonparametric identification of the cost of public funds $\lambda(\cdot)$ from observations (Y, C, P, T, Z).

Proposition 7: *Under A1, B1–B3 and C1-C2, the cost of public funds $\lambda(\cdot)$ is uniquely determined by $\lambda(z) = \mu(z)/[1 - \mu(z)]$, where*

$$\begin{aligned} \mu(z) \quad = \quad & \frac{1}{\mathrm{E}[Y|P=p,Z=z]} \left\{ \frac{\partial \mathrm{E}[T|P=p,Z=z]}{\partial p} \frac{\mathrm{Var}[C|P=p,Z=z]}{\mathrm{Cov}[C,T|P=p,Z=z]} \right. \\ & \left. + \frac{\partial \mathrm{E}[C|P=p,Z=z]}{\partial p} - p \frac{\partial \mathrm{E}[Y|P=p,Z=z]}{\partial p} \right\} \end{aligned} \tag{45}$$

with $\mathrm{Cov}[C, T|P=p, Z=z] < 0$, *for every* $p \in [\underline{p}(z), \overline{p}(z)]$ *and* $z \in \mathcal{Z}$.

Because p can be chosen arbitrarily, (45) shows that $\mu(\cdot)$ and hence $\lambda(\cdot)$ are overidentified. Thus, weaker assumptions than C2 can be exploited to achieve identification of the cost of public funds. For instance, we could assume that C2 holds for the most efficient firm only. In this case, (45) holds at $p = \underline{p}(z)$ only.

4.3 Regulator's Objectives

As discussed in Sections 2 and 3.1, the regulator may not be the utilitarian benevolent maximizer of the Laffont and Tirole's (1986) model and may face various political and legal constraints that affect the regulatory contract, i.e. the price and the transfer. For instance, the regulator may not weight equally the consumers' interests and the firm's profit given that some commissioners are elected or captured by some interest groups. Moreover, in addition to the residual transfer ϵ_t, there may be some deterministic deviation from the optimal transfer that can be explained by some regulator's characteristics or considerations. These aspects can be formally introduced in the model through a weight $\alpha(\cdot)$ on the consumer surplus in the social welfare (8) and a function $m(\cdot) \geq 0$ added to the transfer (21). These two additional functions in the model could respond to some critics of the normative approach taken by Laffont and Tirole (1986). In particular, while acknowledging the rich set of regulation models based on asymmetric information, Joskow and Schmalensee (1986) and Joskow (2005) emphasize the role of political and legal factors as well as the importance of investment and the dynamic aspects of regulation. Our extended model attempts to take into account the former. We maintain the same assumptions as before, namely A1, B1–B3 and C1-C2.

The introduction of a weight $\alpha(\cdot)$ and a deterministic transfer $m(\cdot)$ affects the Ramsey pricing (19) by replacing μ with $\nu = (\lambda + 1 - \alpha)/(1 + \lambda)$ and the transfer (21) by adding $m(\cdot)$ to the right-hand side. The rest of the model remains the same as $m(\cdot)$ modifies the participation constraint (10), namely $U(\theta, Z) \geq m(Z)$, in which case $m(Z)$ should be added to the constant term (16) of the payment. Following Sections 3.2 and 3.3, the identification of the extended model is studied first given $(\lambda(\cdot), \alpha(\cdot), m(\cdot))$. Lemma 4 still holds with $\mu(\cdot)$ replaced by $\nu(\cdot)$ in (26). Thus, $\overline{y}(\cdot, \cdot)$ remains the same while $\overline{c}_0(\cdot, \cdot)$ and $\Delta(\cdot, \cdot)$ depend on $\nu(\cdot)$. Note that (28) and (29) still hold though $\Delta(\cdot, \cdot)$ and $\overline{c}_o(\cdot, \cdot)$ now depend on $\nu(\cdot)$. Proposition 4 then holds with $m(\cdot)$ substracted from the right-hand side of (30) which gives the effort disutility $\psi(\cdot, \cdot)$. Hence, $\overline{y}(\cdot, \cdot)$, $\overline{c}_o(\cdot, \cdot)$, $\psi(\cdot, \cdot)$, $F(\cdot|\cdot)$ are identified given $(\lambda(\cdot), \alpha(\cdot), m(\cdot))$. Consequently, the results of Section 3.3 apply thereby establishing the identification of $y(\cdot, \cdot, \cdot)$, $c_o(\cdot, \cdot, \cdot)$ and $G(\cdot, \cdot, \cdot|\cdot)$ given $(\lambda(\cdot), \alpha(\cdot), m(\cdot))$.

We now turn to the identification of $(\lambda(\cdot), \alpha(\cdot), m(\cdot))$ or equivalently $(\mu(\cdot), \nu(\cdot), m(\cdot))$. We note that (43) remains valid with μ replaced by ν. Thus, Proposition 8 applies thereby identifying $\nu(\cdot)$. Once $\nu(\cdot)$ is identified, the demand $y(\cdot,\cdot,\cdot)$ and the baseline cost $c_o(\cdot,\cdot,\cdot)$ are identified from (37) and (39), respectively. To identify $\mu(\cdot)$ and $m(\cdot)$, we use exclusion restrictions following Guerre, Perrigne and Vuong (2008). Specifically, we partition Z into (Z_1, Z_2, Z_3) with Z_1 affecting only $m(\cdot)$, Z_2 affecting only $\mu(\cdot)$ or equivalently $\lambda(\cdot)$ and Z_3 affecting only $F(\cdot|\cdot)$ and $\psi(\cdot,\cdot)$. For instance, Z_1 could capture some regulator's characteristics or environment such as the regulator's and commissioners' background, and some political and legal variables. The variables Z_2 can represent some local economic conditions that affect the cost of public funds, while the variables Z_3 can be any other exogenous variables. To identify $\mu(\cdot)$, we can use (34), which gives $\underline{\theta}(z) = \underline{\theta}(z_3)$ by the exclusion restriction $F(\cdot|z) = F(\cdot|z_3)$. In particular, the left-hand side $\underline{\theta}(\cdot)$ does not depend on z_2, while the right-hand side depends on $\mu(z_2)$. Taking the derivative of (34) with respect to z_2 can identify $\mu(\cdot)$. Once $\mu(\cdot)$ is identified, $F(\cdot|\cdot)$ is identified from (32). To identify $m(\cdot)$, we note that (30) identifies $\psi(\cdot,\cdot) + m(\cdot)$. Since the exclusion restrictions impose $\psi(\cdot, z_3)$ and $m(z_1)$, $m(\cdot)$ and $\psi(\cdot,\cdot)$ can be identified. Once $m(\cdot)$ is identified, $G(\cdot,\cdot,\cdot|\cdot)$ is identified.

5 Conclusion

This paper studies the nonparametric identification of contract models under asymmetric information and more specifically the Laffont and Tirole's (1986) incentive regulation model with expost observed cost, moral hazard and adverse selection. Exploiting the bijective mapping between the observed price and the firm's unknown type, we show that at a given cost of public funds we can recover the structure of the model, namely the demand and baseline cost, the effort disutility, the firms' type distribution and the joint conditional distribution of the random shocks. We then identify the cost of public funds under a conditional mean independence of cost and residual transfer. We also characterize the restrictions that must be satisfied by the cost of public funds and the observables so that the latter can be rationalized by the model.

Our paper represents a stepping stone in the structural analysis of data subject to incomplete information such as in contracts, insurance and nonlinear pricing. Our results indicate that one does not have to rely on parametric functional forms to identify and estimate such models as our analysis is sufficiently general to extend to other models of incomplete information. As a matter of fact, the incentive regulation model we consider includes some functions such as the effort disutility and the cost of public funds that do not appear in simpler models of incomplete information under adverse selection. This is the case for nonlinear pricing models. See Huang, Perrigne and Vuong (2008) for the nonparametric identification and estimation of a nonlinear pricing model with an application to yellow pages. See also D'Haultfoeuille and Février (2007) for an application to principal-agent models.

Clearly, the problem of estimating and testing such models needs to be addressed. It includes two important questions. First, we need to derive the restrictions imposed by the model on observables to test its validity. Lemma 7 provides a first step toward this goal. Second, incomplete information is generally assumed. It would be interesting to assess the relevance of such an hypothesis. The restrictions imposed by a complete information model would allow one to test which model (incomplete or complete information) is the most accurate to explain the data. The problem of testing adverse selection has known a vivid interest recently and some tests have been developed within the reduced form approach. See Chiappori and Salanié (2000). A test based on the theoretical restrictions imposed by the model on observables would provide a complete answer to this question within the structural approach.

Lastly, it remains to develop a suitable nonparametric estimation procedure and to study its asymptotic properties. Our results show how to express the structural elements of the model from the reduced form probability distribution of the observables through various conditional expectations. The data would consist of several observations of (Y, C, P, T, Z). Thus a multistep estimation procedure could be entertained. A first step would consist in estimating the cost of public funds using standard nonparametric regression estimators on (45) involving the observables (Y, C, P, T, Z). The demand is then estimated by a nonparametric quantile regression of Y on (P, Z) given an arbitrary choice of $p_o(\cdot)$ using (37). A second step would consist in estimating the expected baseline cost from observations on (Y, C, P, Z) using (26). The difficulty here is that a conditional expectation needs to be estimated at the lower boundary of the price support, while this lower boundary also appears in the integral. Nonparametric estimation at the boundary is known to introduce some boundary effects that can be alleviated by local polynomial estimators. With an estimate for the expected baseline cost at hand, we then obtain estimates of the relative cost inefficiency by (27) and the two functions (28) and (29) that are needed in the estimation of the effort disutility (30) and the firm's type (32). This step relies on observations (P, T, Z) and standard nonparametric estimators for the conditional density and distribution of price. Here again, boundary effects have to be controlled as (30) and (32) involve nonparametric estimators at the upper boundary of the price support. A third step would consist in estimating the conditional type density using the estimated types obtained previously following Guerre, Perrigne and Vuong (2000). The baseline cost is estimated by a nonparametric quantile regression of the recovered baseline cost value C_o on (Y, P, Z) given an arbitrary choice of $y_o(\cdot)$ using (39). Lastly, the joint conditional distribution of error terms is obtained from the estimated error terms in (41), (42) and (43). In addition to some boundary problems, the major difficulty in establishing the asymptotic properties of this estimation procedure arises from its multi-step nature and the fact that some functions are estimated from recovered values instead of observed ones. Alternatively, given that nonparametric identification has been established, the analyst can adopt a fully parametric or semiparametric econometric specification. In such case, data on transfer may not be needed as in Perrigne (2002).

Appendix A

This appendix gives the proofs of the propositions and lemmas stated in Section 2.

Proof of Proposition 1: From (8) the Hamiltonian of the optimization problem (P') is

$$\begin{aligned}\mathcal{H} &= \Big\{ \int_p^\infty \overline{y}(v)dv + (1+\lambda)\Big(p\overline{y}(p) - \psi(e) - \mathrm{E}\Big[(\theta - e)c_o(y(p,\epsilon_d),\epsilon_c)\Big]\Big) \\ &\quad -\lambda U(\theta)\Big\} f(\theta) + \gamma(\theta)(-\psi'(e)),\end{aligned}$$

where $p = p(\theta)$ and $e = e(\theta)$ are the control functions, $U(\theta)$ is the state variable, and $\gamma(\theta)$ is the co-state variable. Hence, applying the Pontryagin principle, the FOC are

$$\begin{aligned}\mathcal{H}_p &= \Big\{\lambda\overline{y}(p) + (1+\lambda)p\overline{y}'(p) - (1+\lambda)\mathrm{E}\Big[(\theta-e)c_{o1}(y(p,\epsilon_d),\epsilon_c)y_1(p,\epsilon_d)\Big]\Big\} f(\theta) = 0 \\ \mathcal{H}_e &= \Big\{ -(1+\lambda)\psi'(e) + (1+\lambda)\mathrm{E}\Big[c_o(y(p,\epsilon_d),\epsilon_c)\Big]\Big\} f(\theta) - \gamma(\theta)\psi''(e) = 0 \\ -\mathcal{H}_U &= \lambda f(\theta) = \gamma'(\theta).\end{aligned}$$

The last equation gives $\gamma(\theta) = \lambda F(\theta)$ using the transversality condition $\gamma(\underline{\theta}) = 0$. Thus, rearranging $\mathcal{H}_p$ and $\mathcal{H}_e$, the solutions $p = p^*(\theta)$ and $e = e^*(\theta)$ are given by (12) and (13).□

Proof of Proposition 2: Given the price schedule $p^*(\cdot)$ and the transfer function $t^*(\cdot,\cdot)$, we show that the firm will announce its true type θ and exerts the optimal effort $e^*(\theta)$ by verifying the FOC of the firm's problem (F). Under A1, this problem becomes

$$\begin{aligned}(F^*) \qquad \max_{\tilde{\theta},e}\ & \mathrm{E}\left[t^*\Big(\tilde{\theta},(\theta-e)c_o(y(p^*(\tilde{\theta}),\epsilon_d),\epsilon_c)\Big) \mid \theta\right] - \psi(e) \\ &= \mathrm{E}\left[t^*\Big(\tilde{\theta},(\theta-e)c_o(y(p^*(\tilde{\theta}),\epsilon_d),\epsilon_c)\Big)\right] - \psi(e) \\ &= A(\tilde{\theta}) + \psi'[e^*(\tilde{\theta})]\Big\{\tilde{\theta} - e^*(\tilde{\theta}) - (\theta - e)\Big\} - \psi(e),\end{aligned}$$

where the first equality follows from the independence between θ and (ϵ_d, ϵ_c), while the second equality follows from (15). Thus, using (16) the FOC with respect to $\tilde{\theta}$ and e are

$$\begin{aligned}0 &= \psi'[e^*(\tilde{\theta})]e^{*\prime}(\tilde{\theta}) - \psi'[e^*(\tilde{\theta})] + \frac{d\psi'[e^*(\tilde{\theta})]}{d\tilde{\theta}}\Big\{\tilde{\theta} - e^*(\tilde{\theta}) - (\theta-e)\Big\} + \psi'[e^*(\tilde{\theta})]\Big[1 - e^{*\prime}(\tilde{\theta})\Big] \\ &= \frac{d\psi'[e^*(\tilde{\theta})]}{d\tilde{\theta}}\Big\{\tilde{\theta} - e^*(\tilde{\theta}) - (\theta-e)\Big\} \\ 0 &= \psi'[e^*(\tilde{\theta})] - \psi'(e).\end{aligned}$$

It is easy to see that these FOC are verified if $\tilde{\theta} = \theta$ and $e = e^*(\theta)$.

It remains to show that $[p^*(\cdot), t^*(\cdot,\cdot), e^*(\cdot), U^*(\cdot)]$ solves the FOC of problem (P). In view of the discussion surrounding problem (P'), it suffices to show that the transfer function $t^*(\cdot,\cdot)$ satisfies (6) and (7) where $[p^*(\cdot), e^*(\cdot), U^*(\cdot)]$ solves the FOC of problem (P'). The preceding shows that the transfer function $t^*(\cdot,\cdot)$ satisfies (7). It remains to show that $t^*(\cdot,\cdot)$ also satisfies (6). Using (15), the right-hand side of (6) is

$$A(\theta) + \psi'[e^*(\theta)]\{\theta - e^*(\theta) - (\theta - e^*(\theta))\} - \psi[e^*(\theta)] = A(\theta) - \psi[e^*(\theta)] = U^*(\theta)$$

by (14) and (16), as desired.□

Proof of Lemma 1: From the problem (F), the second partial derivative of the firm's objective function with respect to e is

$$\int U_{33}(\tilde{\theta}, \theta, e, \epsilon_d, \epsilon_c) dG(\epsilon_d, \epsilon_c) \quad = \quad \int t_{22}(\cdot) c_o^2(\cdot) dG(\epsilon_d, \epsilon_c) - \psi''(e),$$

where we have omitted the arguments of the functions to simplify the notation. When the transfer function $t(\cdot,\cdot)$ is weakly decreasing and concave in realized cost c so that $t_2(\cdot) \leq 0$ and $t_{22}(\cdot) \leq 0$, it follows from $\psi''(\cdot) > 0$ that the firm's objective function is *strictly* concave in e for any $(\tilde{\theta}, \theta)$. Hence, the effort $e(\tilde{\theta}, \theta)$, which solves the FOC (3), is uniquely defined and corresponds to a global maximum of the problem (FE).

Next, we show that $0 \leq e_2(\theta, \theta) < 1$. This can be seen by differentiating the FOC (3) defining $e(\tilde{\theta}, \theta)$ with respect to θ. This gives

$$0 = [1 - e_2(\tilde{\theta}, \theta)]\mathrm{E}\Big[t_{22}(\cdot) c_o^2(\cdot)\Big] + \psi''[e(\tilde{\theta}, \theta)] e_2(\tilde{\theta}, \theta).$$

Rearranging and evaluating at $\tilde{\theta} = \theta$ give

$$e_2(\theta, \theta)\Big\{\mathrm{E}\Big[t_{22}(\cdot) c_o^2(\cdot)\Big] - \psi''[e(\theta)]\Big\} = \mathrm{E}\Big[t_{22}(\cdot) c_o^2(\cdot)\Big].$$

Thus the expectation term is nonpositive whenever the transfer function $t(\cdot,\cdot)$ is weakly decreasing and concave in realized cost. Because $\psi''(\cdot) > 0$ by A2-(iii), it follows that $0 \leq e_2(\theta, \theta) < 1$.□

Proof of Lemma 2: As noted before A3, the local SOC (19) is satisfied as soon as $e^{*\prime}(\cdot) \leq 0$. We show that $e^{*\prime}(\cdot) < 0$. By definition $[p^*(\cdot), e^*(\cdot)]$ satisfies the FOC (12)-(13), which can be written as

$$p^*(\theta)\overline{y}'[p^*(\theta)] \quad = \quad (\theta - e^*(\theta))\overline{c}_o'[p^*(\theta)] - \mu \overline{y}[p^*(\theta)] \tag{A.1}$$

$$\psi'[e^*(\theta)] \quad = \quad \overline{c}_o[p^*(\theta)] - \mu \frac{F(\theta)}{f(\theta)} \psi''[e^*(\theta)], \tag{A.2}$$

where we have used A1, the definition of $\overline{c}_o(\cdot)$, and the expression for $\overline{c}_o'(\cdot)$ found earlier. Differentiating (12)-(13) with respect to θ and rearranging equations give

$$\begin{aligned} Ae^{*\prime}(\theta) + Bp^{*\prime}(\theta) &= A && \text{(A.3)} \\ Ce^{*\prime}(\theta) - Ap^{*\prime}(\theta) &= D, && \text{(A.4)} \end{aligned}$$

where

$$\begin{aligned} A &= \overline{c}_o'[p^*(\theta)] \\ B &= (1+\mu)\overline{y}'[p^*(\theta)] + p^*(\theta)\overline{y}''[p^*(\theta)] - (\theta - e^*(\theta))\overline{c}_o''[p^*(\theta)] \\ &= (1-\mu)\overline{V}''[p^*(\theta)] - (\theta - e^*(\theta))\overline{c}_o''[p^*(\theta)] \\ C &= \psi''[e^*(\theta)] + \mu\frac{F(\theta)}{f(\theta)}\psi'''[e^*(\theta)] \\ D &= -\mu\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\psi''[e^*(\theta)], \end{aligned}$$

with $\mu = \lambda/(1+\lambda)$. Under A1–A2, note that $A < 0$, $B < 0$ and $C > 0$. Solving for $e^{*\prime}(\theta)$ gives

$$e^{*\prime}(\theta)\left(C + \frac{A^2}{B}\right) = D + \frac{A^2}{B}.$$

Thus, $e^{*\prime}(\cdot) < 0$ if $-C < A^2/B < -D$, i.e. if

$$\begin{aligned} -\left(\psi''[e^*(\theta)] + \mu\frac{F(\theta)}{f(\theta)}\psi'''[e^*(\theta)]\right) &< \frac{\left\{\overline{c}_o'[p^*(\theta)]\right\}^2}{(1-\mu)\overline{V}''[p^*(\theta)] - (\theta - e^*(\theta))\overline{c}_o''[p^*(\theta)]} \\ &< \mu\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\psi''[e^*(\theta)]. && \text{(A.5)} \end{aligned}$$

Because $-B \geq -(1-\mu)\overline{V}''[p^*(\theta)] > 0$, A3-(i) ensures that

$$-\psi''[e^*(\theta)] < \frac{\left\{\overline{c}_o'[p^*(\theta)]\right\}^2}{(1-\mu)\overline{V}''[p^*(\theta)] - (\theta - e^*(\theta))\overline{c}_o''[p^*(\theta)]},$$

which implies the first inequality in (A.5) by A2. By A2-(iii) and A3-(ii), we have $D < 0$, while $B < 0$ thereby implying the second inequality in (A.5). Lastly, because $e^{*\prime}(\theta) + p^{*\prime}(\theta)B/A = 1$ by (A.3) with $A < 0$ and $B < 0$, it follows from $e^{*\prime}(\cdot) < 0$ that $p^{*\prime}(\cdot) > 0$, as desired.□

Proof of Proposition 3: Recalling that $e(\tilde{\theta}, \theta)$ is the optimal level of effort for a firm with type θ, the firm's expected utility (4) from announcing $\tilde{\theta}$ is

$$U(\tilde{\theta}, \theta) = A(\tilde{\theta}) + \psi'[e^*(\tilde{\theta})]\left\{\tilde{\theta} - e^*(\tilde{\theta}) - (\theta - e(\tilde{\theta}, \theta))\right\} - \psi[e(\tilde{\theta}, \theta)]$$

(see the optimization problem (F^*) in the proof of Proposition 2). To show that $\tilde{\theta} = \theta$ provides a global maximum, we first show that $U_{12}(\tilde{\theta}, \theta) > 0$ for any $(\tilde{\theta}, \theta)$. Using $U_{12}(\tilde{\theta}, \theta) = -\psi''[e(\tilde{\theta}, \theta)]e_1(\tilde{\theta}, \theta)$, this is equivalent to showing $e_1(\tilde{\theta}, \theta) < 0$, where $e(\tilde{\theta}, \theta)$ solves the FOC (3), which can be written under A1 as $0 = \psi'[e^*(\tilde{\theta})] - \psi'[e(\tilde{\theta}, \theta)]$ from the FOC of problem (F^*). Differentiating this FOC with respect to $\tilde{\theta}$ gives $e_1(\tilde{\theta}, \theta)\psi''(\cdot) = \psi''(\cdot)e^{*\prime}(\cdot)$. Because $e^{*\prime}(\cdot) < 0$ by Lemma 2, the right-hand side is strictly negative under A2. Hence $e_1(\tilde{\theta}, \theta) < 0$ implying $U_{12}(\cdot, \cdot) > 0$, as desired. Second, we apply the argument in Appendix A1.4 in Laffont and Tirole (1993) with $\phi(\beta, \hat{\beta})$ equal to $U(\tilde{\theta}, \theta)$. Hence, $\tilde{\theta} = \theta$ provides the global maximum of $U(\tilde{\theta}, \theta)$.

To prove the second part, let $\bar{t}(\theta) \equiv \mathrm{E}\Big[t^*\Big(\theta, (\theta - e^*(\theta))c_o(y(p^*(\theta), \epsilon_d), \epsilon_c)\Big)\Big]$ so that $\bar{t} = \bar{t}(\theta)$. Note that $\mathcal{E}(\theta) \equiv \theta - e^*(\theta)$ is strictly increasing in θ because $d(\theta - e^*(\theta))/d\theta = [1 - e^{*\prime}(\theta)] > 0$ and $e^{*\prime}(\cdot) < 0$. Thus $\theta = \mathcal{E}^{-1}(\mathcal{E})$, where $\mathcal{E}$ is the firm's cost inefficiency. We want to show that $\bar{t}(\theta) = \bar{t}[\mathcal{E}^{-1}(\mathcal{E})]$ is strictly decreasing in $\mathcal{E}$. From (15) and A1, we have $\bar{t}(\theta) = A(\theta)$. Hence, using (16)

$$\frac{d\bar{t}}{d\mathcal{E}} = \frac{A'(\theta)}{\mathcal{E}'(\theta)} = -\frac{\psi'[e^*(\theta)]}{1 - e^{*\prime}(\theta)]} < 0.$$

Thus, the expected transfer is strictly decreasing in $\mathcal{E}$, as desired.□

Appendix B

This appendix gives the proofs of the propositions and lemmas stated in Sections 3 and 4.

Proof of Lemma 3: Let $\tilde{Y}, \tilde{C}, \tilde{P}, \tilde{T}$ denote the endogenous variables under the structure $\tilde{S}$. Let $\tilde{\theta} \equiv \alpha + \beta\theta$ so that $\tilde{\theta}$ is distributed as $\tilde{F}(\cdot|\cdot) = F[(\cdot - \alpha)/\beta|\cdot]$ conditional upon Z. Let $(\tilde{\epsilon_d}, \tilde{\epsilon_c}, \tilde{\epsilon_t}) \equiv (\epsilon_d, \epsilon_c, \epsilon_t)$ so that $(\tilde{\epsilon_d}, \tilde{\epsilon_c}, \tilde{\epsilon_t})$ is jointly distributed as $\tilde{G}(\cdot, \cdot, \cdot|\cdot) = G(\cdot, \cdot, \cdot|\cdot)$ conditional upon Z. We show that $(\tilde{Y}, \tilde{C}, \tilde{P}, \tilde{T}) = (Y, C, P, T)$, which implies the desired result.

Using $\tilde{y}(\cdot, \cdot, \cdot) = y(\cdot, \cdot, \cdot)$, $\tilde{c}_o(\cdot, \cdot, \cdot) = c_o(\cdot, \cdot, \cdot)/\beta$, (22) and (23), we note that

$$\bar{\tilde{y}}(\cdot, \cdot) = \int \tilde{y}(\cdot, \cdot, \tilde{\epsilon}_d)d\tilde{G}(\tilde{\epsilon}_d|\cdot) = \int y(\cdot, \cdot, \epsilon_d)dG(\epsilon_d|\cdot) = \bar{y}(\cdot, \cdot), \tag{B.1}$$

$$\bar{\tilde{c}}_o(\cdot, \cdot) = \int \tilde{c}_o[\tilde{y}(\cdot, \cdot, \tilde{\epsilon}_d), \cdot, \tilde{\epsilon}_c]d\tilde{G}(\tilde{\epsilon}_d, \tilde{\epsilon}_c|\cdot) = \frac{1}{\beta}\int c_o[y(\cdot, \cdot, \epsilon_d), \cdot, \epsilon_c]dG(\epsilon_d, \epsilon_c|\cdot) = \frac{1}{\beta}\bar{c}_o(\cdot, \cdot). \tag{B.2}$$

Now, we consider the FOC (19)–(20) for $(\tilde{P}, \tilde{e})$ and use (B.1)–(B.2), $\tilde{\psi}(\cdot, \cdot) = \psi[(\cdot - \alpha)/\beta, \cdot]$, $\tilde{\theta} = \alpha + \beta\theta$, $\tilde{F}(\cdot|\cdot) = F[(\cdot - \alpha)/\beta|\cdot]$, $\tilde{f}(\cdot|\cdot) = (1/\beta)f[(\cdot - \alpha)/\beta|\cdot]$ and $\tilde{\lambda}(\cdot) = \lambda(\cdot)$ to obtain

$$\begin{aligned} \tilde{P}\bar{y}'(\tilde{P}, Z) + \mu\bar{y}(\tilde{P}, Z) &= \left(\theta - \frac{\tilde{e} - \alpha}{\beta}\right)\bar{c}_o'(\tilde{P}, Z) \\ \psi'\left(\frac{\tilde{e} - \alpha}{\beta}, Z\right) + \mu\frac{F(\theta|Z)}{f(\theta|Z)}\psi''\left(\frac{\tilde{e} - \alpha}{\beta}, Z\right) &= \bar{c}_o(\tilde{P}, Z). \end{aligned}$$

From the solution $p^*(\theta,z)$ and $e^*(\theta,z)$ of (19)–(20), it follows that $\tilde{P} = p^*(\theta,Z) = P$ and $(\tilde{e}-\alpha)/\beta = e^*(\theta,Z) = e$. In particular, the latter implies that $\tilde{e}^*(\tilde{\theta},Z) = \alpha + \beta e^*[(\tilde{\theta}-\alpha)/\beta, Z]$. Moreover, because $\tilde{Y} = \tilde{y}(\tilde{P},Z,\tilde{\epsilon}_d) = y(\tilde{P},Z,\tilde{\epsilon}_d)$, we obtain $\tilde{Y} = Y$ since $\tilde{P} = P$ and $\tilde{\epsilon}_d = \epsilon_d$.

Next, we turn to cost and transfer. From (18), and using (B.2), $\tilde{\theta} = \alpha + \beta\theta$ and $\tilde{e} = \alpha + \beta e$, we have

$$\tilde{C} = (\tilde{\theta}-\tilde{e})\tilde{c}_o(\tilde{Y},Z,\tilde{\epsilon}_c) = (\theta-e)c_o(Y,Z,\epsilon_c) = C,$$

since $\tilde{Y} = Y$ and $\tilde{\epsilon}_c = \epsilon_c$. Moreover, from (20) and the previous results we obtain

$$\begin{aligned}
\tilde{T} &= \tilde{\psi}(\tilde{e},Z) + \int_{\tilde{\theta}}^{\overline{\tilde{\theta}}(Z)} \tilde{\psi}'[\tilde{e}^*(\tilde{u},Z),Z]d\tilde{u} - \tilde{\psi}'(\tilde{e},Z)\left\{\frac{\tilde{C}}{\overline{\tilde{c}}_o(\tilde{P},Z)} - (\tilde{\theta}-\tilde{e})\right\} + \tilde{\epsilon}_t \\
&= \psi(e,Z) + \int_{\alpha+\beta\theta}^{\alpha+\beta\overline{\theta}(Z)} \frac{1}{\beta}\psi'\left[e^*\left(\frac{\tilde{u}-\alpha}{\beta},Z\right),Z\right]d\tilde{u} - \frac{1}{\beta}\psi'(e,Z)\left\{\frac{\beta\tilde{C}}{\overline{c}_o(\tilde{P},Z)} - \beta(\theta-e)\right\} + \tilde{\epsilon}_t \\
&= \psi(e,Z) + \int_{\theta}^{\overline{\theta}(Z)} \psi'\left[e^*(u,Z),Z\right]du - \psi'(e,Z)\left\{\frac{\tilde{C}}{\overline{c}_o(\tilde{P},Z)} - (\theta-e)\right\} + \tilde{\epsilon}_t,
\end{aligned}$$

where the second equality uses (B.2), $\tilde{\psi}(\cdot,\cdot) = \psi[(\cdot-\alpha)/\beta,\cdot]$, $\tilde{\theta} = \alpha+\beta\theta$, $\tilde{e} = \alpha+\beta e$ and $\tilde{e}^*(\tilde{\theta},Z) = \alpha + \beta e^*[(\tilde{\theta}-\alpha)/\beta,Z]$, while the third equality follows from the change of variable $u = (\tilde{u}-\alpha)/\beta$. Thus, (20) implies that $\tilde{T} = T$ since $\tilde{C} = C$, $\tilde{P} = P$ and $\tilde{\epsilon}_t = \epsilon_t$.

Lastly, because of the linear transformation given every value of Z, it is easy to verify that the structure $\tilde{S}$ satisfies A1–A3 and B1 as soon as the structure S satisfies these assumptions.$\square$

Proof of Lemma 4: Recall that $P = p^*(\theta,Z)$, where $p^*(\cdot,\cdot)$ is the optimal price schedule. Assumption B1 implies that P is independent of ϵ_d given Z. Hence, (17) gives $\mathrm{E}[Y|P=p,Z=z] = \mathrm{E}[y(p,z,\epsilon_d)|P=p,Z=z] = \mathrm{E}[y(p,z,\epsilon_d)|Z=z] = \int y(p,z,\epsilon_d)dG(\epsilon_d|z) = \overline{y}(p,z)$ by (22). This establishes (25).

Regarding (23), we recall that θ can be expressed as a function $\theta^*(P,Z) = p^{*-1}(P,Z)$, which is strictly increasing in P since $P = p^*(\theta,Z)$ is strictly increasing in θ by Lemma 2. Thus, e can be expressed as a function $e^*(P,Z)$, which is strictly decreasing in P because $e = e^*(\theta,Z)$ is strictly decreasing in θ by Lemma 2, while $\theta = \theta^*(P,Z)$ is strictly increasing in P. Now, from (18) and using $\theta - e = \theta^*(P,Z) - e^*(P,Z)$, we obtain

$$\begin{aligned}
\mathrm{E}[C|P=p,Z=z] &= (\theta-e)\mathrm{E}\Big[c_o[y(p,z,\epsilon_d),z,\epsilon_c]|P=p,Z=z\Big] \\
&= (\theta-e)\mathrm{E}\Big[c_o[y(p,z,\epsilon_d),z,\epsilon_c]|Z=z\Big] \\
&= (\theta-e)\int c_o[y(p,z,\epsilon_d),z,\epsilon_c]dG(\epsilon_d,\epsilon_c|z) \\
&= (\theta-e)\overline{c}_o(p,z), \qquad \text{(B.3)}
\end{aligned}$$

where $\theta - e = \theta^*(p, z) - e^*(p, z)$. The second equality follows from B1 and the last follows from (23). In particular, (B.3) establishes (27) with $\Delta(\cdot, \cdot)$ satisfying $\Delta(\cdot, \cdot) \geq 1$ and $\partial\Delta(\cdot, \cdot)/\partial p > 0$ because $\theta - e = \theta^*(p, z) - e^*(p, z)$ is strictly increasing in p with strictly positive derivative with respect to p by Lemma 2 implying $\theta^*(p, z) - e^*(p, z) \geq \theta^*[\underline{p}(z), z] - e^*[\underline{p}(z), z] = \underline{\theta}(z) - e^*[\underline{\theta}(z), z] = 1$ by (24). Moreover, writing (B.3) at $p = \underline{p}(z)$, which is the price for the most efficient firm with type $\underline{\theta}(z)$ and exerting the maximal effort $\overline{e}(z) = e^*[\underline{\theta}(z), z]$, we obtain

$$\mathrm{E}[C|P = \underline{p}(z), Z = z] = \overline{c}_o[\underline{p}(z), z], \tag{B.4}$$

because $[\underline{\theta}(z) - \overline{e}(z)] = 1$ by the normalization (24).

Next, we write (19) at $Z = z$ so that $P = p^*(\theta, z) = p$ and $e = e^*(\theta, z)$. Dividing the resulting equation by (B.3) we obtain

$$\frac{p\overline{y}'(p, z) + \mu\overline{y}(p, z)}{\mathrm{E}[C|P = p, Z = z]} = \frac{\overline{c}_o'(p, z)}{\overline{c}_o(p, z)}.$$

Integrating this differential equation from $\underline{p}(z)$ to some arbitrary $p \in [\underline{p}(z), \overline{p}(z)]$, where $\overline{p}(z) \equiv p^*[\overline{\theta}(z), z)]$ is the price for the least efficient type when $Z = z$, we obtain

$$\log\left(\frac{\overline{c}_o(p, z)}{\overline{c}_o(\underline{p}(z), z)}\right) = \int_{\underline{p}(z)}^{p} \frac{\tilde{p}\overline{y}'(\tilde{p}, z) + \mu\overline{y}(\tilde{p}, z)}{\mathrm{E}[C|P = \tilde{p}, Z = z]} d\tilde{p}.$$

Solving for $\overline{c}_o(p, z)$ and using the boundary condition (B.4) give (26).□

Proof of Proposition 4: Because $\theta - e = \theta^*(p, z) - e^*(p, z)$, differentiating (27) with respect to p gives

$$\frac{\partial\theta^*(p, z)}{\partial p} - \frac{\partial e^*(p, z)}{\partial p} = \frac{\partial\Delta(p, z)}{\partial p} > 0, \tag{B.5}$$

where $\partial\theta^*(\cdot, \cdot)/\partial p > 0$ and $\partial e^*(\cdot, \cdot)/\partial p < 0$ from Lemma 2. In particular, $\theta = \theta^*(P, Z)$ and $e = e^*(P, Z)$ are in bijections with P given Z. Thus, taking conditional expectation of (21) given $(P, Z) = (p, z)$, and using (B.3) together with $\mathrm{E}[\epsilon_t|P = p, Z = z] = \mathrm{E}[\epsilon_t|\theta, Z = z] = 0$ by B1, we obtain

$$\mathrm{E}[T|P{=}p, Z{=}z] = \psi(e, z) + \int_{\theta}^{\overline{\theta}(z)} \psi'[e^*(\tilde{\theta}, z), z]d\tilde{\theta}, \tag{B.6}$$

where $\theta = \theta^*(p, z)$, $\overline{\theta}(z) = \theta^*[\overline{p}(z), z]$ and $e = e^*(p, z)$. Differentiating (B.6) gives

$$\frac{\partial \mathrm{E}[T|P{=}p, Z{=}z]}{\partial p} = \psi'(e, z)\left(\frac{\partial e^*(p, z)}{\partial p} - \frac{\partial\theta^*(p, z)}{\partial p}\right), \tag{B.7}$$

where we have used $e^*(\theta, z) = e^*(p, z) = e$. Thus, (B.5), (B.7) and (28) give

$$\psi'(e, z) = \Gamma(p, z) > 0, \tag{B.8}$$

because $\psi'(\cdot, z) > 0$ by A2. Differentiating (B.8) again gives

$$\psi''(e, z)\frac{\partial e^*(p, z)}{\partial p} = \frac{\partial \Gamma(p, z)}{\partial p} < 0, \tag{B.9}$$

because $\psi''(\cdot, z) > 0$ by A2 and $\partial e^*(p, z)/\partial p < 0$ by Lemma 2. Using (B.8)–(B.9) into (20) at $Z = z$ so that $P = p^*(\theta, z) = p$ and $e = e^*(\theta, z)$, we obtain

$$\Gamma(p, z) + \mu \frac{G_{P|Z}(p, z)}{g_{P|Z}(p, z)} \frac{\partial \Gamma(p, z)}{\partial p} \frac{\partial \theta^*(p, z)/\partial p}{\partial e^*(p, z)/\partial p} = \overline{c}_o(p, z), \tag{B.10}$$

where we have used the property that $F(\theta|z)/f(\theta|z) = [G_{P|Z}(p, z)/g_{P|Z}(p, z)]\partial \theta^*(p, z)/\partial p$ because $\theta = \theta^*(p, z)$ is strictly increasing in p from Lemma 2.

We now solve (B.5) and (B.10) for $\partial e^*(p, z)/\partial p$ and $\partial \theta^*(p, z)/\partial p$ to obtain after some algebra

$$\frac{\partial e^*(p, z)}{\partial p} = -R(p, z) < 0 \tag{B.11}$$

$$\frac{\partial \theta^*(p, z)}{\partial p} = \frac{\partial \Delta(p, z)}{\partial p} - R(p, z) > 0, \tag{B.12}$$

where $R(p, z)$ is as given in (29) with $R(p, z) > 0$ because $\partial e^*(\cdot, z)/\partial p < 0$ by Lemma 2. Similarly, the right-hand side of (B.12) must be strictly positive because $\partial \theta^*(\cdot, z)/\partial p > 0$ by Lemma 2 leading to $\partial \Delta(p, z)/\partial p > R(p, z)$. Now, note that $e^*[\overline{p}(z), z] = e^*[\overline{\theta}(z), z] = 0$ by (24). Moreover, from (27), we have $\theta^*[\overline{p}(z), z] - 0 = \Delta[\overline{p}(z), z]$. Thus, integrating (B.11) and (B.12) from some arbitrary $p \in [\underline{p}(z), \overline{p}(z)]$ to $\overline{p}(z)$, and using the preceding boundary conditions, we obtain (31) and (32). As all the functions on the right-hand side of (32) are identified, it follows that the firm's type θ can be recovered from (p, z), and that the conditional distribution $F(\cdot|z)$ of type is identified as the distribution of $\theta = \theta^*(P, z)$, where P is distributed as $G_{P|Z}(\cdot|z)$.

Lastly, let $\underline{e}(z) \equiv e^*[\overline{p}(z), z] = e^*[\overline{\theta}(z), z] = 0$ by the normalization (24), and let $\overline{e}(z) \equiv e^*[\underline{p}(z), z] = e^*[\underline{\theta}(z), z]$. Integrating (B.8) from 0 to some arbitrary $e \in [\underline{e}(z), \overline{e}(z)]$, where $p = p^*(\cdot, z)$ is the inverse function of $e^*(\cdot, z)$, gives

$$\psi(e, z) = \psi(0, z) + \int_0^e \Gamma[p^*(\tilde{e}, z), z] d\tilde{e},$$

which establishes (30) since (B.6) evaluated at $p = \overline{p}(z)$ gives $\mathrm{E}[T|P = \overline{p}(z), Z = z] = \psi(0, z)$ as $e = e^*[\overline{p}(z), z] = 0$ and $\theta = \theta^*[\overline{p}(z), z] = \overline{\theta}(z)$ when $p = \overline{p}(z)$.□

Proof of Lemma 5: Let $\tilde{\epsilon}_d = y(p_o(Z), Z, \epsilon_d)$ for an arbitrary $p_o(\cdot) \in [\underline{p}(\cdot), \overline{p}(\cdot)]$. Thus, $\epsilon_d = y^{-1}[p_o(Z), Z, \tilde{\epsilon}_d]$ since $y(\cdot,\cdot,\cdot)$ is strictly increasing in ϵ_d by assumption. Moreover, let $y(p, z, y^{-1}[p_o(z), z, \tilde{\epsilon}_d]) = \tilde{y}(p, z, \tilde{\epsilon}_d)$, which is strictly increasing in $\tilde{\epsilon}_d$. Thus, we have $\tilde{y}(p_o(z), z, \tilde{\epsilon}_d)$ $= \tilde{\epsilon}_d$ thereby satisfying the first equality in (36). We need to verify that $\overline{y}(p, z) = \overline{\tilde{y}}(p, z)$. In particular,

$$\begin{aligned} \overline{\tilde{y}}(p, z) &= \mathrm{E}[\tilde{y}(p, z, \tilde{\epsilon}_d)|p, z] = \mathrm{E}[y(p, z, y^{-1}[p_0(z), z, \tilde{\epsilon}_d])|p, z] \\ &= \mathrm{E}[y(p, z, \epsilon_d)|p, z] = \overline{y}(p, z). \end{aligned}$$

We can apply the same reasoning for $\tilde{c}_o(y, z, \tilde{\epsilon}_c)$. Let $\tilde{\epsilon}_c = c_o(y_o(Z), Z, \epsilon_c)$ for an arbitrary $y_o(\cdot) \in [\underline{y}(\cdot), \overline{y}(\cdot)]$. Thus, $\epsilon_c = c_o^{-1}[y_o(Z), Z, \tilde{\epsilon}_c]$ since $c_o(\cdot,\cdot,\cdot)$ is strictly increasing in ϵ_c by assumption. Let $c_o(y, z, c_o^{-1}[y_o(z), z, \tilde{\epsilon}_c]) = \tilde{c}_o(y, z, \tilde{\epsilon}_c)$, which is strictly increasing in $\tilde{\epsilon}_c$. Thus, we have $\tilde{c}_o(y_0(z), z, \tilde{\epsilon}_c) = \tilde{\epsilon}_c$ thereby satisfying the second inequality in (36). We need to verify that $\overline{c}_o(p, z) = \overline{\tilde{c}}_o(p, z)$. In particular,

$$\begin{aligned} \overline{\tilde{c}}_o(y, z) &= \mathrm{E}[\tilde{c}_o(y, z, \tilde{\epsilon}_c)|p, z] = \mathrm{E}[c_o(\tilde{y}(p, z, \tilde{\epsilon}_d), z, c_o^{-1}[y_o(z), z, \tilde{\epsilon}_c])|p, z] \\ &= \mathrm{E}[c_o(y(p, z, \epsilon_d)z, \epsilon_c)|p, z] = \overline{c}_o(p, z). \end{aligned}$$

It remains to show that these two models are observationally equivalent, which is straightforward. Since $\tilde{\epsilon}_d = y(p_o(Z), Z, \epsilon_d)$, we have $\tilde{y}(P, Z, \tilde{\epsilon}_d) = y(P, Z, y^{-1}[p_o(Z), Z, y(p_o(Z), Z, \epsilon_d)]) =$ $y(P, Z, \epsilon_d)$. Similarly, since $\tilde{\epsilon}_c = c_o(y_o(Z), Z, \epsilon_c)$, we have $\tilde{c}_o(Y, Z, \tilde{\epsilon}_c) = c_o(Y, Z, c_o^{-1}[y_o(Z), Z,$ $c_o(y_o(Z), Z, \epsilon_c)]) = c_o(Y, Z, \epsilon_c)$. Thus, we have $Y = y(P, Z, \epsilon_d) = \tilde{y}(P, Z, \tilde{\epsilon}_d)$ and $C_o = c_o(Y, Z, \epsilon_c)$ $= \tilde{c}_o(Y, Z, \tilde{\epsilon}_c)$. Lastly, P is conditionally independent of $(\tilde{\epsilon}_d, \tilde{\epsilon}_c)$ given Z because P is conditionally independent of (ϵ_d, ϵ_c) given Z combined with $\tilde{\epsilon}_d = y(p_o(Z), Z, \epsilon_d)$ and $\tilde{\epsilon}_c = c_o(y_o(Z), Z, \epsilon_c)$. Because the latter functions are strictly increasing in ϵ_d and ϵ_c, respectively, it follows that $G_{\tilde{\epsilon}_d|Z}(\cdot|\cdot)$ and $G_{\tilde{\epsilon}_c|\tilde{\epsilon}_d}, Z(\cdot|\cdot,\cdot)$ are strictly increasing in their first arguments.□

Proof of Proposition 5: Part (i) follows Matzkin (2003) identification argument. Because ϵ_d is independent of θ given Z by B1, then ϵ_d is independent of $P = p^*(\theta, Z)$ given Z. Thus, if $G_{\epsilon_d|P,Z}(\cdot|\cdot,\cdot)$ denotes the conditional distribution of ϵ_d given (P, Z), then for every (p, z) we have $G_{\epsilon_d|Z}(\cdot|z) = G_{\epsilon_d|P,Z}(\cdot|p, z) = G_{Y|P,Z}[y(p, z, \cdot)|p, z]$ because $y(p, z, \cdot)$ is strictly increasing in ϵ_d. In particular, this shows that $G_{Y|P,Z}(\cdot|p, z)$ is strictly increasing in its first argument in view of B3-(ii). Hence, for every (p, z) we have

$$y(p, z, \cdot) = G_{Y|P,Z}^{-1}[G_{\epsilon_d|Z}(\cdot|z)|p, z]. \tag{B.13}$$

Moreover, letting $p = p_o(z)$ we obtain $G_{\epsilon_d|Z}(\cdot|z) = G_{Y|P,Z}[y(p_o(z), z, \cdot)|p_o(z), z] = G_{Y|P,Z}[\cdot|p_o(z),$

$z]$, where the second equality follows from the first normalization in (36). This establishes (38) and hence (37) using (B.13).

To prove (ii) we extend Matzkin's argument as $Y = y(P, Z, \epsilon_d)$ is not independent from ϵ_c given Z in $C_o = c_o(Y, Z, \epsilon_c)$. On the other hand, we exploit the fact that P is independent from ϵ_c given (ϵ_d, Z) because $P = p^*(\theta, Z)$ and θ is independent of ϵ_c given (ϵ_d, Z) by B1. Thus, similarly to above, we obtain $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d, z) = G_{\epsilon_c|\epsilon_d,P,Z}(\cdot|\epsilon_d, p, z) = G_{C_o|\epsilon_d,P,Z}\{c_o[y(p, z, \epsilon_d), z, \cdot]|\epsilon_d, p, z\} = G_{C_o|Y,P,Z}[c_o(y, z, \cdot)|y, p, z]$ because $c_o(y, z, \cdot)$ is strictly increasing in ϵ_c and $y \equiv y(p, z, \epsilon_d)$. In particular, $G_{C_o|Y,P,Z}(\cdot|y, p, z)$ is strictly increasing in its first argument in view of B3-(ii). Hence, we have

$$c_o(y, z, \cdot) = G^{-1}_{C_o|Y,P,Z}[G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d, z)|y, p, z], \tag{B.14}$$

for every (y, p, z, ϵ_d) satisfying $y = y(p, z, \epsilon_d)$. We now exploit the second normalization in (36). Given the second part of B3-(i), let $p = p_\dagger(z, \epsilon_d)$ in (B.14). We obtain

$$\begin{aligned} G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d, z) &= G_{C_o|Y,P,Z}\Big[c_o\{y[p_\dagger(z), z, \epsilon_d)], z, \cdot\} \mid y[p_\dagger(z), z, \epsilon_d)], p_\dagger(z, \epsilon_d), z\Big] \\ &= G_{C_o|Y,P,Z}\Big[c_o\{y_o(z), z, \cdot\} \mid y_o(z), p_\dagger(z, \epsilon_d), z\Big] \\ &= G_{C_o|Y,P,Z}\left[\cdot \mid y_o(z), p_\dagger, z\right], \end{aligned}$$

where the second equality follows from the second part of B3-(i), while the third equality follows from the second normalization in (36). This establishes (40). Equation (39) follows from (B.14). Lastly, $p_\dagger(\cdot, \cdot)$ is identified as $y_o(\cdot)$ is known and $y(\cdot, \cdot, \cdot)$ is identified by (i).□

Proof of Lemma 6: Equation (41) follows directly from (37) noting that $Y = y(P, Z, \epsilon_d)$. Similarly, noting that $C_0 = c_o(Y, Z, \epsilon_c)$, (39) gives

$$\epsilon_c = G^{-1}_{C_o|Y,P,Z}\Big(G_{C_o|Y,P,Z}(C_o|Y, P, Z)|y_o(Z), p_\dagger(Z, \epsilon_d), Z\Big).$$

Recall that $C_o = C/\Delta(P, Z)$ so that $G_{C|Y,P,Z}(\cdot|Y, P, Z) = G_{C_o|Y,P,Z}(\cdot/\Delta(P, Z)|Y, P, Z)$ and $G^{-1}_{C|Y,P,Z}(\cdot|Y, P, Z) = \Delta(P, Z)G^{-1}_{C_o|Y,P,Z}(\cdot|Y, P, Z)$. This establishes (42).

To establish (43), we make the change of variable $\tilde{p} = p^*(\tilde{e}, z)$ in (30), we obtain

$$\begin{aligned} \psi(e, z) &= \mathrm{E}[T|P = \overline{p}(z), Z = z] + \int_{p^*(0,z)}^{p^*(e,z)} \Gamma(\tilde{p}, z)\frac{\partial e^*(\tilde{p}, z)}{\partial p}d\tilde{p} \\ &= \mathrm{E}[T|P = \overline{p}(z), Z = z] + \int_{p}^{\overline{p}(z)} \Gamma(\tilde{p}, z)R(\tilde{p}, z)d\tilde{p}, \end{aligned} \tag{B.15}$$

where the second equality follows from $p^*(0, z) = \overline{p}(z)$ and (B.11). Thus, using (27), (B.6) and (B.8) into (21) we obtain

$$T = \mathrm{E}[T|P, Z] - \frac{\Gamma(P, Z)}{\overline{c}_o(P, Z)}\Big(C - \mathrm{E}[C|P, Z]\Big) + \epsilon_t. \tag{B.16}$$

We now compute $\Gamma(P,Z)/\overline{c}_o(P,Z)$. From (27), we note that $\Delta(P,Z)\overline{c}_o(P,Z) = \mathrm{E}[C|P,Z]$. Thus, differentiating with respect to p gives

$$\overline{c}_o(P,Z)\frac{\partial\Delta(P,Z)}{\partial p} = \frac{\partial \mathrm{E}[C|P,Z]}{\partial p} - \Delta(P,Z)\frac{\partial \overline{c}_o(P,Z)}{\partial p},$$

where

$$\frac{\partial \overline{c}_o(P,Z)}{\partial p} = \frac{\overline{c}_o(P,Z)}{\mathrm{E}[C|P,Z]}\Big[P\overline{y}'(P,Z)+\mu\overline{y}(P,Z)\Big] = \frac{1}{\Delta(P,Z)}\Big[P\overline{y}'(P,Z)+\mu\overline{y}(P,Z)\Big]$$

from (26)-(27). Hence,

$$\overline{c}_o(P,Z)\frac{\partial\Delta(P,Z)}{\partial p} = \frac{\partial \mathrm{E}[C|P,Z]}{\partial p} - \Big[P\overline{y}'(P,Z)+\mu\overline{y}(P,Z)\Big].$$

It follows from (28) that

$$\frac{\Gamma(P,Z)}{\overline{c}_o(P,Z)} = -\frac{\partial \mathrm{E}[T|P,Z]/\partial p}{\partial \mathrm{E}[C|P,Z]/\partial p - [P\overline{y}'(P,Z)+\mu\overline{y}(P,Z)]}.$$

This establishes (43) in view of (B.16).□

Proof of Lemma 7: To prove the first part, we need to show that the conditional distribution of (Y,C,P,T) given Z induced by a structure $S\in\mathcal{S}$ satisfying P1–P5 also satisfy C1. Regarding C1-(i), the first condition follows from (22) and A1 together with B3-(ii). The second condition follows by taking the conditional expectation of (18) given (P,Z) and using $\theta-e>0$ by A1 and $\overline{c}_o(p,z)>0$. The latter inequality follows from (23) using $c_o(\cdot,\cdot,\cdot)\geq 0$ together with B3-(ii). Regarding C1-(ii), from the second part of the proof of Proposition 3, we have $d\overline{t}/d(\theta-e^*(\theta,z)) = -\psi'[e^*(\theta,z),z] < 0$ implying $\psi'[e^*(\theta,z),z] > 0$. From (30), $\psi'(e,z) = \Gamma[p(e,z),z]$. Hence, $\Gamma[p(e,z),z]>0$. Regarding $\Gamma'[p(e,z),z]$, the proof of Lemma 1 shows that the second partial derivative of the firm's objective function is equal to $-\psi''(e)$ because of the linearity in C of the transfer (21). Thus, strict concavity implies $\psi''(e)>0$. As $\psi''(e) = \Gamma'[p(e,z),z]p'(e,z)$ from (30), where $p'(e,z)<0$ because $p'(\theta,z)>0$ and $e'(\theta,z)<0$ by assumption, it follows that $\Gamma'[p(e,z),z]<0$.

Regarding C1-(iii), we have $\Delta'(p,z) = \theta'(p,z)-e'(p,z) > 0$ because $p'(\theta,z)>0$ and $e'(p,z) < 0$ by assumption. Moreover, from (32), we have $\theta'(p,z) = \Delta'(p,z) - R(p,z) > 0$ by assumption. Thus $\Delta'(p,z) > R(p,z)$ or equivalently $\Gamma'(p,z) < \overline{c}_o(p,z)$ as discussed in the text. Regarding C1-(iv), Lemma 6 shows that (ϵ_d,ϵ_c) can be recovered through identified functions $\phi_d(Y,P,Z)$ and $\phi_c(Y,C,P,Z)$. Since P and θ are in a bijective relationship, the latter are conditionally independent of P given Z in view of B1. Moreover, (43) shows that $\mathrm{E}[\epsilon_t|P,Z]=0$. Regarding

C1-(v), the first part follows from $p'(\theta, z) > 0$ and $f_{\theta|Z}(\cdot|\cdot) > 0$, while the second part follows from B3-(ii), $G_{Y|P,Z}(\cdot|p, z) = G_{\epsilon_d|Z}[y^{-1}(p, z, \cdot)|z]$ and $G_{C|Y,P,Z}(\cdot|y, p, z) = G_{\epsilon_c|\epsilon_d,Z}[c_o^{-1}(y, z, \cdot/(\theta - e))|\epsilon_d, z]$, where the latter uses the bijective mapping bewteen P and θ given z and B1-B2.

Turning to the second part, let the observations (Y, P, C, T, Z) and a function $\mu(\cdot)$ satisfy C1. We need to define $[y, c_o, \psi, F, G, \mu]$ from the observables and show that these functions satisfy A1, B1–B3, as well as P1–P5. In view of Proposition 5 with $\bigcap_{\epsilon_d \in [\underline{y}(p_o(z)z), \overline{y}(p_o(z), z)]} \{y = y(p, z, \epsilon_d), p \in [\underline{p}(z), \overline{p}(z)]\}$ nonempty for every $z \in \mathcal{Z}$, let

$$
\begin{aligned}
y(p, z, \epsilon_d) &= G^{-1}_{Y|P,Z}[G_{Y|P,Z}(\epsilon_d|p_o(z), z)|p, z] \\
c_o(y, z, \epsilon_c) &= G^{-1}_{C_o|Y,P,Z}\Big[G_{C_o|Y,P,Z}[\epsilon_c|y_o(z), p_\dagger(z, \epsilon_d), z]|y, p, z\Big],
\end{aligned}
$$

for some $[p_o(\cdot), y_o(\cdot)]$, where $y = y(p, z, \epsilon_d)$ and $C_o = C/\Delta(P, Z)$ with $\Delta(p, z)$ given in (27). Note that B3-(i) is satisfied by construction. By C1-(v), $y(p, z, \cdot)$ and $c_o(y, z, \cdot)$ are strictly increasing in their last arguments. Thus, we can define $\epsilon_d = y^{-1}(P, Z, Y)$ and $\epsilon_c = c_o^{-1}(Y, Z, C_o)$. Note that $C = \Delta(P, Z)C_o = \Delta(P, Z)c_o(Y, Z, \epsilon_c)$, where $\Delta(P, Z) = \theta - e$ and $c_o(Y, Z, \epsilon_c) \geq 0$ because $C \geq 0$. Using C1-(i), $\mathrm{E}[C|P, Z] = \Delta(P, Z)\overline{c}_o(P, Z) > 0$ leading to $\Delta(P, Z) > 0$ because $\overline{c}_o(P, Z) > 0$ by (26). Thus the first part of A1 is satisfied. Using C1-(iv), we have $G_{\epsilon_d|Z}(\cdot|z) = G_{\epsilon_d|P,Z}(\cdot|p_o(z), z) = G_{Y|P,Z}(\cdot|p_o(z), z)$. Thus, $G_{\epsilon_d|Z}(\cdot|z)$ is nondegenerated and strictly increasing in its first argument by C1-(v). Similarly, using C1-(iv), we have $G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d, z) = G_{\epsilon_c|\epsilon_d,P,Z}(\cdot|\epsilon_d, p_\dagger(z, \epsilon_d), z) = G_{\epsilon_c|Y,P,Z}(\cdot|y_o(z), p_\dagger(z, \epsilon_d), z) = G_{C_o|Y,P,Z}[\cdot|y_o(z), p_\dagger(z, \epsilon_d), z]$, which is nondegenerated and strictly increasing in its first argument by C1-(v) and $C_o = C/\Delta(P, Z)$ with $\Delta(P, Z) > 0$. Hence, B3-(ii) is satisfied.

In view of Proposition 4, let

$$
\begin{aligned}
\psi(e, z) &= E[T|P = \overline{p}(z), Z = z] + \int_0^e \Gamma[p(\tilde{e}, z)]d\tilde{e} \\
\theta &= \theta(P, Z) \equiv \Delta(P, Z) + \int_P^{\overline{p}(Z)} R(\tilde{p}, Z)d\tilde{p},
\end{aligned}
$$

where $p(\cdot, z)$ is the inverse of $e(\cdot, z) = \int_{\cdot}^{\overline{p}(z)} R(\tilde{p}, z)d\tilde{p}$, which is strictly decreasing as $R(\cdot, \cdot) > 0$ from C1-(iii). The distribution $F(\theta|Z)$ is then defined as the distribution of $\theta(P, Z)$ given Z. Note that $\theta'(p, z) > 0$ because $\Delta'(p, z) > R(p, z)$ by C1-(iii). Thus $F(\theta|z)$ admits a density, which is strictly positive on $[\underline{\theta}(z), \overline{\theta}(z)]$ as defined in (34) and (35) by C1-(v). Moreover, P2-P3 hold from C1-(iii). For, P3 follows from $\theta'(p, z) > 0$ and P2 follows from $e'(p, z) = -R(p, z) < 0$ by C1-(iii). In addition, B2 is satisfied as $\underline{\theta}(z) - e(\underline{\theta}(z), z) = \Delta(\underline{p}(z), z) = 1$ by (26) and (27), while $e(\overline{\theta}(z), z) = 0$ because $e(\overline{p}(z), z) = 0$ by construction and $p(\cdot, z)$ strictly increasing leading to $\overline{p}(z) = p(\overline{\theta}, z)$.

Note that (ϵ_d, ϵ_c) are conditionally independent of P given Z by Lemma 6 and C1-(iv), and hence of θ given Z. Moreover, (43) implies $\mathrm{E}[\epsilon_t|P, Z] = 0$. Thus, B1 and the second part of A1 are satisfied. Lastly, P1 and P4-P5 hold from C1-(ii). For, with the transfer (43), the proof of Lemma 1 shows that P1 is equivalent to $\psi''(e, z) > 0$, which is ensured by $\psi''(e, z) = \Gamma'(p, z)e'(p, z)$, $\Gamma'(p, z) < 0$ by C1-(ii) and $e'(p, z) < 0$ as above. Similarly, the proof of Proposition 3 shows that P4 holds when $e'(\theta, z) < 0$, which is established above. Moreover, the proof of Proposition 3 shows that P5 holds when $\psi'(e, z)) > 0$, which follows from C1-(ii). □

Proof of Proposition 6: Let $S = [y, c_o, \psi, F, G, \lambda]$ be a structure inducing a distribution for (Y, C, P, T) given Z satisfying C1. Define $\tilde{\lambda}(\cdot) = \lambda(\cdot) + \epsilon$, with $\epsilon \neq 0$ sufficiently small so that $\tilde{\lambda}(\cdot)$ and the distribution of (Y, C, P, T) given Z satisfy C1-(ii,iii). Note that C1-(i,v) are still satisfied as these assumptions do not involve $\tilde{\lambda}(\cdot)$. We need to show that C1-(iv) is satisfied as well. We have $\tilde{\phi}_d(\cdot, \cdot, \cdot) = \phi_d(\cdot, \cdot, \cdot)$ and $\tilde{\phi}_c(\cdot, \cdot, \cdot, \cdot) = phi_c(\cdot, \cdot, \cdot, \cdot)$ from (41) and (42), respectively. On the other hand, $\tilde{\phi}_t(\cdot, \cdot, \cdot, \cdot, \cdot) \neq \phi_t(\cdot, \cdot, \cdot, \cdot, \cdot)$ but $\mathrm{E}[\tilde{\phi}_t(Y, C, P, T, Z)|P, Z] = 0$ by (43). Thus, from Lemma 7, there exists a structure $\tilde{S} = [\tilde{y}, \tilde{c}_o, \tilde{\psi}, \tilde{F}, \tilde{G}, \tilde{\lambda}]$ that satisfies A1, B1–B3 as well as P1–P5 and that rationalizes the observables (Y, C, P, T) given Z. The structure $\tilde{S}$ differs from S because $\tilde{\lambda}(\cdot) \neq \lambda(\cdot)$. Though $\tilde{y}(\cdot, \cdot, \cdot) = y(\cdot, \cdot, \cdot)$, $\tilde{G}_{\epsilon_d|Z}(\cdot|\cdot) = G_{\epsilon_d|Z}(\cdot|\cdot)$ and $\tilde{G}_{\epsilon_c|\epsilon_d, Z}(\cdot|\cdot, \cdot) = G_{\epsilon_c|\epsilon_c, Z}(\cdot|\cdot, \cdot)$, the remaining functions differ.□

Proof of Proposition 7: It suffices to show that C is a valid instrument as (45) follows from (44) with $W = C$. By C2, we have $\mathrm{Cov}[C, \epsilon_t|P, Z] = 0$, while C is perfectly correlated with itself given (P, Z). Lastly, $\mathrm{Cov}[C, T|P, Z] = \mathrm{E}\{C(T - \mathrm{E}[T|P, Z])|P, Z\} = -[\Gamma(P, Z)/\overline{c}_o(P, Z)]\mathrm{Var}[C|P,$ $Z]$ by (B.16). Thus, $\mathrm{Cov}[C, T|P, Z] < 0$ as $\Gamma(\cdot, \cdot) > 0$ and $\overline{c}_o(\cdot, \cdot) > 0$.□

References

Ballard, C., **J. Shoven** and **J. Whalley** (1985): "General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States," *American Economic Review*, 75, 128-138.

Baron, D. (1989): "Design of Regulatory Mechanisms and Institutions," in R. Schlamensee and R. Willig, eds., *Handbook of Industrial Organization, Volume II*, North Holland.

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

Besanko, D. (1984): "On the Use of Revenue Requirements Regulation Under Imperfect Information," in M. Crew, ed., *Analyzing the Impact of Regulatory Change in Public Utilities*, Lexington Books.

Blundell, R. and **J. Powell** (2003): "Endogeneity in Semiparametric and Nonparametric Regression Models," in M. Dewatripont, L.P. Hansen and S. Turnovky, eds., *Advances in Economics and Econometrics: Eighth World Congress, Volume II*, Cambridge University Press.

Bolton, P. and **M. Dewatripont** (2005): *Contract Theory*, MIT Press.

Bresnahan, T. (1989): "Empirical Studies of Industries with Market Power," in R. Schmalensee and R. Willig, eds., *Handbook of Industrial organization, Volume II*, North Holland.

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

Chernozhukov, V., **H. Hong** and **E. Tamer** (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243-1284.

Chiappori, P.A. and **B. Salanié** (2000): "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 108, 56-78.

Chiappori P.A. and **B. Salanié** (2003): "Testing Contract Theory: A Survey of Some Recent Work," in M. Dewatripont, L. Hansen and S. Turnovsky, eds., *Advances in Eco-*

nomics and Econometrics: Eighth World Congress, Volume II, Cambridge University Press.

D'Haultfoeuille, X. and **P. Février** (2007): "Identification and Estimation of Incentive Problems: Adverse Selection," Working Paper, CREST-INSEE.

Février, P., **R. Preget** and **M. Visser** (2004): "Econometrics of Share Auctions," Working Paper, CREST.

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

Guerre, E., **I. Perrigne** and **Q. Vuong** (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525-574.

Guerre, E., **I. Perrigne** and **Q. Vuong** (2008): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," Working Paper, Pennsylvania State University.

Hastie, T. and **R. Tibshirani** (1990): *Generalized Additive Models*, Chapman and Hall.

Huang, Y., **I. Perrigne** and **Q. Vuong** (2008): "Nonlinear Pricing in Yellow Pages," in progress.

Ivaldi, M. and **D. Martimort** (1994): : "Competition under Nonlinear Pricing," *Annales d'Economie et de Statistique*, 34, 71-114.

Jorgenson, D. and **K.Y. Yun** (1991): "The Excess Burden of US Taxation," *Journal of Accounting, Auditing and Finance*, 6, 487-509.

Joskow, P. (2005): "Regulation and Deregulation After 25 Years: Lessons Learned for Research In Industrial Organization," *Review of Industrial Organization*, 26, 169-193.

Joskow, N. and **N. Rose** (1989): "The Effects of Economic Regulation," in R. Schlamensee and R. Willig, eds., *Handbook of Industrial Organization, Volume II*, North Holland.

Joskow, P. and **R. Schmalensee** (1986): "Incentive Regulation for Electric Utilities," *Yale Journal on Regulation*, 4, 1-49.

Laffont, J.J. (1994): "The New Economics of Regulation Ten Years After," *Econometrica*, 62, 507-537.

Laffont, J.J. (2005): *Regulation and Development*, Cambridge University Press.

Laffont, J.J. and **D. Martimort** (2001): *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press.

Laffont, J.J. and **J. Tirole** (1986): "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, 94, 614-641.

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

Laffont, J.J. and **Q. Vuong** (1996): "Structural Analysis of Auction Data," *American Economic Review, Papers and Proceedings,* 86, 414-420.

Li, T., **I. Perrigne** and **Q. Vuong** (2000): "Conditionally Independent Private Information in OCS Wildcat Auctions," *Journal of Econometrics,* 98, 129-161.

Manski, C. and **E. Tamer** (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519-547.

Matzkin, R. (1994): "Restrictions of Economic Theory in Nonparametric Methods," in R. Engle and D. McFadden, eds., *Handbook of Econometrics, Volume IV*, North Holland.

Matzkin, R. (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339-1375.

Matzkin, R. (2005): "Identification in Nonparametric Simultaneous Equations," Working Paper, Northwestern University.

Miravete, E. (2002): "Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans," *Review of Economic Studies*, 69, 943-971.

Paarsch, H. (1992): "Deciding Between the Common Value and Private Value Paradigms in Empirical Models of Auctions," *Journal of Econometrics*, 51, 191-215.

Paarsch, H. and **B. Shearer** (2000): "Piece Rates, Fixed Wages and Incentive Effects: Statistical Evidence From Payroll Records," *International Economic Review*, 41, 59-92.

Park, B., **R. Sickles** and **L. Simar** (1998): "Stochastic Panel Frontiers: A Semiparametric Approach," *Journal of Econometrics*, 84, 273-301.

Perrigne, I. (2002): "Incentive Regulatory Contracts in Public Transportation: An Empirical Study," Working Paper, Pennsylvania State University.

Perrigne, I. and **S. Surana** (2004): "Politics and Regulation: The Case of Public Transit," Working Paper, Pennsylvania State University.

Perrigne, I. and **Q. Vuong** (2008a): "Nonidentification of the Incentive Regulation Model with a General Cost Function," in progress.

Perrigne, I. and **Q. Vuong** (2008b): "Estimating Cost Efficiency Under Asymmetric Information," in progress.

Prakasa Rao, B. (1992): *Identifiability in Stochastic Models*, Academic Press.

Roehrig, C. (1988): "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56, 433-447.

Thomas, A. (1995): "Regulating Pollution Under Asymmetric Information: The Case of Industrial Wastewater Treatment," *Journal of Environmental Economics and Management*, 28, 357-373.

Walters, M. and **E. Auriol** (2007): "The Marginal Cost of Public Funds in Developing Countries: An Application to 38 African Countries," Working Paper, IDEI.

Wolak, F. (1994): "An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction," *Annales d'Economie et de Statistiques*, 34, 13-69.