Econometrics of Incentive Regulation: Nonparametric Identification*

Isabelle Perrigne
Pennsylvania State University

Quang Vuong
Pennsylvania State University

September 2004
Preliminary and Incomplete Draft

---

* This paper is in memory of J.J. Laffont. Jean-Jacques Laffont's dedication to his profession has been a source of inspiration to us. We dedicate this paper to him, which arose from a term paper we wrote for a course that he taught in Spring 2003. A preliminary version of this paper was presented at Johns Hopkins University and PennState University. We thank the participants for much constructive comments.

Abstract

This paper studies the nonparametric identification of the incentive regulation model with ex post observed costs developed in Laffont and Tirole (1986). We first extend such a basic model to general random demand and cost functions, while considering a monopolist producing a private good. We then map the resulting model into a structural econometric model that allows for some unobserved heterogeneity. We establish the nonparametric identification of the cost of public funds, the demand, cost and effort functions, 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 unobserved heterogeneity.

Fields: Incentive Regulation, Optimal Contracts, Cost Efficiency, Nonparametric Identification, Unobserved Heterogeneity.

Address: Pennsylvania State University, Department of Economics, University Park, PA 16802-3306.

Econometrics of Incentive Regulation: Nonparametric Identification

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. A simple example of information asymmetries is an auction, in which the seller does not know bidders' values for the auctioned object and the bidders do not know their competitors' values. As a response, bidders play strategically. Game theory provides a useful tool for analyzing such behavior through the Bayesian Nash equilibrium. Contracts provide another important example of how information asymmetry governs relationships between (say) a principal who designs the contract and an agent. Two types of imperfect information 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.[1] The agent plays strategically as he can cheat on his own characteristics and he can minimize effort. Thus, the principal has to give to the agent the right incentives to alleviate such problems through the terms of the contract.[2]

---

[1] An auction model can be viewed as a model with adverse selection where bidders' values can be considered as hidden characteristics. The parallel between auctions and contracts will be emphasized later in the paper.

[2] See the book by Laffont and Martimort (2001) on the theory of incentives.

Contracts are widely used in the economic world. Just to name a few, agriculture, insurance, retailing and management provide many examples of contractual relationships. In this paper, we are interested in regulatory contracts between a regulatory authority or regulator and a monopolistic firm. The government regulates firms to prevent natural monopolistic behavior. Public agencies may also be concerned by redistribution aspects. We developed such an interest for several reasons. First, regulation governs many industries such as utilities (electricity, water, gas, telecommunications) as well as many services (postal service, public transit, railroad), which represent an important component of the economy. Second, data on regulatory contracts are more readily available. Given the public nature of regulatory commissions, data are in general accessible to the analyst in contrast to "private" contracts, in which data may be subject to confidentiality issues. Third, regulatory contracts are in general well defined in terms of the objectives assigned to the regulated firm and the compensation arrangements made by the regulator.[3] Fourth, the economic literature provides a solid background to analyze regulation in a framework of imperfect information as surveyed by Baron (1989), Laffont and Tirole (1993) and Laffont (1994).[4] The incentive regulation model introduced by Laffont and Tirole (1986) represents a breakthrough in the new economics of regulation. This paper explains the trade-off faced by the regulator between firm's rent extraction and efficiency. This trade-off is the key issue in incentive regulation in presence of information asymmetries. Namely, the regulator has to provide some incentives to the firm to reveal his information or type as well as to exert appropriate effort. Though the regulator would like to extract all the firm's rent as leaving rents to the firm is costly to the society, the regulator must give up some rent to achieve revelation and efficiency from the firm. Moreover, subsidizing the firm through some monetary transfer requires additional taxes, which are costly to the society.[5] Laffont and Tirole (1986) have shown that the regulator can achieve such a

---

[3] Private contracts can be under an implicit form or subject to many unspecified terms such as in incomplete contracts.

[4] Laffont (1994) refers to the new economics of regulation in opposition to the classical theory of regulation neglecting information asymmetries. The latter literature has provided, however, important pricing rules such as the Ramsey-Boiteux pricing and the peak-load pricing rules.

[5] The Laffont and Tirole (1986) model considers a case when transfers are legally possible. By federal law, such transfers may be forbidden. Transfers in this case may take a different form.

trade-off while using the expost observable costs in the transfer paid ex post to the firm.[6] This results in an increase of the social welfare relatively to the Baron and Myerson (1982) model in which no expost cost information is used.[7]

In contrast to these important theoretical developments and the economic importance of regulation in our economies, very few empirical studies relying on a structural modeling have been performed. As a matter of fact, the empirical literature is mostly limited to a reduced form approach as surveyed by Joskow and Rose (1989).[8] Though this literature acknowledges the presence of information asymmetries, very few papers have attempted to estimate a regulatory contract model. Notable exceptions are Wolak (1994) for the regulation of water utilities, Gagnepain and Ivaldi (2002) and Perrigne (2002) for public transportation.[9] This clearly did not meet the expectation of theorists. Laffont (1994, p. 532) believes 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) state 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 of the complexity of the models to be estimated. Asymmetric information models lead to highly nonlinear models whose estimation requires suitable econometric tools. Moreover, the issue of identification needs to be addressed.[10] Parametric identification can in principle be achieved but such results

---

[6] Firms are usually submitted to an annual audit of their financial results and costs by the regulatory commission. In western economies, accounting rules are well defined and such data are reliable. In developing countries, the problem is somewhat different. Morever, the cost of public funds may be large. Both characteristics give rise to different incentive rules as shown by Laffont (2004).

[7] In addition to these nice features, the Laffont and Tirole (1986) model combines both adverse selection and moral hazard, while reducing adverse selection also known as the false moral hazard problem.

[8] Empirical studies have analyzed the effects of regulation on price, product quality, innovation and productivity growth to name a few. The determinants of regulatory mechanisms have also been analyzed. Some of these studies rely on natural experiments when a change in the regulatory process takes place.

[9] The situation is quite similar for the analysis of contract data in general. See the survey by Chiappori and Salanie (2003).

[10] Another important related question is to derive the restrictions imposed by the model on observables to test the model validity. Without such restrictions, any model could explain the data.

can be questioned as the identification closely depends on particular functional forms.[11] Moreover, important misspecification issues may arise. The recent literature on the structural analysis of auction models constitutes a stepping stone from which the econometrics of contract models can develop.

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).[12] Our results are general in the sense that other contract models can be identified using a similar approach. We need first to adapt the model. In particular, these authors consider the regulation of a public good with a fixed demand, while a private good with a random demand seems to be the most prevalent case in regulation.[13] The difficulty lies in that the contract design needs to consider expected demand, while the firm will have to fullfill the realized demand. Additional difficulties lie in the contract implementation and in checking whether the local second-order conditions are globally satisfied. In this respect, some assumptions need to be strengthened.

We then derive the corresponding structural econometric model. In particular, the error terms need to arise naturally from the theoretical model. We then face a number of complications. The theoretical model is roughly defined by a demand function subject to some random shock, a cost function depending on the unobserved firm's type and effort subject to some random shock, the cost of public funds, the effort disutility function and the firms' type distribution. The observables are the (ex post) demand, the (ex post) cost, the price decided by the regulator and the transfer paid 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 is related 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 transfer). Thus,

---

[11] The econometrics of auction models provides interesting examples in this respect. For instance, the common value auction model can be estimated parametrically using specific distributions for the structure of the model, while this model is not identified in general. See Paarsch (1992) and Fevrier, Preget and Visser (2004) for parametric estimation of common value auction models. See Laffont and Vuong (1996) and Li, Perrigne and Vuong (2000) for the nonindentification of the common value model.

[12] The problem of deriving restrictions is left for future research. The nonparametric estimation of the model is treated in a separate paper. See Perrigne and Vuong (in progress).

[13] See the previous examples, in which the consumer needs to pay to get access to it.

the econometric model is singular. We thus introduce an additional term representing unobserved heterogeneity, which can be added equivalently to the effort disutility function or the transfer function. An additional advantage of such an error term is that it can be used to assess the explanatory power of the Laffont and Tirole's model. The econometric model is then based on five equations, which are the demand, the cost, the *generalized* Ramsey pricing, the optimal effort level and the transfer.[14]

The econometric model allows for exogenous variables capturing observed firm, regulator, and/or market heterogeneity, which can affect all the functions and distributions in the model. We first show that the model is nonparametrically identified given the cost of public funds under a multiplicative decomposition of the cost function into a base cost and a cost inefficiency level. To give a flavor of our results, two assumptions are made. The firm's type is assumed to be conditionally independent of the three other random shocks in the models, while a natural normalization is imposed for the cost inefficiency level.[15] The nonparametric identification will then rely on the one-to-one mapping defined by the model between the price and the firm's type. Using the price distribution will allow us to identify the optimal effort level as well as the firm's type. The analogy with the auction model becomes clear at this point. In particular, the bidder's private value and his bid in an auction model are equivalent to the firm's type and the price in a regulatory contract. Guerre, Perrigne and Vuong (2000) recover the bidder's private value using the bid distribution as the equilibrium defines a monotonic function relating the bid to the private value. A similar idea is exploited here. Given such results, we then address the nonparametric identification of the cost of public funds.

This paper represents an important step towards the development of the econometrics of contract models. First, we show that the Laffont and Tirole (1986) model is nonparametrically identified from observables. The complexity of the model and its technical difficulties represent numerous challenges to the econometrician as described above. In particular, we show that the estimation of a cost frontier as performed typically in production frontier analyses (see, e.g. Gagnepain and Ivaldi (2002)) does not exploit all the

---

[14] The *generalized* Ramsey pricing and the optimal effort level equations are directly derived from the first-order conditions of the regulator's maximization problem.

[15] Such a normalization is necessary as a linear transformation of the firm's type will lead to an observationally equivalent model causing the nonidentification of the model.

information in the theoretical model and may produce a biased estimate of the base cost function.[16] Second, given the many ingredients included in this model, our identification result can be extended to other contract models. A contract model with adverse selection is a simplified version of this model as the effort disutility and the cost of public funds are not part of the model structure. As such, nonlinear pricing and many contractual data besides regulatory contract data could be entertained. Third, the nonparametric nature of our identification result will lead to an estimation method that is robust to misspecification. In contrast, Perrigne (2002) develops a parametric estimation based on parametric identification of the model.[17] Our paper is also parsimonious in error terms introduced in the model in contrast to Wolak (1994). While considering the estimation of the Baron and Myerson (1982) model, Wolak (1994) considers a fully parametric estimation method while adding "econometrician" error terms to the first-order conditions of the regulator's maximization problem in addition to the demand and cost error terms and the firm's type. Lastly, our results offer a new approach to the estimation of cost efficiency.[18] As a matter of fact, in view of informational asymmetries, it is reasonable to consider that firms's costs are subject to adverse selection and moral hazard issues.

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 the nonparametric identification of the model including the definition of the econometric model, while considering first the identification of the structure for a given cost of public funds. Section 4 concludes.

---

[16] See Lemma 4 and the following discussion.

[17] Note, however, that this paper discusses some technical difficulties addressed here such as the generalization of the Laffont and Tirole's model, the necessary normalization of the cost inefficiency index and the conditional independence of the firm's type and the demand shock among others. This paper does not use information on the transfer. Thus the additional term of unobserved heterogeneity is not necessary. This paper is based on three observed endogenous variables (demand, cost, price) determined by three random variables (demand and cost random shocks and firm's type). The paper also discusses the need of the transfer information to identify possibly semiparametrically the model.

[18] See Park and Simar (1994) and Park, Sickles and Simar (1998) for some recent developments in the semiparametric estimation of frontier models.

2 The Model

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

$$\begin{aligned} y &= y(p, \epsilon_d) \geq 0 \\ c &= c(y, \theta - e, \epsilon_c) \geq 0, \end{aligned}$$

where y is the quantity of private good, c is the corresponding costs, p is the price per unit of private good, θ represents the firm's type, e is the level of effort exerted by the firm, and (ϵ_d, ϵ_c) are random shocks to demand and costs. As usual, θ and e are private information to the firm, where θ is the (scalar) adverse selection parameter known to be distributed as $F(\cdot)$ with density $f(\cdot) > 0$ on its support $[\underline{\theta}, \overline{\theta}]$. The random shocks (ϵ_d, ϵ_c) are known to be jointly distributed as $G(\cdot, \cdot)$ *independently* from θ.[19]

The regulator offers a price schedule $p(\tilde{\theta})$ 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. The random shocks (ϵ_d, ϵ_c) are realized *ex post*, i.e. after contractual arrangements have been made between the regulator and the monopolist. 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.

Throughout, 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 $a_k(\cdot)$. All discussions of assumptions and second order conditions are relegated to subsection 2.5.

2.1. The Firm's Problem

Given the price $p(\cdot)$ and transfer $t(\cdot, \cdot)$ functions chosen by the regulator, the realized

---

[19] Note that ϵ_c is assumed to be independent of θ in Laffont and Tirole (1986), while ϵ_d is void because the demand y is fixed. Laffont and Tirole (1986) also consider a constant marginal cost function, namely $c = (\theta - e)y + \epsilon_c$, while Laffont and 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 cost function we consider starting from Section 2.4.

utility for the firm with type θ when it announces $\tilde{\theta}$ and exerts effort e is

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

where $\psi(e)\geq 0$ 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

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

where the independence between θ and (ϵ_d,ϵ_c) has been used.

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

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

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

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

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

$$\int t_2\Big(\tilde{\theta},c\{y[p(\tilde{\theta}),\epsilon_d],\theta-e,\epsilon_c\}\Big)c_2\{y[p(\tilde{\theta}),\epsilon_d],\theta-e,\epsilon_c\}\; dG(\epsilon_d,\epsilon_c)=-\psi'(e). \tag{3}$$

Denote the corresponding expected utility by

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

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, we must have $U_1(\theta,\theta)=0$ for any θ. Equivalently, denoting $U(\theta) \equiv U(\theta,\theta)$ and $e(\theta) \equiv e(\theta,\theta)$, and using $U'(\theta) = U_1(\theta,\theta) + U_2(\theta,\theta)$ and then (4), we obtain

$$\begin{aligned}
U'(\theta) &= U_2(\theta,\theta) \\
&= \int U_2(\theta,\theta,e(\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) + e_2(\theta,\theta)\int U_3(\theta,\theta,e(\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) \\
&= \int U_2(\theta,\theta,e(\theta),\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) \\
&= \int t_2\Big(\theta, c\{y[p(\theta),\epsilon_d],\theta - e,\epsilon_c\}\Big)c_2\{y[p(\theta),\epsilon_d],\theta - e,\epsilon_c\}\ dG(\epsilon_d,\epsilon_c),
\end{aligned}$$

where the third equality follows from (2) since $e(\theta) = e(\theta,\theta)$, and the fourth equality follows 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

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

2.3. THE REGULATOR'S PROBLEM

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

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

where $\lambda > 0$ is the shadow cost of public funds. Thus, using the independence of θ and (ϵ_d,ϵ_c), the expected social welfare is

$$\int SW(\theta,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c)dF(\theta) =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} \Big\{ \int \Big[\int_{p(\theta)}^{\infty} y(v,\epsilon_d)dv + (1+\lambda)p(\theta)y(p(\theta),\epsilon_d)$$
$$-(1+\lambda)\Big(\psi(e(\theta)) + c[y(p(\theta),\epsilon_d),\theta - e(\theta),\epsilon_c]\Big)\Big]dG(\epsilon_d,\epsilon_c) - \lambda U(\theta)\Big\}dF(\theta), \quad (8)$$

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 constraint and the participation constraint, i.e. to

$$U'(\theta) \;=\; -\psi'(e(\theta)) \quad (9)$$
$$U(\theta) \;\geq\; 0, \quad (10)$$

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 $U'(\theta) < 0$ under the condition that $\psi'(\cdot) > 0$, which is assumed hereafter. Hence, the participation constraint (10) can be written equivalently as $U(\overline{\theta}) \geq 0$ or

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

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

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

Proposition 1: *The functions $p^*(\cdot)$ and $e^*(\cdot)$ that solve the FOC of the optimization problem (P') satisfy*

$$\begin{aligned} \frac{p - \widetilde{mc}}{p} &= \frac{\lambda}{1+\lambda}\frac{1}{\tilde{\eta}} && (12) \\ \psi'(e) &= \overline{cs}_e - \frac{\lambda}{1+\lambda}\frac{F(\theta)}{f(\theta)}\psi''(e), && (13) \end{aligned}$$

where $p = p^*(\theta)$, $e = e^*(\theta)$ and

$$\begin{aligned} \widetilde{mc} &= \frac{E_\epsilon\{c_1[y(p,\epsilon_d), \theta - e, \epsilon_c] y_1(p,\epsilon_d)\}}{E_\epsilon\{y_1(p,\epsilon_d)\}} \\ \tilde{\eta} &= -\frac{p\overline{y}'(p)}{\overline{y}(p)} \\ \overline{cs}_e &= E_\epsilon\{c_2[y(p,\epsilon_d), \theta - e, \epsilon_c]\}. \end{aligned}$$

Note that $\widetilde{mc}$ differs from the expected marginal cost $\overline{mc} \equiv E_\epsilon\{c_1[y(p,\epsilon_d), \theta - e, \epsilon_c]\}$ for producing one additional unit to satisfy the random demand $y(p,\epsilon_d)$ at price p. Moreover, $\tilde{\eta}$ is the elasticity of the expected demand $\overline{y}(p)$, which differs from the expected elasticity of demand $\overline{\eta} \equiv E_\epsilon[-py_1(p,\epsilon_d)/y(p,\epsilon_d)]$. On the other hand, $\overline{cs}_e$ 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. In particular, when $\theta = \underline{\theta}$ so that $F(\theta) = 0$, (13) gives $\psi'(e) = \overline{cs}_e$ 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 $\widetilde{mc} = \overline{mc}$ and $\tilde{\eta} = \overline{\eta}$ so that (12) and (13) reduce to the FOC in Laffont and Tirole (1986) with a constant marginal cost function with additive random shock $(\theta - e)y + \epsilon_c$ considered there.

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

Hereafter, we assume that the cost function is multiplicatively separable in $\theta - e$ (see Laffont and Tirole (1993, p.171).

Assumption A1: *The random cost function is of the form*

$$c(y, \theta - e, \epsilon_c) = H(\theta - e)\ c_o(y, \epsilon_c), \tag{15}$$

for some functions $H(\cdot) \geq 0$ *and* $c_o(\cdot, \cdot) \geq 0$.

The function $c_o(\cdot, \cdot)$ can be viewed as the (random) *base cost*, while $H \equiv H(\theta - e)$ can be interpreted as the *cost inefficiency level* of the firm. Let the *expected base cost* for satisfying the random demand $y(p, \epsilon_d)$ at price p be $\overline{c}_o(p) \equiv E_\epsilon\{c_o[y(p, \epsilon_d), \epsilon_c]\}$.

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

$$t^*(\tilde{\theta}, c) \quad = \quad A(\tilde{\theta}) - \frac{\psi'[e^*(\tilde{\theta})]}{H'[\tilde{\theta} - e^*(\tilde{\theta})]} \left\{ \frac{c}{\overline{c}_o[p^*(\tilde{\theta})]} - H[\tilde{\theta} - e^*(\tilde{\theta})] \right\}, \tag{16}$$

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

$$A(\tilde{\theta}) = \psi[e^*(\tilde{\theta})] + \int_{\tilde{\theta}}^{\overline{\theta}} \psi'[e^*(\beta)]\ d\beta. \tag{17}$$

Thus, given the price schedule $p^*(\cdot)$ *and the transfer function* $t^*(\cdot, \cdot)$*, 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 discrpancy between the realized cost and the expected cost. In particular, (16) can be viewed as a menu of linear cost-reimbursement rules in realized cost c with slopes and intercepts depending on the firm's announcement $\tilde{\theta}$. Moreover, when $\theta = \underline{\theta}$, (13) and (15) imply that $\psi'[e^*(\underline{\theta})] = H'[\underline{\theta} - e^*(\underline{\theta})]\overline{c}_o[p^*(\underline{\theta})]$ 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, which announces its true type $\underline{\theta}$, chooses a fixed-price contract.

2.5. SECOND-ORDER CONDITIONS

Up to now, we have considered only the first-order conditions (FOC). In this subsection we verify that our optimal solution corresponds to a global maximum. In particular, it is fundamental to verify that announcing the true type θ holds not only locally but globally. As usual, we do this ex post by verifying that our solution satisfies the second-order conditions (SOC) for a local maximum, and that these SOC extend globally.

First, we make explicit our assumptions on the demand, cost and effort functions. Let $V(p, \epsilon_d)$ be the social value of producing the quantity demanded at price p given demand shock ϵ_d

$$V(p, \epsilon_d) = \int_p^\infty y(v, \epsilon_d)dv + (1 + \lambda)py(p, \epsilon_d),$$

i.e., $V(p, \epsilon_d)$ is the sum of the net consumer surplus and the revenue for the regulator computed at the shadow cost of public funds (see Laffont and Tirole (1993, p.132) when the good is private). Let the expected social value be

$$\overline{V}(p) \equiv \int V(p, \epsilon_d)dG(\epsilon_d) = \int_p^\infty \overline{y}(v)dv + (1 + \lambda)p\overline{y}(p),$$

where we have used the definition of $\overline{y}(p)$.

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

Assumption A2-(i) is standard (see Laffont and Tirole (1986, 1993)). Since

$$\begin{aligned} \overline{V}'(p) &= \lambda\overline{y}(p) + (1 + \lambda)p\overline{y}'(p) &&= (1 + \lambda)\overline{y}(p)\left(\frac{\lambda}{1 + \lambda} - \tilde{\eta}(p)\right) \\ \overline{V}''(p) &= (1 + 2\lambda)\overline{y}'(p) + (1 + \lambda)p\overline{y}''(p) &&= (1 + \lambda)p\overline{y}'(p)\left(\frac{1 + 2\lambda}{(1 + \lambda)p} + \frac{\overline{y}''(p)}{\overline{y}'(p)}\right), \end{aligned}$$

it follows that 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 curved, i.e. $-\overline{y}''(p)/\overline{y}'(p) < (1 + 2\lambda)/[(1 + \lambda)p]$ when $\overline{y}(p) > 0$ and $\overline{y}'(p) < 0$ as expected. Regarding assumption A2-(ii), the definition of the expected base cost $\overline{c}_o(p) = E_\epsilon\{c_o[y(p, \epsilon_d), \epsilon_c)]\}$ gives

$$\begin{aligned} \overline{c}_o'(p) &= E_\epsilon\{c_{o,1}[y(p, \epsilon_d), \epsilon_c)]y_1(p, \epsilon_d)\} \\ \overline{c}_o''(p) &= E_\epsilon\{c_{o,11}[y(p, \epsilon_d), \epsilon_c)]y_1^2(p, \epsilon_d)\} + E_\epsilon\{c_{o,1}[y(p, \epsilon_d), \epsilon_c)]y_{11}(p, \epsilon_d)\}. \end{aligned}$$

Thus, assumption A2-(ii) is satisfied if $c_{o,1}(\cdot,\cdot) > 0$, $c_{o,11}(\cdot,\cdot) \geq 0$, $y_1(\cdot,\cdot) < 0$ and $y_{11}(\cdot,\cdot) \geq 0$, i.e. if the base cost function is strictly increasing and convex in quantity and demand is strictly decreasing and convex in price. The latter conditions are satisfied in general. Assumption A2-(iii,iv) follows Laffont and Tirole (1993, p. 171). In particular, the cost inefficiency level is strictly increasing and convex in $\theta - e$, while the effort cost is strictly increasing and strcitly convex in e.

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

Lemma 1: *Suppose that the transfer function $t(\cdot,\cdot)$ is weakly decreasing and concave in realized cost c. Given assumptions 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, $0 \leq e_2(\theta, \theta) < 1$.*

Note that $t^*(\cdot,\cdot)$ is weakly decreasing and concave in realized cost c, as it is linear in c with slope $-\psi'[e^*(\tilde{\theta})]/\{H'[\tilde{\theta} - e^*(\tilde{\theta})]\overline{c}_o[p^*(\tilde{\theta})]\} < 0$. Thus, Lemma 1 applies.

Next, we turn to the incentive constraint (5). The local SOC for $\tilde{\theta} = \theta$ to be a local maximum is $U_{11}(\theta, \theta) \leq 0$, where $U(\tilde{\theta}, \theta)$ is given by (4). As is well known, using the FOC: $U_1(\theta, \theta) = 0$ which must hold for any θ, this SOC is equivalent to $U_{12}(\theta, \theta) \geq 0$. But differentiating (4) and using (1) give

$$U_2(\tilde{\theta}, \theta) = E_\epsilon\Big\{t_2(\cdot)c_2(\cdot)[1 - e_2(\tilde{\theta}, \theta)]\Big\} - \psi'[e(\tilde{\theta}, \theta)]e_2(\tilde{\theta}, \theta) = -\psi'[e(\tilde{\theta}, \theta)],$$

where the second equality follows from (3) where $e = e(\tilde{\theta}, \theta)$. Hence

$$U_{12}(\tilde{\theta}, \theta) = -\psi''[e(\tilde{\theta}, \theta)]e_1(\tilde{\theta}, \theta). \tag{18}$$

Because $\psi''(\cdot) > 0$, the local SOC: $U_{12}(\theta, \theta) \geq 0$ is equivalent to $e_1(\theta, \theta) \leq 0$, i.e.

$$e'(\theta) \leq e_2(\theta, \theta), \tag{19}$$

since $e(\theta) = e(\theta, \theta)$ implies $e'(\theta) = e_1(\theta, \theta) + e_2(\theta, \theta)$.

When the transfer function $t(\cdot,\cdot)$ is weakly decreasing and concave in realized cost, as is the case for the linearly decreasing transfer $t^*(\cdot,\cdot)$ given by (16), Lemma 1 implies that a *sufficient* condition for the local SOC (19) 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) $\psi''[e^*(\theta)]\overline{V}''[p^*(\theta)] + (1+\lambda)\Big\{H'[\theta - e^*(\theta)]\overline{c}'_o[p^*(\theta)]\Big\}^2 < 0$
(ii) $[(1+\lambda)/\lambda]H''[\theta - e^*(\theta)]\overline{c}_o[p^*(\theta)]/\psi''[e^*(\theta)] \leq d[F(\theta)/f(\theta)]/d\theta.$

Condition A3-(i) is reminiscent of assumption 1-(iii) in Laffont and Tirole (1986) for the case where $c(y, \theta - e, \epsilon_c) = (\theta - e)y + \epsilon_c$ and y is nonrandom. Condition A3-(ii) is slightly stronger than the usual condition that $F(\cdot)$ is log-concave as in Laffont and Tirole (1993, assumption 1.2). It actually reduces to it when $H(\theta - e) = \theta - e$ so that $H''(\cdot) = 0$.

Lemma 2: *Given assumptions A1–A3 and the transfer* $t^*(\cdot,\cdot)$ *and price* $p^*(\cdot)$ *functions, the local SOC (19) for truth telling is satisfied as* $e^{*\prime}(\cdot) < 0$. *Moreover,* $p^{*\prime}(\cdot) > 0$.

In particular, effort $e(\cdot)$ is strictly decreasing in firm's type θ, while price $p(\cdot)$ is strictly increasing in firm's type. These agree with the fact that the cost inefficiency level $H^* \equiv H[\theta - e^*(\theta)]$ of the firm is strictly increasing with its type θ because $\theta - e^*(\theta)$ is strictly increasing in θ and $H'(\cdot) > 0$.

It remains to show that $\tilde{\theta} = \theta$ provides a *global* maximum of the firm's utility (4) under the optimal transfer (16). This is accomplished by the next result.

Proposition 3: *Given assumptions A1–A3 and the transfer* $t^*(\cdot,\cdot)$ *and price* $p^*(\cdot)$ *functions, truth telling provides the global maximum of the expected utility function* $U(\tilde{\theta}, \theta)$ *given in (4). Moreover, the expected transfer* $\overline{t} \equiv E_\epsilon\Big[t^*\{\theta, c[y(p^*(\theta), \epsilon_d), \theta - e^*(\theta), \epsilon_c]\}\Big]$ *for a firm with type* θ *announcing its true type* θ *and thus exerting the optimal effort* $e^*(\theta)$ *is strictly decreasing and convex in the firm's cost inefficiency level* H^*.

In particular, the second part of Proposition 3 ensures that the regulator can use a menu of linear cost-reimbursement rules, as was proposed in subsection 2.4 on implementation. Moreover, because the firm's cost inefficiency level H^* is strictly increasing in firm's type, the expected transfer is strictly decreasing in firm's type, as expected.

3 Nonparametric Identification

The structural approach relies on the maintained assumption that the regulator offers the optimal price schedule $p^*(\cdot)$ and optimal transfer function $t^*(\cdot,\cdot)$ to the monopolist who then reveals its true type θ. The incentive regulation model of Section 2 then determines the price $p = p^*(\theta)$ per unit of private good, the effort $e = e^*(\theta)$ exerted by the monopolist, the quantity $y = y(p, \epsilon_d)$ of private good given the realized demand shock ϵ_d, the cost $c = c(y, \theta - e, \epsilon_c)$ for producing y given the realized cost shock ϵ_c, as well as the (net) transfer $t = t^*(\theta, c)$ to the monopolist. 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, which are the demand, base cost, cost efficiency and effort disutility functions, the distributions of the firm's type, demand shock and cost shock, as well as the shadow cost of public fund from the distribution of the observables. Throughout, assumptions A1–A3 are maintained.

3.1. The Structural Econometric Model

A number of complications arise. First, the effort exerted by the monopolist is typically unobserved as are the firm's type θ and the demand and cost shocks (ϵ_d, ϵ_c). Hereafter, we thus assume that only (Y, C, P, T) are observed, where we use capital letters to distinguish random variables from their realizations.

Second, the demand, cost and effort disutility functions typically depend on a vector of exogenous variables $Z \in I\!R^d$, where Z includes some characteristics of the firm, regulator and/or market. To allow for such dependencies, the demand, cost and effort disutility functions are defined hereafter as $y(p, z, \epsilon_d)$, $H(\theta - e, z)c_o(y, z, \epsilon_c)$ and $\psi(e, z)$ when $Z = z$. Similarly, the firm's type θ, the demand shocks ϵ_d and the cost shock ϵ_c may depend on 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 fund λ may depend on z, i.e. $\lambda = \lambda(z)$ for some nonnegative function $\lambda(\cdot)$.[20] From such dependencies on z, it follows that the optimal price, transfer and 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

---

[20] Clearly, not all of the variables in Z may affect these functions and distributions. For instance, some demand shifters may affect only the demand. The cost function may depend on some firm's specific characteristics, while the cost of public funds may depend on economic activity and the political environment. Though possibly helpful, exclusion restrictions are not exploited in this paper to achieve nonparametric identification.

let $\mathcal{Z}$ denote the support of the distribution of Z.

Third, 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). Indeed, 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 avoid such a difficulty, it is thus necessary to introduce another source of randomness.

There are two simple ways to do so. A first method, which is compatible with the structural approach, is to assume that some heterogeneity entering the effort disutility function but not the demand and cost functions is unobserved by the econometrician. This is reasonable as the determinants of the effort disutility function are likely to be less known that the determinants of the demand and cost functions. Assuming that such unobserved heterogeneity in the effort disutility function can be summarized by an additive term ϵ_t so that the effort disutility function is now $\psi(\cdot, z) + \epsilon_t$ instead of $\psi(\cdot, z)$, it can be seen from the form of (16) that the transfer $T = t^*(\theta, C, Z)$ will include ϵ_t as an additive term. Alternatively, a second method for introducing another source of randomness, which is less structural but retains the main content of the theoretical model of Section 2 is to consider that the observed transfer T differs from the optimal transfer $T^* = t^*(\theta, C, Z)$ by an additive random term ϵ_t. Such a random term ϵ_t may arise from measuring T^* with error, as data on transfers are likely to be imprecise. The random term ϵ_t may also represent extra transfers from the regulator to the firm that do not depend on cost efficiency considerations. In particular, the second approach is useful when one believes that the observed transfer is not equal to the optimal transfer. In this case, the Laffont–Tirole (1986) model can be viewed as a source for the observed transfer, while ϵ_t can then be used to assed deviations from the theoretical incentive regulation model.

Collecting the preceding remarks, rearranging (12) and(13), and combining (16) and (17) evaluated at $\tilde{\theta} = \theta$, the structural econometric model for the endogenous variables (Y, C, P, T) given the exogenous variables Z is defined, under assumptions A1–A3, by the nonlinear nonparametric simultaneous equation model generally implicit in P with nonadditive error terms $(\theta, \epsilon_d, \epsilon_c)$

$$\begin{aligned} Y &= y(P, Z, \epsilon_d) && (20)\\ C &= H(\theta - e, Z) c_o(Y, Z, \epsilon_c) && (21)\\ P\overline{y}'(P, Z) + \frac{\lambda}{1+\lambda}\overline{y}(P, Z) &= H(\theta - e, Z)\overline{c}_o'(P, Z) && (22)\\ \psi'(e, Z) + \frac{\lambda}{1+\lambda}\frac{F(\theta|Z)}{f(\theta|Z)}\psi''(e, Z) &= H'(\theta - e, Z)\overline{c}_o(P, Z) && (23)\\ T &= \psi(e, Z) + \int_{\theta}^{\overline{\theta}(Z)} \psi'[e^*(\tilde{\theta}, Z), Z] d\tilde{\theta} \\ &\quad - \frac{\psi'(e, Z)}{H'(\theta - e, Z)}\left\{\frac{C}{\overline{c}_o(P, Z)} - H(\theta - e, Z)\right\} + \epsilon_t, && (24) \end{aligned}$$

where $\lambda = \lambda(Z)$, $P = p^*(\theta, Z)$ and $e = e^*(\theta, Z)$ solve (22)–(23), and a prime denotes a derivation with respect to the first argument of a function. Following Section 2, $\overline{y}(p, z)$ and $\overline{c}_o(p, z)$ in (21)–(24) are, conditional upon $Z = z$, the expected demand at price p and the expected base cost for producing the random quantity $y(p, \epsilon_d)$ at price p, i.e.

$$\begin{aligned} \overline{y}(p, z) &= \int y(p, z, \epsilon_d) dG(\epsilon_d | z) && (25)\\ \overline{c}_o(p, z) &= \int c_o[y(p, z, \epsilon_d), z, \epsilon_c] dG(\epsilon_d, \epsilon_c | z). && (26) \end{aligned}$$

To complete the specification of the econometric model, we make the following assumption on the random elements $(\theta, \epsilon_d, \epsilon_c, \epsilon_t)$.[21]

Assumption B1: *θ is independent of $(\epsilon_d, \epsilon_c, \epsilon_t)$ conditional upon Z with $E[\epsilon_t|Z] = 0$.*[22]

The condition $E[\epsilon_t|Z] = 0$ is a normalization. When ϵ_t arises from some heterogeneity in the effort disutility function that is unobserved by the econometrician but known by the regulator and the firm, then (ϵ_d, ϵ_c) must be independent of θ given (Z, ϵ_t) for the theoretical model of Section 2 to apply. Assumption B1 then holds under the normalization $E[\epsilon_t|Z] = 0$ if ϵ_t and θ are independent conditional upon Z, i.e. if the firm's type does

---

[21] While estimating the Baron and Myerson (1982) model, Wolak (1994) considers so called "econometrician" error terms in the FOC of the regulator's maximization problem, i.e. in (22) and (23). In contrast, our error terms arise naturally from the model following the structural approach.

[22] A weaker requirement than assumption B1 is that θ is independent of (ϵ_d, ϵ_c) conditional upon Z, and $E[\epsilon_t|\theta, Z] = 0$ for every value z of Z. Indeed, only the condition $E[\epsilon_t|\theta, Z] = 0$ is used in establishing (B.6) in Appendix B. For simplicity, we consider Assumption B1.

not depend on the unobserved heterogeneity conditional on the observed heterogeneity Z. Alternatively, when ϵ_t is interpreted as a random term directly added to the optimal transfer due to measurement errors and/or extra cost-unrelated transfers, it is reasonable to assume that ϵ_t is independent of θ conditional upon $(Z, \epsilon_d, \epsilon_c)$. Because (ϵ_d, ϵ_c) is independent of θ conditional upon Z for the theoretical model of Section 2 to apply, it follows that θ is independent of $(\epsilon_d, \epsilon_c, \epsilon_t)$ conditional upon Z and hence that assumption B1 is satisfied under the normalization $E[\epsilon_t|Z] = 0$.

To summarize, the observables are (Y, C, P, T, Z), where the endogenous variables (Y, C, P, T) are determined by (21-25), while $(e, \theta, \epsilon_d, \epsilon_c, \epsilon_t)$ are unobserved. The structural elements of the model are the cost-of-public-fund function $\lambda(\cdot)$, the demand function $y(\cdot, \cdot, \cdot)$, the base cost function $c_o(\cdot, \cdot, \cdot)$ with its cost inefficiency function $H(\cdot, \cdot)$, the effort disutility function $\psi(\cdot, \cdot)$, the conditional distribution $F(\cdot|\cdot)$ of type given Z, and the joint distribution $G(\cdot, \cdot, \cdot|\cdot)$ of the random terms $(\epsilon_d, \epsilon_c, \epsilon_t)$ conditional upon Z. 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 parametric and nonparametric contexts, see e.g. Koopmans (1949), Roehrig (1988) and Prakasa Rao (1992). In subsections 3.2–3.4 we study such a problem for the simpler case where the cost inefficiency function $H(\cdot, z)$ is the identity function so that $H(\theta - e, z)$ and $H'(\theta - e, z)$ are replaced by $(\theta - e)$ and 1, respectively, in (21)–(24). The resulting model is called the *basic model*. The general model is then studied in subsection 3.5.

3.2. NONPARAMETRIC IDENTIFICATION OF $\psi(\cdot, \cdot)$ AND $F(\cdot|\cdot)$ GIVEN $\lambda(\cdot)$

In this subsection we consider the basic model and study the nonparametric identification of the effort disutility function $\psi(\cdot, \cdot)$ and the conditional distribution of firm's type $F(\cdot|\cdot)$ assuming that the public fund cost function $\lambda(\cdot)$ is known. Identification of $\lambda(\cdot)$ is addressed in subsection 3.4.

To begin, it must be noted that a *location–scale normalization* is necessary despite the restriction that the cost inefficiency function $H(\cdot, \cdot)$ is the identity function in the basic model. Intuitively, this arises because θ, the base cost function $c_0(\cdot, \cdot, \cdot)$ and the effort disutility function are unknown. The next result formalizes the necessity of a location-scale normalization.

Lemma 3: *Let $\alpha \geq 0$ and $\beta > 0$ be some functions $\alpha(\cdot)$ and $\beta(\cdot)$ of Z. Consider the* *two structures* $\mathcal{S} \equiv [y, c_o, \psi, F, G, \lambda]$ *and* $\tilde{\mathcal{S}} \equiv [\tilde{y}, \tilde{c}_o, \tilde{\psi}, \tilde{F}, \tilde{G}, \tilde{\lambda}]$ *in the basic model with assumptions 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* $\tilde{\lambda}(\cdot) = \lambda(\cdot)$*. Then, the structures* $\mathcal{S}$ *and* $\tilde{\mathcal{S}}$ *lead to the same conditional distribution of* (Y, C, P, T) *given* Z*, i.e. the structures* $\mathcal{S}$ *and* $\tilde{\mathcal{S}}$ *are observationally equivalent.*

As the proof of Lemma 3 indicates, the observational equivalence between $\mathcal{S}$ and $\tilde{\mathcal{S}}$ 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. Many location-scale normalizations can be employed. For instance, one can fix how two quantiles of θ varies with z. A natural choice for these quantiles are $\underline{\theta}(z)$ and $\overline{\theta}(z)$, which correspond to the most and less efficient firms, 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) in the basic model to ensure that $C = [\theta - e^*(\theta, Z)]c_o(Y, Z, \epsilon_c) \geq 0$.

A more convenient location-scale normalization, which is used hereafter, 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 0, irrespective of the value of Z. Formally, we impose

Assumption B2: For every value z of Z

$$\underline{\theta}(z) - e^*[\underline{\theta}(z), z] = 1 \qquad \text{and} \qquad e^*[\overline{\theta}(z), z] = 0. \tag{27}$$

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 (27) actually determines $\underline{\theta}(z)$ and $\overline{\theta}(z)$ as in the preceding direct location-scale normalization, though (27) fixes those boundaries endogenously through the optimal effort function $e^*(\cdot, z)$. Moreover, the normalization (27) is much convenient as it ensures that $\theta - e^*(\theta, z)$ and $e^*(\theta, z)$ of the firm with type θ are both always nonnegative, as desired. Then, from (21) with $H(x, z) = x$, it follows that $c_o(y, z, \epsilon_c)$ can be interpreted as the *cost frontier* for producing y given (z, ϵ_c), while $\theta - e = \theta - e^*(\theta, z) \geq \underline{\theta}(z) - e^*[\underline{\theta}(z), z] = 1$ can be viewed as the *relative cost inefficiency* of a firm with type θ relative to the cost efficient firm with type $\underline{\theta}(z)$.

We now turn to the nonparametric identification of the effort disutility function $\psi(\cdot,\cdot)$ and the conditional distribution of type $F(\cdot|\cdot)$. We need a preliminary result, which establishes that the expected demand function $\overline{y}(\cdot,\cdot)$ and the expected base cost $\overline{c}_o(\cdot,\cdot)$ are identified nonparametrically from observations on quantity, price and costs given $\lambda(\cdot)$. Moreover, it is shown that the relative cost inefficiency of the firm can be recovered uniquely from the 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 $\lambda(\cdot)$ is known in the basic model with assumptions A1–A3 and B1–B2. Thus the expected demand function $\overline{y}(\cdot,\cdot)$ and the expected base cost $\overline{c}_o(\cdot,\cdot)$ are uniquely determined by $\underline{p}(\cdot)$ and the conditional means of (Y,C) given (P,Z) as*

$$\begin{aligned}\overline{y}(p,z) &= E[Y|P\!=\!p,Z\!=\!z] && (28)\\ \overline{c}_o(p,z) &= E[C|P\!=\!\underline{p}(z),Z\!=\!z]\exp\left\{\int_{\underline{p}(z)}^{p}\frac{\tilde{p}\overline{y}'(\tilde{p},z)+\mu\overline{y}(\tilde{p},z)}{E[C|P\!=\!\tilde{p},Z\!=\!z]}d\tilde{p}\right\}, && (29)\end{aligned}$$

where $\mu=\mu(z)\equiv\lambda(z)/[1+\lambda(z)]$. Moreover, the relative cost inefficiency is

$$\theta-e^*(\theta,z)=\Delta(p,z)\equiv E[C|P\!=\!p,Z\!=\!z]/\overline{c}_o(p,z), \tag{30}$$

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 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 Z in the demand equation (20) under assumption B1, as the latter only ensures that ϵ_d is independent of θ and hence of P given Z.[23] 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 base cost (26).[24] Nevertheless, by exploiting the generalized Ramsey pricing rule (22), Lemma 4 indicates

---

[23] For instance, consider a demand that is additively separable in ϵ_d, i.e. $Y=r(P,Z)+\epsilon_d$. Then, the regression of Y on (P,Z) recovers the expected demand $\overline{y}(p,z)$ at price p despite the possible correlation between ϵ_d and Z. This is so because $r(p,z)\neq\overline{y}(p,z)=r(p,z)+E[\epsilon_d|Z]=E[Y|P=p,Z=z]$.

[24] For instance, consider a typical cost specification of the form $\log C=\log[c_o(Y,Z,\epsilon_c)]+\log(\theta-e)=s(Y,Z)+\epsilon_c+\log(\theta-e)$, where $\log(\theta-e)\geq 0$ in view of (27). The composite error term $\epsilon_c+\log(\theta-e)$ is typically correlated with both $Y=y(P,Z,\epsilon_d)$ and Z under Assumption B1. Regression, IV estimation, and ML estimation of this model attempts to estimate the cost frontier $s(y,z)$, which is different from the expected base cost $\overline{c}_o(p,z)$ that is relevant in the FOC (22)–(23) of price and effort.

that the expected base cost $\overline{c}_o(\cdot,\cdot)$, which is the expected base cost for the most efficient firm given the normalization (27), can be estimated from (29) by combining appropriately the regressions of Y and C on (P,Z) with the knowledge of $\underline{p}(\cdot)$. Moreover, (30) shows that the relative cost inefficiency $\theta - e^*(\theta,z)$ of a firm can be recovered from the firm's individual values (y,p,c,z) as the function $\Delta(\cdot,\cdot)$ is known from the regression of C given (P,Z) and the expected base cost $\overline{c}_o(\cdot,\cdot)$.

Using Lemma 4, the next result establishes the nonparametric identification of the effort disutility function $\psi(\cdot,\cdot)$ and the conditional distribution of type $F(\cdot|\cdot)$ from observations on quantity, price, costs and transfers given $\lambda(\cdot)$. To this end, define the functions

$$\Gamma(p,z) \;=\; -\frac{\partial E[T|P\!=\!p, Z\!=\!z]/\partial p}{\partial \Delta(p,z)/\partial p} \tag{31}$$

$$R(p,z) \;=\; \frac{\mu[G_{P|Z}(p|z)/g_{P|Z}(p|z)] \times \partial\Gamma(p,z)/\partial p \times \partial\Delta(p,z)/\partial p}{\Gamma(p,z) - \overline{c}_0(p,z) + \mu[G_{P|Z}(p|z)/g_{P|Z}(p|z)] \times \partial\Gamma(p,z)/\partial p}, \tag{32}$$

where $G_{P|Z}(\cdot|\cdot)$ and $g_{P|Z}(\cdot|\cdot)$ denote the conditional distribution and density of P given Z. In particular, note that 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. As seen in the proof, the functions $\Gamma(\cdot,\cdot)$ and $R(\cdot,\cdot)$ exploit the expected optimal transfer from (24) and the FOC for optimal effort (23).

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

$$\psi(e,z) \;=\; E[T|P\!=\!\overline{p}(z), Z\!=\!z] + \int_0^e \Gamma[p^*(\tilde{e},z),z]\; d\tilde{e}, \tag{33}$$

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

$$e^*(p,z) = \int_p^{\overline{p}(z)} R(\tilde{p},z)\; d\tilde{p}, \tag{34}$$

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 distribution $F(\cdot|z)$ of type given $Z=z$ as the distribution of

$$\theta = \theta^*(P,z) \equiv \Delta(P,z) + \int_P^{\overline{p}(z)} R(\tilde{p},z)\; d\tilde{p}, \tag{35}$$

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

In particular, while the minimal effort $e^*[\overline{\theta}(z), z] = 0$ by the normalization (27), (34) implies that the maximal effort exerted by the efficient firm with type $\underline{\theta}(z)$ is

$$e^*[\underline{\theta}(z), z] = \int_{\underline{p}(z)}^{\overline{p}(z)} R(\tilde{p}, z)\ d\tilde{p} > 0. \tag{36}$$

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

$$\begin{aligned}\underline{\theta}(z) &= 1 + \int_{\underline{p}(z)}^{\overline{p}(z)} R(\tilde{p}, z)\ d\tilde{p} > 1 && (37)\\ \overline{\theta}(z) &= \frac{E[C|P{=}\overline{p}(z), Z{=}z]}{E[C|P{=}\underline{p}(z), Z{=}z]} \exp\left\{-\int_{\underline{p}(z)}^{\overline{p}(z)} \frac{\tilde{p}\overline{y}'(\tilde{p}, z) + \mu\overline{y}(\tilde{p}, z)}{E[C|P{=}\tilde{p}, Z{=}z]} d\tilde{p}\right\} > \underline{\theta}(z), && (38)\end{aligned}$$

from (35) in view of (29)–(30) and $\Delta[\underline{p}(z), z] = 1$.

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). Moreover, such a property is analogous to that used in Guerre, Perrigne and Vuong (2000) where observed bids and unobserved private values play the role of observed prices and unobserved types. Specifically, (35) is analogous to the fundamental equation (3) in that paper, which expresses any unobserved private value in terms of the corresponding observed bid and functions that are identified from observables. Though more complicated, (35) expresses the unobserved firm's type in terms of the observed price and functions that are identified from observables. Hence, as in that paper, the unobserved firm's type θ can be recovered uniquely through (35) from the firm's observed price p and characteristics z once the various unknown functions have been estimated from data on (Y, C, P, T, Z). That is, (35) 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 (34) from the firm's observed value (p, z).

3.3. NONPARAMETRIC IDENTIFICATION OF $y(\cdot,\cdot,\cdot)$, $c_o(\cdot,\cdot,\cdot)$ AND $G(\cdot,\cdot,\cdot|\cdot)$ GIVEN $\lambda(\cdot)$

Lemma 3 establishes the identification of the expected demand and expected base cost functions $\overline{y}(\cdot)$ and $\overline{c}_o(\cdot)$ at every price p. For counterfactual exercises or policy evaluations, one may need to identify the remaining structural elements of the model, which are the demand function $y(\cdot,\cdot,\cdot)$, the base cost function $c_o(\cdot,\cdot,\cdot)$ and the conditional distribution

$G(\cdot,\cdot,\cdot|\cdot)$ of $(\epsilon_d,\epsilon_c,\epsilon_t)$ given Z. This is the purpose of the present subsection, which still assumes that the public fund cost function $\lambda(\cdot)$ is known.

Unlike the random term ϵ_t, which enters additively in the transfer equation (24), the demand shock ϵ_d and cost shock ϵ_c do not enter additively in the demand equation (20) and cost equation (21). In this, we follow Matzkin (2003) who argues that the structural specification of a random demand or a cost function seldom leads to an additive random term as several references given in that paper indicate. When the random term does not enter additively into the relationship between the endogoneous variable and the exogenous variables, Matzkin (1993) shows that the model is nonidentified nonparametrically and that some normalization is needed. Several normalizations can be entertained.[25] Hereafter, we use Matzkin (2003) first normalization, though we could have used the other two normalizations studied in her paper, which exploit some restrictions imposed by economic theory (see also Matzkin (1994)).

Let $[\underline{\epsilon}_d(z),\overline{\epsilon}_d(z)]$ be the support of the conditional distribution $G_{\epsilon_d|Z}(\cdot|z)$ of ϵ_d 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 natural strict monotonocity conditions on the demand and cost shocks (ϵ_d,ϵ_c).

Assumption B3: *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*

$$y[p_o(z),z,\epsilon_d]=\epsilon_d \qquad \text{and} \qquad c_o[y_o(z),z,\epsilon_c]=\epsilon_c \tag{39}$$

where $p_o(\cdot)$ and $y_o(\cdot)$ are known. Moreover, the demand and base cost functions $y(p,z,\cdot)$ and $c_o(y,z,\cdot)$ are strictly increasing in ϵ_d and ϵ_c, respectively, for all values (y,p,z), while

---

[25] For instance, consider the demand equation (20), namely $Y=y(P,Z,\epsilon_d)$, where $P=p^*(\theta,Z)$ is independent of ϵ_d given Z in view of Assumption B1. Clearly, the model is nonparametrically identified as a monotonic transformation of the demand shock ϵ_d can be compensated by an appropriate transformation of the function $y(\cdot,\cdot,\cdot)$. Thus, an obvious normalization is to impose simply that the distribution of ϵ_d given Z is known and equal to $G^o_{\epsilon_d|Z}(\cdot|\cdot)$ say. With such a normalization, then $y(\cdot,\cdot,\cdot)$ is identified since $G^o_{\epsilon_d|Z}(\cdot|z)=G_{\epsilon_d|P,Z}(\cdot|p,z)=G_{Y|P,Z}[y(p,z,\cdot)|p,z]$ so that $y(p,z,\cdot)=G^{-1}_{Y|P,Z}[G^o_{\epsilon_d|Z}(\cdot|z)|p,z]$. In addition to requiring that the distribution of ϵ_d given Z be chosen, such a similar normalization does not seem to be useful to identify the cost equation (21) as Y is not independent of ϵ_c given Z.

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 strictly increasing in their first arguments.*

The first condition in (39) actually says that the demand Y when the price is $p_o(z)$ and $Z=z$ is the demand shock ϵ_d, while the second condition says that the base cost when the output is $y_o(z)$ and $Z=z$ is the cost shock ϵ_c. The first condition is exactly Matzkin (2003) normalization (see equation (2.5) in that paper), where we have chosen $p_o(z)$ as the reference point when $Z=z$. As indicated there, under the additively separable demand $y(p,z,\epsilon_d)=r(p,z)+\epsilon_d$, such a condition is satisfied if $r[p_o(z),z]=0$. Thus, a choice for $p_o(z)$ can be the highest price $\overline{p}(z)$ so that $r[\overline{p}(z),z]=0$ after possible location relabelling of $r(p,z)$ and ϵ_d as $r(p,z)-r[\overline{p}(z),z]$ and $\epsilon_d+r[\overline{p}(z),z]$, respectively. Note that $\overline{p}(\cdot)$ is known from the knowledge of $G_{P|Z}(\cdot|\cdot)$, as required. The second condition in (39) is similar, where the reference point is $y_o(z)$. For instance, under the additively separable base cost $c_o(y,z,\epsilon_c)=s(y,z)+\epsilon_c$, the second condition is satisfied if $s[y_o(z),z]=0$. A choice for $y_o(z)$ is the lowest output $\underline{y}[\overline{p}(z),z]=y[\overline{p}(z),z,\underline{\epsilon}_d(z)]$, after possible relabelling of $s(y,z)$ and ϵ_c through a location shift, as previously.

The next result establishes the nonparametric identification of the demand function $y(\cdot,\cdot,\cdot)$, the base cost function $c_o(\cdot,\cdot,\cdot)$ and the conditional distribution $G(\cdot,\cdot|\cdot)$ of (ϵ_d,ϵ_c) given Z from observations on quantity, price, costs given $\lambda(\cdot)$. Let $C_o=c_o(Y,Z,\epsilon_c)$ be the (random) *base cost*. Because $C=[\theta-e^*(\theta,Z)]C_o$, where C is observed and $\theta-e^*(\theta,Z)=\Delta(P,Z)$, which is identified by Lemma 4, it follows that the base cost C_o can be recovered and its conditional distribution $G_{C_o|Y,P,Z}(\cdot|\cdot,\cdot,\cdot)$ given (Y,P,Z) is identified. More formally, because $\Delta(P,Z)\geq 1>0$, 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). This information is used next.

Proposition 5: *Suppose that* $\lambda(\cdot)$ *is known in the basic model with assumptions A1–A3 and B1–B3.*
(i) The demand function $y(\cdot,\cdot,\cdot)$ *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*

$$y(p,z,\epsilon_d) \;=\; G^{-1}_{Y|P,Z}\left\{G_{Y|P,Z}[\epsilon_d|p_o(z),z]|p,z\right\} \tag{40}$$

$$G_{\epsilon_d|Z}(\cdot|z) \;=\; G_{Y|P,Z}[\;\cdot\;|p_o(z), z]. \tag{41}$$

(ii) Suppose that for all $z \in \mathcal{Z}$*, and all* $\epsilon_d \in [\underline{\epsilon}_d(z), \overline{\epsilon}_d(z)]$*, there exists* $p_\dagger(z,\epsilon_d) \in [\underline{p}(z), \overline{p}(z)]$ *such that* $y_o(z) = y[p_\dagger(z,\epsilon_d), z, \epsilon_d]$*. Then, the base cost function* $c_o(\cdot,\cdot,\cdot)$ *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*

$$c_o(y,z,\epsilon_c) \;=\; G^{-1}_{C_o|Y,P,Z}\left\{G_{C_o|Y,P,Z}[\epsilon_c|y_o(z), p_\dagger(z,\epsilon_d), z]|y,p,z\right\} \tag{42}$$

$$G_{\epsilon_c|\epsilon_d,Z}(\cdot|\epsilon_d,z) \;=\; G_{C_o|Y,P,Z}[\;\cdot\;|y_o(z), p_\dagger(z,\epsilon_d), z], \tag{43}$$

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

The proof of (i) follows Matzkin(2003). Despite the nonindependence of $Y = y(P,Z,\epsilon_d)$ and ϵ_c given Z in $C_o = c_o(Y,Z,\epsilon_c)$, the proof of (ii) is only slightly more involved as it exploits the additional condition in (ii). This condition says roughly that for any value of the demand shock 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 Assumption B3, when $z = z$. The additional condition in (ii) can be relaxed. In this case, following 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 $y_o(z) = y[p_\dagger(z,\epsilon_d), z, \epsilon_d]$.

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). Indeed, the demand and cost shocks (ϵ_d,ϵ_c) can be recovered from (Y,C,P,Z) through (20)–(21) as $y(\cdot,\cdot,\cdot)$ and $c_o(\cdot,\cdot,\cdot)$ are identified by Proposition 5. The identification of $G_{\epsilon_t|\epsilon_d,\epsilon_c,Z}(\cdot|\cdot,\cdot,\cdot)$ follows immediately from (24), since ϵ_t can be expressed as a function of (C,P,T,Z) and functions that are identified and thus estimable from observations on (Y,C,P,T,Z) by Lemma 4 and Proposition 4. For a simpler expression than (24), see Lemma 5 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.

3.4. NONPARAMETRIC IDENTIFICATION OF $\lambda(\cdot)$

The previous identification results can be used when the cost of public fund $\lambda(\cdot)$ is known.

In the US, $\lambda = 0.3$ is a well accepted value among economists, while the cost of public funds takes larger values in developing countries. On the other hand, in microeconometric studies, one may want to distinguish regulatory contracts according to the regulator and/or the market. In this case, 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 subsection.

To this end, we need a lemma that expresses the firm's rent and the received transfer directly from observations on (Y, C, P, T, Z) and identifiable functions.

Lemma 5: *In the basic model with assumptions A1–A3 and B1–B2, the firm's transfer can be written as*

$$T = E[T|P,Z] - \frac{\partial E[T|P,Z]/\partial p}{\partial E[C|P,Z]/\partial p - [P\overline{y}'(P,Z) + \mu\overline{y}(P,Z)]}\Big(C - E[C|P,Z]\Big) + \epsilon_t \quad (44)$$

where $E[\epsilon_t|P,Z] = 0$. *Moreover, the firm's expected rent is*

$$U^*(\theta) = E[T|P{=}p, Z{=}z] - E[T|P{=}\overline{p}(z), Z{=}z] - \int_p^{\overline{p}(z)} \Gamma(\tilde{p},z)R(\tilde{p},z)d\tilde{p} \ \geq \ 0 \qquad (45)$$

where $\theta = \theta^*(p,z)$ *as given by (35).*

Equation (44) is interesting for several reasons. First, note that the fraction in (44) is strictly positive as it is equal $\psi'(e,Z)/\overline{c}_o(P,Z)$. Thus, the ex post transfer T is, conditional on (P,Z), equal to its expectation $E[T|P,Z]$ plus a combined residual η_t, which is the difference between the random term ϵ_t and a positive fraction of the excess cost C over its expectation $E[C|P,Z]$. Moreover, because the firm's ex post rent is equal to ex post transfer minus the disutility of effort $\psi(e,z)$, while the expected rent is equal to the expected transfer minus $\psi(e,z)$, it follows that the ex post rent is equal to (45) plus the combined residual η_t. Second, as indicated in the previous subsection, when $\lambda(\cdot)$ and hence μ are known, (44) can be used to recover ϵ_t and hence establish the identification of $G_{\epsilon_t|\epsilon_d,\epsilon_c,Z}(\cdot|\cdot,\cdot,\cdot)$. Third, (44) suggests how μ and hence $\lambda(\cdot)$ can be identified.

We make the following identifying assumption.

Assumption B4: *C and ϵ_t are independent conditional upon (θ, Z).*[26]

---

[26] As a matter of fact, we only need that cost and residual transfer are uncorrelated given (θ, Z), i.e. $E[C\epsilon_t|\theta,Z] = 0$. In line with assumption B1, which can be weakened as indicated in footnote 22, we use again a conditional independence requirement that is stronger.

In particular, because $P = p^*(\theta, Z)$ and $e = e^*(\theta, Z)$, it follows from (20)-(21) that Assumption B4 is satisfied if (ϵ_d, ϵ_c) are independent of ϵ_t given (θ, Z). Alternative identifying assumptions can be made to identify $\lambda(\cdot)$. For instance, one could exploit the conditional independence of θ and ϵ_t given $(\epsilon_d, \epsilon_c, Z)$ that follows from assumption B1. We use assumption B4 because it leads to a simple way for identifying $\lambda(\cdot)$. Moreover, if we adopt the second method discussed in subsection 3.1 for introducing ϵ_t, then ϵ_t can be interpreted as the *residual transfer*. Because such residual transfer are unrelated to cost efficiency, it is natural to assume that C and ϵ_t are independent conditional upon (θ, Z) so that assumption B4 is satisfied.

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

Proposition 6: *In the basic model with assumptions A1–A3, B1–B2 and B4, the cost-of-public-fund function* $\lambda(\cdot)$ *is uniquely determined by* $\lambda(z) = \mu(z)/[1 - \mu(z)]$, *where*

$$\begin{aligned}\mu(z) \quad = \quad & \frac{1}{E[Y|P{=}p, Z{=}z]} \left\{ \frac{\partial E[T|P{=}p, Z{=}z]}{\partial p} \frac{Var[C|P{=}p]}{Cov[C, T|P{=}p, Z{=}z]} \right. \\ & \left. - \frac{\partial E[C|P{=}p, Z{=}z]}{\partial p} + p \frac{\partial E[Y|P{=}p, Z{=}z]}{\partial p} \right\} \qquad (46)\end{aligned}$$

with $Cov[C, T|P{=}p, Z{=}z] < 0$, *for any* $p \in [\underline{p}(z), \overline{p}(z)]$ *such that* $E[Y|P{=}p, Z{=}z] > 0$.

Because p can be chosen arbitrarily, (46) shows that $\mu(\cdot)$ and hence $\lambda(\cdot)$ are overidentified. Thus, weaker assumptions than assumption B4 can be exploited to achieve identification of the cost of public funds.

3.5. NONPARAMETRIC IDENTIFICATION OF THE GENERAL MODEL

To be completed.

4 Conclusion

To be completed.

Appendix A

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

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

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

where $p = p(\theta)$ and $e = e(\theta)$ are the control functions, $U(\theta)$ is the state variable, and $\mu(\theta)$ is the Lagrange multiplier associated with the constraint (9). Hence the FOC are:

$$\begin{aligned}\mathcal{H}_p &= \Big\{\lambda\overline{y}(p) + (1+\lambda)p\overline{y}'(p) - (1+\lambda)E_\epsilon\{c_1[y(p,\epsilon_d),\theta-e,\epsilon_c]y_1(p,\epsilon_d)\}\Big\}f(\theta) = 0 \\ \mathcal{H}_e &= \Big\{-(1+\lambda)\psi'(e) + (1+\lambda)E_\epsilon\{c_2[y(p,\epsilon_d),\theta-e,\epsilon_c]\}\Big\}f(\theta) - \mu(\theta)\psi''(e) = 0 \\ -\mathcal{H}_U &= \lambda f(\theta) = \mu'(\theta)\end{aligned}$$

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

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 assumption A1, this problem becomes

$$\begin{aligned}(F^*) \qquad \max_{\tilde{\theta},e} \; & E\Big\{t^*\Big(\tilde{\theta}, c[y(p^*(\tilde{\theta}),\epsilon_d),\theta-e,\epsilon_c]\Big) \mid \theta\Big\} - \psi(e) \\ &= E_\epsilon\Big\{t^*\Big(\tilde{\theta}, c[y(p^*(\tilde{\theta}),\epsilon_d),\theta-e,\epsilon_c]\Big)\Big\} - \psi(e) \\ &= A(\tilde{\theta}) + \frac{\psi'[e^*(\tilde{\theta})]}{H'[\tilde{\theta}-e^*(\tilde{\theta})]}\Big\{H[\tilde{\theta}-e^*(\tilde{\theta})] - H(\theta-e)\Big\} - \psi(e)\end{aligned}$$

where the first equality follows from the independence between θ and (ϵ_d,ϵ_c), while the second equality follows from (15) and (16). Thus, using (17) the FOC with respect to $\tilde{\theta}$ and e are respectively

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

$$= \left\{\frac{d}{d\tilde{\theta}}\left(\frac{\psi'[e^*(\tilde{\theta})]}{H'[\tilde{\theta}-e^*(\tilde{\theta})]}\right)\right\}\Big\{H[\tilde{\theta}-e^*(\tilde{\theta})]-H(\theta-e)\Big\}$$

$$0 = \frac{\psi'[e^*(\tilde{\theta})]}{H'[\tilde{\theta}-e^*(\tilde{\theta})]}H'(\theta-e)-\psi'(e)$$

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 (16), the right-hand side of (6) is

$$\begin{aligned} A(\theta)+\frac{\psi'[e^*(\theta)]}{H'[\theta-e^*(\theta)]}\{H[\theta-e^*(\theta)]-H[\theta-e^*(\theta)]\}-\psi[e^*(\theta)] &= A(\theta)-\psi[e^*(\theta)] \\ &= U^*(\theta) \end{aligned}$$

by (14) and (17), as desired. QED

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

$$\int U_{33}(\tilde{\theta},\theta,e,\epsilon_d,\epsilon_c)dG(\epsilon_d,\epsilon_c) = \int[t_{22}(\cdot)c_2^2(\cdot)+t_2(\cdot)c_{22}(\cdot)]dG(\epsilon_d,\epsilon_c)-\psi''(e)$$

where we have omitted the arguments of the functions to simplify the notation. But $c_{22}(\cdot) = H''(\theta-e)c_o[y(p(\tilde{\theta}),\epsilon_d),\epsilon_c] \geq 0$ by assumptions A1 and A2-(iii). 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

$$0=[1-e_2(\tilde{\theta},\theta)]E_\epsilon\Big[t_{22}(\cdot)c_2^2(\cdot)+t_2(\cdot)c_{22}(\cdot)\Big]+\psi''[e(\tilde{\theta},\theta)]e_2(\tilde{\theta},\theta)$$

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

$$e_2(\theta,\theta)\Big\{E_\epsilon\Big[t_{22}(\cdot)c_2^2(\cdot)+t_2(\cdot)c_{22}(\cdot)\Big]-\psi''[e(\theta)]\Big\}=E_\epsilon\Big[t_{22}(\cdot)c_2^2(\cdot)+t_2(\cdot)c_{22}(\cdot)\Big]$$

Under assumptions A1 and A2-(iii), we have $c_{22}(\cdot) \geq 0$ as noted above. 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 assumption A2-(iv), it follows that $0 \leq e_2(\theta,\theta) < 1$. QED

Proof of Lemma 2: As noted before assumption 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

$$\begin{aligned} p^*(\theta)\overline{y}'[p^*(\theta)] &= H[\theta - e^*(\theta)]\overline{c}_o'[p^*(\theta)] - \frac{\lambda}{1+\lambda}\overline{y}[p^*(\theta)] && \text{(A.1)} \\ \psi'[e^*(\theta)] &= H'[\theta - e^*(\theta)]\overline{c}_o[p^*(\theta)] - \frac{\lambda}{1+\lambda}\frac{F(\theta)}{f(\theta)}\psi''[e^*(\theta)] && \text{(A.2)} \end{aligned}$$

where we have used assumption A1, the definition of $\overline{c}_o(\cdot)$, and the expression for $\overline{c}_o'(\cdot)$ found earlier. Differentiating (20)-(21) with respect to θ and rearranging 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

$$\begin{aligned} A &= H'[\theta - e^*(\theta)]\overline{c}_o'[p^*(\theta)] \\ B &= \frac{1+2\lambda}{1+\lambda}\overline{y}'[p^*(\theta)] + p^*(\theta)\overline{y}''[p^*(\theta)] - H[\theta - e^*(\theta)]\overline{c}_o''[p^*(\theta)] \\ &= \frac{\overline{V}''[p^*(\theta)]}{1+\lambda} - H[\theta - e^*(\theta)]\overline{c}_o''[p^*(\theta)] \\ C &= \psi''[e^*(\theta)] + \frac{\lambda}{1+\lambda}\frac{F(\theta)}{f(\theta)}\psi'''[e^*(\theta)] + H''[\theta - e^*(\theta)]\overline{c}_o[p^*(\theta)] \\ D &= H''[\theta - e^*(\theta)]\overline{c}_o[p^*(\theta)] - \frac{\lambda}{1+\lambda}\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\psi''[e^*(\theta)] \end{aligned}$$

Under assumptions 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

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

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

$$-\psi''[e^*(\theta)] < \frac{\Big\{H'[\theta - e^*(\theta)]\overline{c}_o'[p^*(\theta)]\Big\}^2}{\frac{\overline{V}''[p^*(\theta)]}{1+\lambda} - H[\theta - e^*(\theta)]\overline{c}_o''[p^*(\theta)]}$$

which implies the first inequality in (A.5) by assumption A2. Condition A3-(ii) is equivalent to

$$H''[\theta - e^*(\theta)]\overline{c}_o[p^*(\theta)] \leq \frac{\lambda}{1+\lambda}\frac{d}{d\theta}\bigg(\frac{F(\theta)}{f(\theta)}\bigg)\,\psi''[e^*(\theta)]$$

which implies the second inequality in (A.5) because $B < 0$ and $H'[\theta - e^*(\theta)]\overline{c}^{o\prime}[p^*(\theta)] \neq 0$ under assumption A2.

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. QED

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

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

(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 $(\tilde{\theta}, \theta)$. From (18), 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 assumption A1 as

$$0 = \frac{\psi'[e^*(\tilde{\theta})]}{H'[\tilde{\theta} - e^*(\tilde{\theta})]}H'[\theta - e(\tilde{\theta}, \theta)] - \psi'[e(\tilde{\theta}, \theta)]$$

from the FOC of problem (F^*). Differentiating this FOC with respect to $\tilde{\theta}$ gives

$$e_1(\tilde{\theta}, \theta)\Big\{\psi''(\cdot) + \frac{\psi'(\cdot)}{H'(\cdot)}H''(\cdot)\Big\} = H'(\cdot)\Big\{\frac{\psi''(\cdot)e^{*\prime}(\cdot)}{H'(\cdot)} - \frac{\psi'(\cdot)H''(\cdot)[1 - e^{*\prime}(\cdot)]}{H'^2(\cdot)}\Big\}$$

Because $e^{*\prime}(\cdot) < 0$ by Lemma 2, it is easy to verify that the right-hand side is strictly negative while the term in braces is strictly positive under assumption 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 $U(\tilde{\theta}, \theta)$. This establishes that $\tilde{\theta} = \theta$ provides the global maximum of $U(\tilde{\theta}, \theta)$.

To prove the second part, let $\bar{t}(\theta) \equiv E_\epsilon\Big[t^*\{\theta, c[y(p^*(\theta), \epsilon_d), \theta - e^*(\theta), \epsilon_c]\}\Big]$ so that $\bar{t} = \bar{t}(\theta)$. Let $H_\dagger(\theta) \equiv H[\theta - e^*(\theta)] = E_\epsilon\{c[y(p^*(\theta), \epsilon_d), \theta - e^*(\theta), \epsilon_c]\}/\overline{c}_o[p^*(\theta)]$ so that the firm's cost inefficiency level H^* satisfies $H^* = H_\dagger(\theta)$. Note that $H_\dagger(\cdot)$ is strictly increasing in θ because $dH_\dagger/d\theta = [1 - e^{*\prime}(\theta)]H'[\theta - e^*(\theta)] > 0$ as $H'(\cdot) > 0$ and $e^{*\prime}(\cdot) < 0$. Thus $\theta = H_\dagger^{-1}(H^*)$. We want to show that $\bar{t}_\dagger(H^*) \equiv \bar{t}[H_\dagger^{-1}(H^*)]$ is strictly decreasing and convex in H^*. From (16) and assumption A1, we have $\bar{t}(\theta) = A(\theta)$. Hence, using (17)

$$\frac{d\bar{t}_\dagger}{dH^*} = \frac{A'(\theta)}{H_\dagger'(\theta)} = -\frac{\psi'[e^*(\theta)]}{H'[\theta - e^*(\theta)]}$$

which is strictly negative. Thus, the expected transfer is strictly decreasing in H^*, as desired. Moreover,

$$\frac{d^2\bar{t}_\dagger}{dH^{*2}} = -\frac{\psi''[e^*(\theta)]e^{*\prime}(\theta)}{H'^2[\theta - e^*(\theta)][1 - e^{*\prime}(\theta)]} + \frac{\psi'[e^*(\theta)]}{H'^2[\theta - e^*(\theta)]}H''[\theta - e^*(\theta)][1 - e^{*\prime}(\theta)]$$

It is easy to see that $d^2\bar{t}_\dagger/dH^{*2} > 0$ under assumption A2 because $e^{*\prime}(\theta) < 0$ by Lemma 2. Thus $\bar{t}_\dagger(\cdot)$ is strictly convex in H^*, as desired. QED

Appendix B

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

Proof of Lemma 3: Let $\tilde{Y}, \tilde{C}, \tilde{P}, \tilde{T}$ denote the endogenous variables under the structure $\tilde{\mathcal{S}}$. 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 $\tilde{y}(\cdot,\cdot,\cdot) = y(\cdot,\cdot,\cdot)$ and $\tilde{c}_o(\cdot,\cdot,\cdot) = c_o(\cdot,\cdot,\cdot)/\beta$ and (25) and (26), note that

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

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

Now, consider the FOC for $(\tilde{P}, \tilde{e})$, namely (22)–(23), and use (B.1)–(B.2), $\tilde{\psi}(\cdot,\cdot) = \psi[(\cdot - \alpha)/\beta]$, $\tilde{\theta} = \alpha + \beta\theta$, $\tilde{F}(\cdot|\cdot) = F[(\cdot - \alpha)/\beta|\cdot]$, $\tilde{f}(\cdot|\cdot) = (1/\beta)f[(\cdot - \alpha)/\beta|\cdot]$ and $\tilde{\lambda}(\cdot) = \lambda(\cdot)$ to obtain

$$\begin{aligned} \tilde{P}\overline{y}'(\tilde{P}, Z) + \frac{\lambda}{1+\lambda}\overline{y}(P, Z) &= \left(\theta - \frac{\tilde{e} - \alpha}{\beta}\right)\overline{c}_o'(P, Z) \\ \psi'\left(\frac{\tilde{e} - \alpha}{\beta}, Z\right) + \frac{\lambda}{1+\lambda}\frac{F(\theta|Z)}{f(\theta|Z)}\psi''\left(\frac{\tilde{e} - \alpha}{\beta}, Z\right) &= \overline{c}_o(\tilde{P}, Z) \end{aligned}$$

since $H(x, z) = x$ and $H'(x, z) = 1$ in the basic model. From the solution $p^*(\theta, z)$ and $e^*(\theta, z)$ of (22)–(23), 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 (21) with $H(x, z) = x$, 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 (23) 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), $\tilde{\psi}(\cdot, \cdot) = \psi[(\cdot - \alpha)/\beta]$, $\tilde{\theta} = \alpha + \beta\theta$, $\tilde{e} = \alpha + \beta e$ 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, (23) 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 $\tilde{\mathcal{S}}$ satisfies assumptions A1–A3 and B1 as soon as the structure $\mathcal{S}$ satisfies these assumptions. QED

Proof of Lemma 4: Recall that $P = p^*(\theta, Z)$, where $p^*(\cdot, \cdot)$ is the optimal price schedule. Assumption B1 thus implies that P is independent of ϵ_d given Z. Hence, (20) gives $E[Y|P = p, Z = z] = E[y(p, z, \epsilon_d)|P = p, Z = z] = E[y(p, z, \epsilon_d)|Z = z] = \int y(p, z, \epsilon_d)dG(\epsilon_d|z) = \overline{y}(p, z)$ by (25). This establishes (28).

Regarding (26), 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 (20)–(21) with $H(x, z) = x$ and using that $\theta - e = \theta^*(P, Z) - e^*(P, Z)$, we obtain

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

where $\theta - e = \theta^*(p,z) - e^*(p,z)$. The second equality follows from assumption B1 and the last equality follows from (26). In particular, (B.3) establishes (30) 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 (27). 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

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

because $[\underline{\theta}(z) - \overline{e}(z)] = 1$ by the normalization (27).

Next, we write (22) with $H(x,z) = x$ at $Z = z$ so that $P = p^*(\theta,z) = p$ and $e = e^*(\theta,z)$. Dividing the resulting equation by (B.3) we obtain

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

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

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

Solving for $\overline{c}_o(p,z)$ and using the boundary condition (B.4) give (29). QED

Proof of Proposition 4: Because $\theta - e = \theta - e^*(\theta,z) = \theta^*(p,z) - e^*(p,z)$, differentiating (30) 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 (24) given $(P,Z) = (p,z)$, and using (B.3) together with $E[\epsilon_t|P = p, Z = z] = E[\epsilon_t|\theta, Z = z] = E[\epsilon_t|Z = z] = 0$ by assumption B1, we obtain

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

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

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

where we have used $e^*(\theta,z)=e^*(p,z)=e$. Thus, (B.5) and (31) gives

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

because $\psi'(\cdot,z)>0$ by assumption 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 assumption A2 and $\partial e^*(p,z)/\partial p<0$. Using (B.8)–(B.9) into (23) with $H'(x,z)=1$ 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, as in Guerre, Perrigne and Vuong (2000), 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.

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

$$\begin{aligned}\frac{\partial e^*(p,z)}{\partial p} &= -R(p,z)<0 && \text{(B.11)}\\ \frac{\partial \theta^*(p,z)}{\partial p} &= \frac{\partial\Delta(p,z)}{\partial p}-R(p,z)>0 && \text{(B.12)}\end{aligned}$$

where $R(p,z)$ is as given in (32) 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. Now, note that $e^*[\overline{p}(z),z]=e^*[\overline{\theta}(z),z]=0$ by (27). Moreover, from (30), 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 (34) and (35). As all the functions on the right-hand side of (35) 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 (27), and let $\overline{e}(z)\equiv e^*[\underline{p}(z),z]=e^*[\overline{\theta}(z),z]$. 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 (33) since (B.6) evaluated at $p=\overline{p}(z)$ gives $E[T|P=\overline{p}(z),Z=z]=\psi(0,z)$ as $e=e^*[\overline{p}(z),z]=0$ and $\theta=\theta^*[\overline{p}(z),z]=\overline{\theta}(z)$ when $p=\overline{p}(z)$. QED

Proof of Proposition 5: Part (i) follows exactly Matzkin (2003) identification argument. Because ϵ_d is independent of θ given Z by assumption B1, then ϵ_d is independent of $P = p^*(\theta, Z)$ given Z. Thus, if $G_{\epsilon_d|P,Z}(\cdot|\cdot,\cdot)$ denote 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 the second part of assumption B3. Hence, for every (p, z) we have

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

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 (39). This establishes (41) and hence (40) 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 assumption 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 the second part of assumption B3. 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 (39). Given the additional conditional in (ii), let $p = p_\dagger(z, \epsilon_d)$ in (B.14). We obtain

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

where the second equality follows from the additional condition in (ii), while the third equality follows from the second normalization in (39). This establishes (43). Equation (42) 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). QED

Proof of Lemma 5: From (B.8), we have $\psi'(e) = \Gamma[p^*(e, z), z]$, where $e = e^*(\theta, z)$. Hence, from (14) and making the change of variable $\tilde{p} = p^*[e^*(\tilde{\theta}, z), z] = p^*(\tilde{\theta}, z)$, we obtain

$$U^*(\theta) = \int_{p}^{\overline{p}(z)} \Gamma(\tilde{p}, z)\frac{\partial\theta^*(\tilde{p}, z)}{\partial p}d\tilde{p} = \int_{p}^{\overline{p}(z)} \Gamma(\tilde{p}, z)\left[\frac{\partial\Delta(\tilde{p}, z)}{\partial p} - R(\tilde{p}, z)\right]d\tilde{p}$$

where the second equality follows from (B.12). Then, (45) follows from the participation constraint (10) and

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

which follows from (B.5) and (B.7).

Similarly, making the change of variable $\tilde{p} = p^*(\tilde{e},z)$ in (33), we obtain

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

where the second equality follows from $p^*(0,z) = \overline{p}(z)$ and (B.11). Thus, using (14), (30), (45), (B.8) and (B.15) into (24) we obtain

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

We now compute $\Gamma(P,Z)/\overline{c}_o(P,Z)$. From (30), we note that $\Delta(P,Z)\overline{c}_o(P,Z) = E[C|P,Z]$. Thus, differentiating with respect to p gives

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

where

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

from (29)–(30). Hence,

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

It follows from (31) that

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

This establishes (44) in view of (B.16). Lastly, $E[\epsilon_t|P,Z] = E[\epsilon_t|Z] = 0$ by Assumption B1. QED

Proof of Proposition 6: Because θ and P are in a bijection given Z, we have $E[C\epsilon_t|P,Z] = E[C\epsilon_t|\theta,Z] = E[C|\theta,Z]E[\epsilon_t|\theta,Z] = 0$, where the second equality follows from assumption B4, while the third equality follows from $E[\epsilon_t|\theta, Z] = 0$ by assumption B1. Now, multiply (44) by C and take the expectation of the resulting equation conditional on (P, Z). Using $E[C\epsilon_t|P, Z] = 0$, this gives

$$\begin{aligned} E\{C(T - E[T|P,Z])|P,Z\} &= -\frac{\partial E[T|P,Z]/\partial p}{\partial E[C|P,Z]/\partial p - [P\overline{y}'(P,Z) + \mu\overline{y}(P,Z)]} \\ &\qquad \times \Big(E\{C(C - E[C|P,Z])|P,Z\}\Big) \end{aligned}$$

Since $E\{C(T - E[T|P,Z])|P,Z\} = Cov[C,T|P,Z]$ and $E\{C(C - E[C|P,Z])|P,Z\} = Var[C|P,Z]$, solving for μ gives (46) when $(P,Z) = (p,z)$. Then $\lambda(z)$ is obtained from $\mu(z) = \lambda(z)/[1 + \lambda(z)]$. Moreover, $Cov[C,T|P,Z] = E\{C(T - E[T|P,Z])|P,Z\} = -[\Gamma(P,Z)/\overline{c}_o(P,Z)]Var[C|P,Z]$ by (B.16) and Assumption B4. Thus, $Cov[C,T|P,Z] < 0$ as $\Gamma(\cdot,\cdot) > 0$ and $\overline{c}_o(\cdot,\cdot) > 0$. QED

References

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.

Chiappori P.A. and **B. Salanie** (2003): “Testing Contract Theory: A Survey of Some Recent Work,” in M. Dewatripont, L. Hansen and S. Turnovsky, eds., *Advances in Economics and Econometrics, Vol. I*, Cambridge University Press.

Fevrier, 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.

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.

Koopmans, T. (1949): “Identification Problems in Economic Model Construction,” *Econometrica*, 17, 125-144.

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

Laffont, J.J. (2004): *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 Aanalsyis 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.

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.

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

Park, B. and **L. Simar** (1994): "Efficient Semiparametric Estimation in Stochastic Frontier Models," *Journal of the American Statistical Association*, 89, 929-936.

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 **Q. Vuong** (in progress): "Econometrics of Incentive Regulation: Nonparametric Estimation," Working Paper, Pennsylvania State University.

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.

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