Centre de Recherche en Économie et Statistique

CREST

Série des Documents de Travail

n° 2011-27

Identification of a Class of Adverse Selection Models with Contracts Variation

X. D'HAULTFOEUILLE[1]
Ph. FÉVRIER[2]

1 CREST. Email : xavier.dhaultfoeuille@ensae.fr
2 CREST. Email : philippe.fevrier@ensae.fr

Identification of a Class of Adverse Selection Models with Contracts Variation *

Xavier D'Haultfoeuille † Philippe Février‡

December 2010

Abstract

We study the empirical content of an adverse selection model defined by the objective function of the principal, the agents' cost function and the distribution of agents' types. This model, though simple, encompasses the basic regulation, nonlinear pricing and price discrimination models, first price auctions and simple insurance settings. We analyse identification of the model in the presence of exogenous changes which affect the contracts but not the agents' cost and types. Our main result is that one or two such changes are sufficient to obtain full identification of the model. To establish this, we rely on functions that we call horizontal and vertical transforms and which allow us to identify the functions of interest by induction. Partial or full identification is then achieved by using either a fixed point strategy or results from group theory and dynamical systems.

Keywords: Adverse Selection, Nonparametric Identification, Fixed Points, Group Theory.

JEL classification numbers: C14, D82, D86

*We are very grateful to Steve Berry, Pierre Février, Etienne Ghys, Thierry Magnac, Martin Pesendorfer, Rob Porter, Jean-Charles Rochet, Bernard Salanié, and the participants of various seminars and conferences for helpful discussions and comments.

†CREST (INSEE). E-mail address: xavier.dhaultfoeuille@ensae.fr.

‡CREST (ENSAE). E-mail address: fevrier@ensae.fr.

1 Introduction

Over the past three decades, extensive attention has been devoted to asymmetries of information and their consequences in economics. A canonical example where these asymmetries play a fundamental role is the adverse selection model. This model has been helpful, for instance, to better understand theoretically nonlinear pricing, regulation, financial contracts or taxation theory. In recent years, the empirical literature on adverse selection models has also grown rapidly.[1] However, apart from the auction literature, most of these papers adopt a parametric framework.[2] Such parametric restrictions obscure what can be identified nonparametrically from the model and data. This issue is yet important to safely investigate optimality of contracts or do policy exercises, without being sensitive to a particular choice of parametrization.

In this paper, we analyze the nonparametric inference on a simple adverse selection model when the econometrician observes the contract and the associated trades. The model we consider is characterized by the objective function of the principal, the distribution of agents' types and the cost function of the agents. The space of trades available to the agent is supposed to be an interval of the real line, the agent's type is unidimensional and the cost is assumed to be separable. Such conditions discard discrete choices, which are common in price discrimination for instance, as well as multidimensional screening, and reduce the dimensionality of the cost function. Despite these limitations, the model is well suited to several settings, such as the basic regulation, nonlinear pricing and price discrimination models, financial contracts, delegation of tasks by firms, first price auctions or simple insurance models. Thus, even if our model cannot take the particular features of each possible application into account, we believe that our results deliver useful insights on the empirical content of adverse selection models, in a spirit close to what has been done in auctions, building on the work of Guerre et al. (2000). Understanding the econometrics of

[1] Applications include auction models (see e.g. Paarsch, 1992, and Guerre et al., 2000), regulatory contracts (see, among others, Wolak, 1994, Gagnepain and Ivaldi, 2002, Perrigne, 2002, Perrigne and Vuong, 2008, and Lavergne and Thomas, 2005), nonlinear pricing (see Ivaldi and Martimort, 1994, Miravete, 2002, Miravete and Roller, 2005, and Perrigne and Vuong, 2010) and price discrimination models (see Leslie, 2004, and Crawford and Shum, 2007, 2007). Adverse selection models have also been used to study the provision of incentives in firms, see e.g. Ferrall and Shearer (1999) and Paarsch and Shearer (2000).

[2] Notable exceptions are the papers of Perrigne and Vuong (2008) and Perrigne and Vuong (2010). The first studies the Laffont and Tirole (1986) regulation model in which ex-post costs are observed, and shows that such a model is nonparametrically identified. The second considers a nonlinear pricing model, slightly different from the one discuss below, and also proves nonparametric identifiability.

the most simple common structure of these models is helpful before studying more complex ones.

As shown in a companion paper (see D'Haultfoeuille and Février, 2010), the model is not nonparametrically identified without variation in the contracts. On the other hand, exogenous changes where the contracts vary while the cost function and the distribution of the agents' type remain constant can be used to recover the functions of interest. This amounts to observing an instrumental variable which affects the principal's objective function but not the agent. In the nonlinear pricing model, for instance, one may use any cost shifter of the firm (the principal), since it induces changes in its objective function but is unlikely to affect the consumers directly. Similarly, in the delegation of a production to agents, any demand shifter on the produced good is likely to be a valid instrument. In a first price auction setting, the number of bidders also satisfies this requirement if it is independent of the valuation of the good (Guerre et al., 2009).

Within this framework, our main result is that one or two exogenous changes are sufficient to yield full identification of the model. To establish this result, we extensively use the first-order condition of the agent, which defines the optimal choices of the agents, and the link between the observed distribution of the trades and the unobserved distribution of the agents' types. The first equation allows us to define what we call horizontal transforms whereas the second one yields vertical transforms. These transforms are identified in the data and can be combined to identify recursively the functions of interest.

Building on this idea, we show that the model is fully identified with only one change if two transfer functions cross. The idea is to used fixed points on the horizontal transform to recover the structural functions elsewhere. This result is related to the recent result of Guerre et al. (2009) on the nonparametric identification of risk aversion in first price auctions. We extend here their result to other adverse selection frameworks. We also prove that with two or more changes, the same idea can be used even if no pair of transfer functions cross, because fixed points for compositions of the horizontal transforms may still exist.

Nevertheless, the fixed point strategy breaks down in several settings, such as the one considered by D'Haultfoeuille and Février (2010). We study identification in this case by characterizing the set of points which can be reached using horizontal and vertical transforms, starting from an initial point. In the language of group theory, this amounts to studying the orbit of a point for the group generated by the horizontal transforms. Using this strategy, we show that the model is set identified with one exogenous change in the contracts, but point identified under a very mild restriction with two or more changes. Basically, this latter result states that the orbit of a point will be of the same topological nature as additive subgroups of the real line, which are known to be dense under minimal restrictions. As a result, the functions of interest are identified on a dense subset, and thus everywhere by continuity.

An important feature of our identification procedure is that the cost function and the distribution of the agents' types are recovered using the agent's program solely. This is convenient when the optimality of the principal is questionable. For instance, the common knowledge assumption on the distribution function of the agents' types or their cost function may fail to hold, the principal may also be risk averse (see Lewis and Sappping-ton, 1995) and the costs of implementing nonlinear contracts may modify significantly his program (see Ferrall and Shearer, 1999). Our results are not affected by these problems. Actually, our identification results may be used as a first step for testing the optimality of the contracts, when theory provides restriction on the principal's objective function, such as in regulation or financial contracts (see, e.g., Baron and Myerson, 1982 and Freixas and Laffont, 1990).

The paper is organized as follows. Section 2 recalls the main theoretical results for a principal-agent model with adverse selection. Section 3 is devoted to the nonparametric identification of this model. Section 4 develops some extensions. Section 5 concludes. All proofs are deferred to the appendix.

2 The Adverse selection model

Since the seminal work of Myerson (1979, 1981), extensive attention has been devoted to the theoretical properties of adverse selection models. We follow closely here the presentation of Laffont and Martimort (2002) and consider a basic adverse selection model where a principal trades y with some agents and provides them with a monetary transfer t. Agents are heterogeneous with a quasi-linear utility function $U(t, y, \theta) = t - C(y, \theta)$.[3] The monetary cost $C(y, \theta)$ of implementing y depends on their type θ which is unobserved

[3]The convention, here, is that y is produced by the agents as in the regulatory model. Equivalently, we could assume that the agents consume y and that the utility function takes the form $U(t, y, \theta) = U(y, \theta) - t$ as in the price discrimination model. Note also that since the agent's program is determinist, we could suppose him to be risk averse (i.e. $U(t, y, \theta) = u(t - C(y, \theta))$ for a concave increasing function u) without modifying the results of the program. We omit u from the discussion since it is obviously not identified.

by the principal. We suppose that θ is real and nonnegative, so that we can interpret it as a measure of the agent's intrinsic efficiency, a smaller θ corresponding to a more efficient agent. We denote by $F_\theta(.)$ (resp. $f_\theta(.)$) the distribution function (resp. density function) of θ and suppose it to be common knowledge. Our first assumption restricts the functional form of the cost function.

Assumption 1 *(cost separability)* $C(y,\theta) = \theta C(y)$ *where* $C(.)$ *is three times continuously differentiable,* $C'(y) > 0$ *for all* $y > 0$ *and* $C''(.) > 0$.

Basically, cost separability is a restriction that reduces the dimensionality of the problem. In general, such a condition is necessary to obtain identification results. This assumption is quite common in the theoretical literature (see e. g. Wilson, 1993, or Laffont and Tirole, 1993) as well as in empirical works (see Wolak, 1994, Ferrall and Shearer, 1999, Lavergne and Thomas, 2005). It is also assumed by Perrigne and Vuong (2008) in their nonparametric analysis of the regulation model. We come back to this assumption in Section 4, and show that our results can be extended to other restrictions on the cost function, and even be relaxed in particular settings.

We now analyze separately the agent's program from the firm's one. Indeed, if everyone is usually ready to believe that the agents behave optimally, it is less clear that the econometrician always wants to impose the optimality of the contracts. Hence we differentiate in the paper the results that only rely on the optimal behavior of the agent from those that also require the contracts to be optimal.

2.1 The agent's program

The agent faces a set of contracts of the form $[(y, t(y)); y \in \mathbb{R}^+, t(y) \in \mathbb{R}^+]$. The agent of type θ can either refuse all contracts or accept one of them. If he accepts a contract (y, t), the agent delivers y and receives a transfer t. If he refuses, he obtains his outside opportunity utility level $\underline{U}$. Hence, the agent of type θ chooses the trade $y(\theta)$ satisfying

$$y(\theta) \in \arg\max_y [t(y) - \theta C(y)]. \tag{2.1}$$

Moreover, $y(\theta)$ is implemented if and only if the agent participates, ie $\max_y t(y) - \theta C(y) \geq \underline{U}$. In the following, we rely on the first-order condition of the agent. For this approach to be valid, the regular conditions below are imposed.

Assumption 2 *(regular conditions 1)* $t(.)$ *is three times differentiable and for all* $\theta > 0$, $t''(.)-\theta C''(.) < 0$, $t'(0)-\theta C'(0) > 0$, $\lim_{y\to+\infty} t'(y)-\theta C'(y) < 0$ *and* $\max_y t(y)-\theta C(y) \geq \underline{U}$. *Finally,* $F_\theta(.)$ *is continuous and strictly increasing on* $\mathbb{R}^+$.

The main condition is on the smoothness of t, which rules out kinks in the transfer functions and thus bunching, for which the contracts targeted for different types coincide.[4] The other conditions on $t(.)$ hold for instance if $t(.)$ is concave, $C'(0) = 0$ and $\lim_{y\to\infty} C'(y) = \infty$.

Under Assumptions 1 and 2, every agent participates and the agent's program (2.1) admits a unique solution which is defined by the first-order condition

$$t'(y(\theta)) = \theta C'(y(\theta)). \tag{2.2}$$

Moreover, by differentiating this equation, $y'(\theta)$ satisfies

$$[t''(y(\theta)) - \theta C''(y(\theta))]y'(\theta) = C'(y(\theta)).$$

Thus, $y'(\theta) < 0$ and there is indeed no bunching of types.

2.2 The principal's program

Given the agent's program (2.1), the principal chooses the transfer function in order to maximize his objective function. We suppose here the principal to be risk neutral and his objective function to be quasi linear, $W(t, y, \theta) = S(y, \theta) - t$. Let $t^*(.)$ denote the optimal contract for the principal. $t^*(.)$ satisfies

$$t^*(.) \in \arg\max_{t(.)} \int [S(y(\theta), \theta) - t(y(\theta))]\, f_\theta(\theta) d\theta \quad \text{s.t. } y(\theta) \text{ satisfies (2.1)}.$$

Without further restriction, $t^*(.)$ does not necessarily satisfy Assumption 2. The optimal contract may lead to bunching, for instance. Besides, the first-order condition of the principal may neither be necessary nor sufficient to describe the optimal contracts. To avoid these technicalities, we impose the following regularity conditions.

Assumption 3 *(regularity conditions 2)* S *is three times differentiable with* $\partial S/\partial y > 0$, $\partial^2 S/\partial y^2 \leq 0$ *and* $\partial^2 S/\partial y \partial\theta \leq 0$. f_θ *is continuously differentiable with* $f_\theta > 0$ *and* $\theta \mapsto \theta + F_\theta(\theta)/f_\theta(\theta)$ *is strictly increasing. For all* $\theta > 0$, $\frac{\partial S}{\partial y}(0, \theta) - [\theta + F(\theta)/f(\theta)]C'(0) > 0$ *and* $\lim_{y\to+\infty} \frac{\partial S}{\partial y}(y, \theta) - [\theta + F(\theta)/f(\theta)]C'(y) < 0$.

[4]Because bunching leads to rather different results both in theory and in terms of identification, we leave this case for future research.

These regularity conditions state that the objective function of the principal is increasing and concave, and that the distribution of θ satisfies a technical condition which holds for most single-peaked densities.

Theorem 2.1 *Under Assumptions 1 and 3,*

1. *The trade $y^*(\theta)$ corresponding to the optimal transfer is defined by*
$$\frac{\partial S}{\partial y}(y^*(\theta),\theta) = \left[\theta + \frac{F_\theta(\theta)}{f_\theta(\theta)}\right] C'(y^*(\theta)). \tag{2.3}$$

2. *The optimal transfer function is defined by $t^{*\prime}(y^*(\theta)) = \theta C'(y^*(\theta))$ and the border condition $\lim_{\theta\to\infty} t(y^*(\theta)) - \theta C(y^*(\theta)) = \underline{U}$. It satisfies Assumption 2.*

The theorem is proved for instance in Laffont and Martimort (2002). Integrating (2.2) shows that, compared to the symmetric optimal contract where agents'type is observed by the principal, the firm has to leave a positive information rent $\int_\theta^{+\infty} C(y^*(\tau))d\tau$ to agents of type θ for them to reveal their types. This information rent increases with the efficiency of the agent and creates inefficiencies in production (the term $F_\theta(\theta)/f_\theta(\theta)C'(y^*(\theta))$ in Equation (2.3)). Besides, one can remark that if contracts are optimal, Assumption 3 automatically entails Assumption 2. Hence, the first-order approach of the agent's program is valid and there is no bunching at equilibrium.

2.3 Examples

As mentioned in the introduction, this model encompasses several classical settings. A first one is price discrimination. In Mussa and Rosen (1978), the principal is a firm that produces a good of quality y at a cost $H(y)$. Agents have heterogenous preferences for quality θ and have a utility $U = \theta y - t$ if they pay t for a good of quality y. The same model can be used to study nonlinear pricing by a monopoly (see e.g. Maskin and Riley, 1984). A second example is financial contracts. In Freixas and Laffont (1990) framework, the principal is a lender who provides the borrower with a loan y. His utility is $S(y) = t - Ry$, where R denotes the risk-free interest rate. Agents are firms with profit $U = \theta f(y) - t$, where $\theta f(y)$ is the production of the firm, y represents the units of capital and θ is a productivity index. A third example is regulation. In the Baron and Myerson (1982) model, the regulator maximizes a weighted sum of the consumers' surplus and the regulated firms defined by heterogenous cost functions of the form $\theta C(y)$. In our notations,

$$S(y,\theta) - t(y) = (1-\alpha)\left[\int_0^y p(u)du - t(y)\right] + \alpha\left[t(y) - \theta C(y)\right], \tag{2.4}$$

where $p(.)$ denotes the price function.

Auctions also fit within this framework. In this case, the focus is primary put on the agent's program. The parameter θ (usually denoted v in these models) is the valuation for the good and y corresponds to the bid. In a first price auction with n risk-neutral bidders for example, the utility of the agents takes the form $U(y,\theta) = \theta F_y^{n-1}(y) - yF_y^{n-1}(y)$, where F_y is the cumulative distribution function of y and $F_y^{n-1}(y)$ corresponds to the probability of winning the auction with a bid y. Thus, we recover a separable form of the kind $U(y,\theta) = \theta U(y) - t(y)$ that corresponds to the agent's model.

Finally, this agent's model is also useful for simple insurance settings without moral hazard. The insurance company proposes to agents contracts of the form $(y, d(y))$ where y and $d(y)$ denote respectively the premium and the corresponding deductible. Agents are heterogenous with respect to the probability p of facing an accident. Letting $u(.)$ denote their vNM utility, the expected utility of agent p when choosing y is given by

$$U(y,p) = u(-y) + p\left[u(-y-d(y)) - u(-y)\right],$$

and the first-order condition satisfies

$$1 - d'(y) = \frac{p-1}{p}\frac{u'(-y)}{u'(-y-d(y))}.$$

The model fits within the previous framework by letting $t'(y) = 1 - d'(y)$, $C'(y) = \frac{u'(-y)}{u'(-y-d(y))}$ and $\theta = \frac{p-1}{p}$.

The model we consider has however some limits. For instance, the space of trades available to the agent is supposed to be an interval of the real line and the agent's type is unidimensional. Such conditions discard discrete choices, which are common in price discrimination for instance, as well as multidimensional screening (see Pioner, 2009, and Aryal et al., 2010, for identification results on this case). Similarly, the participation constraint does not depend on the agent's type θ, since the outside opportunity utility level $\underline{U}$ is constant. Finally, this model assumes that agents perfectly control the output y. No supplementary error term enters in the program, which might sometimes be restrictive. In all these cases, our results should be seen as a first step to understand the econometrics of these more complex models.

3 Identification

3.1 The setting

We now turn to the empirical content of the model. We suppose that the econometrician observes the trades at equilibrium $y_{X,Z}(\theta)$ for an infinite sample of agents and for several menus of contracts indexed by covariates X and instruments Z whose role is clarified below. We also suppose that the corresponding transfers $t_{X,Z}(y_{X,Z}(\theta))$ are observable, for each menu of contracts.[5] The trades and transfers enable one to identify the cumulative distribution function of $y_{X,Z}(\theta)$, $F_{y_{X,Z}}(.)$ and the transfer function $t_{X,Z}(.)$. The question is whether the cost functions $C_{X,Z}(.)$, the distributions of types $F_{\theta,X,Z}(.)$ and the principal's objective function $S_{X,Z}(.)$ can be recovered from these functions and the model.

In general, variation in the menus of contracts are due to changes in the cost function, the distribution of types or the principal objective function. Without exclusion restriction, such variations do not have any effect on identification since only the data for $X = x$ and $Z = z$ can be used to identify $C_{x,z}$, $F_{\theta,x,z}$ and $S_{x,z}$. We impose here the exclusion restriction $(C_{X,Z}, F_{\theta,X,Z}) = (C_X, F_{\theta,X})$. In other terms, Z does not affect neither the cost function nor the distribution of types. Variation in the contracts only stems from a change in the principal's function through a change in Z.

We suppose here that Z has a finite support, because we believe that in most empirical applications, only few changes in contracts are observed. The value of the instrument does not play any role, so we can suppose without loss of generality that $Z \in \{1, ..., K\}$, $K \geq 2$.[6] Finally we suppress X in our notations for the ease of exposition. Agents are thus supposed to be homogenous except for their unknown types; if they differ by observed characteristics, our results below must be understood to be conditional on these characteristics.[7] With these notations, the exclusion restriction takes the following form:

Assumption 4 *(exclusion restriction)* $C_1 = ... = C_K = C$ *and* $F_{\theta,1} = ... = F_{\theta,K} = F_\theta$.

There are several situations where Assumption 4 is likely to hold. In the monopoly price

[5] This assumption may be strong (see Wolak, 1994, and Ferrall and Shearer, 1999, for examples where the transfers are unknown).

[6] We only focus on the case where an instrument is available because the model is not identified otherwise (see D'Haultfoeuille and Février, 2010). The cases where a continuous instrument is available is considered in Subsection 4.3.

[7] The case where they differ by unobserved characteristics is postponed until Subsection 4.2.

discrimination model, the price of an input may vary, inducing a change in the cost function of the monopoly and thus in S. However, this variation does not affect the utility function of the consumer. As usual in this literature, any cost shifter Z may play the role of the instrument. Similarly, in the delegation of a production to agents, exogenous variations of the market value of the product affects the principal's objective function but not the agents' one. In such cases, any demand shifter is a valid instrument Z. An example of this kind is developed in D'Haultfoeuille and Février (2010). In the regulation context, one may use for example changes in the government color. Such a change is likely to induce variation in the weight that the principal put on the firms profit compared to the consumers surplus, but not in $C(.)$ and $F_\theta(.)$ (Gagnepain and Ivaldi, 2007). Finally, in an auction context, one may rely on changes in the number of bidders, following Guerre et al. (2009).

Our framework can sometimes be applied even if the changes in the menus are due to modifications of $C(.)$ and $F_\theta(.)$. Indeed, suppose that these modifications appear continuously, while the principal only modifies his menu of contracts from time to time, because of menu costs. This situation typically arises in nonlinear pricing or price discrimination. Then trades and transfers just before and after the menu change satisfy Assumption 4. This idea is close to the one of regression discontinuity (see e.g. Hahn et al., 2001). In this case, the menu of contracts just before the change is inoptimal in the sense that it does not correspond to the one defined in Theorem 2.1. However, as will become clear below, this does not preclude identification of $C(.)$ and $F_\theta(.)$.

Other examples involving inoptimal variations of contracts are experiments, in which different menus of contracts are proposed to people in a random way. For instance, the Rand Health Insurance experiment (see Manning et al., 1987) randomly assigned families who participate in the experiment to 14 different insurance plans. Similarly, Ausubel (1999) and Karlan and Zinman (2009) analyse the market for bank credit by using randomized mailed solicitations. The propositions vary among several dimensions such as the interest rate or the duration of the loan.

3.2 The horizontal and vertical transforms

First, note that a normalization is necessary since we can replace $(\theta, C(.))$ by $(\alpha\theta, C(.)/\alpha)$ and leave the model unchanged. Hence, for a given $y_0 > 0$, we can choose any $\theta_0 > 0$ such that $\theta_1(y_0) = \theta_0$, where $\theta_k(.)$ denotes the inverse of $y_k(.)$.[8] Now, our identification results are based on two transforms. First, under Assumption 2, $\theta_k(.)$ is strictly decreasing. As a result, for all $y \geq 0$,

$$1 - F_{y_k}(y) = P(y_k(\theta) > y) = P(\theta \leq \theta_k(y)) = F_\theta(\theta_k(y)), \tag{3.1}$$

where the second equality stems from the fact that the distribution of θ is atomless, by Assumption 2. This equation, together with Assumption 3, also implies that $F_{y_k}(.)$ is strictly increasing and for all $(i, j) \in \{1, ..., K\}^2$,

$$F_{y_j}(y_j(\theta)) = F_{y_i}(y_i(\theta)).$$

Hence, letting $H_{ij}(y) = F_{y_j}^{-1}[F_{y_i}(y)]$ denote the quantile-quantile transformation between the distribution of $y_j(\theta)$ and $y_i(\theta)$, we get

$$y_j(\theta) = H_{ij}(y_i(\theta)). \tag{3.2}$$

Because $H_{ij}(.)$ is identified, the knowledge of $y_i(\theta)$ implies the knowledge of $y_j(\theta)$. From an economic perspective, this equality simply states that it is possible to recover the output of an agent of type θ when $Z = j$ if we know which production he chooses when $Z = i$. To do so, even if his type θ is unobserved, it suffices to pick the quantile of F_{y_j} corresponding to $F_{y_i}(y_i(\theta))$. In Figure 1, we can recover point (1) for instance if we know point (0). Another consequence is that it suffices to identify $y_1(.)$ (or equivalently $\theta_1(.)$) to recover the other functions $y_k(.)_{2\leq k\leq K}$ (or $\theta_k(.)_{2\leq k\leq K}$).

The second transform relies on the agent's first-order condition, which defines the unique solution of the program under Assumption 2. Taking Equation (2.2) at $\theta_k(y)$, we obtain

$$t'_k(y) = \theta_k(y)C'(y). \tag{3.3}$$

This implies in turn that for all $y > 0$,

$$\frac{t'_i(y)}{\theta_i(y)} = \frac{t'_j(y)}{\theta_j(y)}.$$

If we define the vertical transform $V_{ij}(., .)$ by $V_{ij}(\theta, y) = t'_j(y) \times \theta / t'_i(y)$, we get

$$\theta_j(y) = V_{ij}(\theta_i(y), y). \tag{3.4}$$

[8]In our setting, the choice of y_0 will be innocuous except in the partial identification case corresponding to Theorem 3.3 (see D'Haultfoeuille and Février, 2010 for a discussion in this particular case). Besides, once a normalization has been done on $\theta_1(.)$, no other normalization on the $\theta_k(.)$, $k \geq 2$, is needed. Indeed, the normalization on $\theta_1(y_0)$ induces a normalization on $C'(.)$ and $F_\theta(.)$ by Equations (3.3) and (3.1). This normalization then applies to all other menus of contracts.

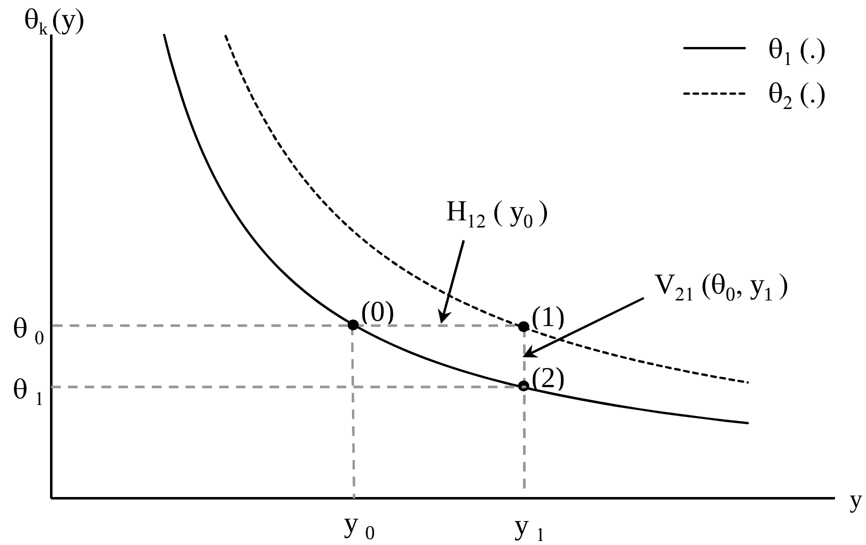

Figure 1: The horizontal and vertical transforms.

Because $V_{ij}(.,.)$ is identified on $\mathbb{R}^{+2}$, the knowledge of $\theta_i(y)$ implies the knowledge of $\theta_j(y)$. Contrary to the horizontal transform which links different outputs that similar agents choose facing different values of Z, the vertical transform links different types of agents who chooses the same level of output in those different situations. Knowing the type of an agent with an optimal output of y when $Z = i$, it is possible to recover the type of the agent that chooses the same level y when $Z = j$.[9]

Coming back to Figure 1, we can identify point (2) starting from point (1). Hence, starting from $(y_0, \theta_1(y_0))$, we can identify $(y_1, \theta_1(y_1))$ where $y_1 = H_{12}(y_0)$ and $\theta_1(y_1) = V_{21}(\theta_1(y_0), y_1)$. By induction, we can then identify $\theta_1(.)$ on a set of points. Then we also recover $C'(.)$ and $F_\theta(.)$ on some points, as the following lemma shows.

Lemma 3.1 *Suppose that for a given y, $\theta_1(y)$ is identified. Then $C'(y)$ and $F_\theta(\theta_1(y))$ are identified.*

Proof: the first-order condition (3.3) shows that $C'(.)$ is identified on y. Besides, because $y_1(.)$ is identified and $F_\theta(\theta_1(y)) = 1 - F_{y_1}(y)$ by (3.1), $F_\theta(.)$ is identified at $\theta_1(y)$. □

Although straightforward, this lemma has some interesting consequences. First, it shows that it suffices to focus on the identification of $\theta_1(.)$. Second, this lemma can be useful to

[9] The cost separability assumption is essential to define the vertical transforms. As shown in Subsection 4.1, it is possible however to define other transforms similar to the vertical one if other restrictions are imposed.

identify the model without exogenous variation but when side information is available. If for instance $C'(.)$ (resp. $F_\theta(.)$) is known or identified, we can recover $\theta_1(.)$ by (3.3) (resp. by (3.1)) and thus also $F_\theta(.)$ (resp. $C'(.)$).[10]

The way we use horizontal and vertical transforms varies depending on how contracts change. As a result, we separate two main kind of changes (referred to as crossing and noncrossing cases) where both proofs and identification results are quite different.

3.3 The crossing case

We first consider the case where two functions $t_i'(.)$ and $t_j'(.)$ cross.

Assumption 5 *(Crossing condition 1) There exists $i \neq j \in \{1, ..., K\}^2$ and a finite set of points $0 < y_0 < ... < y_M$ such that $t_i'(y_l) = t_j'(y_l)$ for $l = 0, ..., M$.*

In this case, the model can be fully recovered thanks to the intersection points. The intuition of this can be explained as follows. By the normalization, the value $\theta_1(y_0)$ of an intersection point y_0 of $t_1'(.)$ and $t_2'(.)$ can always be fixed to any $\theta_0 > 0$. For any y_α and θ^0, define the sequence $(\theta^n)_{n\in\mathbb{N}}$ as in Figure 2. We show that $(\theta^n)_{n\in\mathbb{N}}$ always converges, but reaches θ_0 if and only if $\theta^0 = \theta_1(y_\alpha)$. This allows us to recover $\theta_1(y_\alpha)$, since θ_0 is known. Because y_α was arbitrary, this proves that $\theta_1(.)$ is fully identified. Then by Lemma 3.1, the functions of interest can be recovered.

[10] An example is regulation. Suppose that total costs are observable ex post, as in Wolak (1994) and Perrigne and Vuong (2008), i.e. that $\theta(y)C(y)$ is identified. Because $\theta(y)C'(y) = t_1'(y)$ is also identified, $[C'/C](.)$ is identified. Then, by integration, $C'(.)$ can be recovered up to a multiplicative constant, which is then given by the normalization $\theta(y_0) = \theta_0$. As a result, $\theta_1(.)$, and thus $F_\theta(.)$, are identified.

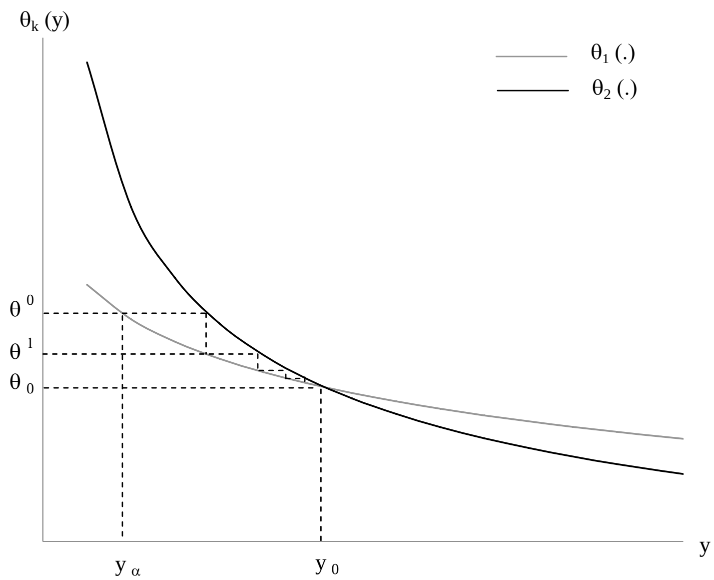

Figure 2: Identification when $t'_1(.)$ and $t'_2(.)$ cross.

Theorem 3.1 *If Assumptions 1, 2, 4 and 5 hold, $C'(.)$ and $F_\theta(.)$ are identified.*

Theorem 3.1 shows that full identification can be achieved provided that Assumption 5 holds. It is reminiscent of the result of Guerre et al. (2009) in the context of first-price auctions with risk averse bidders. They also use exogenous variations (namely, variation in the number of bidders) to obtain identification of the model at the limit, using a converging sequence (see their Proposition 1). We provide further details on the link between the result of Guerre et al. (2009) and ours in Subsection 4.1.

Assumption 5 may be considered restrictive. We now show that actually, it is possible to weaken it substantially when $K \geq 3$. To see this, remark that identification is secured above thanks to the existence of a fixed point on $H_{ij}(.)$ (since $H_{ij}(y) = y$ if and only if $t'_i(y) = t'_j(y)$). But in this reasoning, we need not restrict ourselves to $H_{ij}(.)$. Any identified transform may be used instead. We thus consider the following condition.[11]

Assumption 6 *(Crossing condition 2) There exists $(j_1, k_1, ..., j_p, k_p) \in \{1, ..., K\}^{2p}$ and $(l_1, ..., l_p) \in \mathbb{N}^p$ such that $G = H^{l_1}_{j_1k_1} \circ ... \circ H^{l_p}_{j_pk_p}$ admits a positive and finite number of fixed points in $\mathbb{R}^{*+}$.*

When $K = 2$, Assumption 6 is equivalent to Assumption 5. However, when $K \geq 3$, it may hold even if t'_i and t'_j never cross, for all $i \neq j \in \{1, ..., K\}^2$. As an illustration, consider

[11] We let subsequently $\circ$ denote the composition operator so that $f \circ g(x) = f(g(x))$. For any function f, we also let $f^k = f \circ ... \circ f$ if $k \geq 1$, $f^0(x) = x$ and $f^{-k} = f^{-1} \circ ... \circ f^{-1}$ (for $k \geq 1$) if f is one to one.

the case where $K = 3$, $C'(y) = y$, and, for all $y \in \mathbb{R}^+$, $t'_1(y) = \delta_1/(1+y)$, $t'_2(y) = \delta_2$ and $t'_3(y) = \delta_3$, with $\delta_1 < \delta_2 < \delta_3$. These functions do not cross. Besides, some algebra show that $H_{23}(y) = (\delta_3/\delta_2)y$ and

$$H_{21}(y) = \frac{\sqrt{1+\frac{4\delta_1 y}{\delta_2}}-1}{2}.$$

Thus, for all $k \geq 1$,

$$H_{23}^k \circ H_{21}(y) = \left(\frac{\delta_3}{\delta_2}\right)^k \frac{\sqrt{1+\frac{4\delta_1 y}{\delta_2}}-1}{2}. \tag{3.5}$$

Let $k_0 \geq 1$ be such that $(\delta_3/\delta_2)^{k_0}\delta_1/\delta_2 > 1$. It follows from (3.5) that $H_{23}^{k_0} \circ H_{21}(y)/y$ is continuous, decreasing, tends to $(\delta_3/\delta_2)^{k_0}\delta_1/\delta_2 > 1$ at zero and to zero at infinity. Thus, by the intermediate value theorem, $H_{23}^{k_0} \circ H_{21}$ admits a unique fixed point and Assumption 6 is satisfied, though the marginal transfer functions do not satisfy Assumption 6.

Theorem 3.2 *If Assumptions 1, 2, 4 and 6 hold, $C'(.)$ and $F_\theta(.)$ are identified.*

Theorem 3.2 extends Theorem 3.1 by establishing full identification under the weaker crossing condition. It shows that the previous idea can be applied as soon as there exists an identified function, perhaps different from the horizontal transforms, which admits a fixed point.

3.4 The noncrossing case

We now turn to the noncrossing case, which formally corresponds to the following assumption.[12]

Assumption 7 *(Noncrossing condition) For all $p \in \mathbb{N}$, $(j_1, k_1, ..., j_p, k_p) \in \{1, ..., K\}^{2p}$ and $(l_1, ..., l_p) \in \mathbb{N}^p$, $H_{j_1k_1}^{l_1} \circ ... \circ H_{j_pk_p}^{l_p}(y) = y$ for a given $y \in \mathbb{R}^{*+}$ implies that $H_{j_1k_1}^{l_1} \circ ... \circ H_{j_pk_p}^{l_p}$ is the identity function.*

Assumption 7 states that no composition of the horizontal transforms admits any fixed point, unless this composition is the identity function. When $K = 2$, this condition is actually equivalent to the fact that $t'_1(.)$ and $t'_2(.)$ do not cross. It is more restrictive when $K \geq 3$ because, as discussed above, Assumption 6 may hold even if none of the

[12] Assumption 7 is the contrary of Assumption 6, if we rule out the pathological cases where some compositions $H_{j_1k_1}^{l_1} \circ ... \circ H_{j_pk_p}^{l_p}$ admit an infinite number of fixed points without being the identity function, while all other compositions do not admit any fixed point.

marginal transfer functions cross. Still, there are important examples where Assumption 7 is satisfied, such as separable transfer functions. If, indeed, $t'_k(y) = \delta_k m(y)$ for all y and $k = 1...K$, then $\theta_k(y) = (\delta_k/\delta_1)\theta_1(y)$, so that $H_{jk} = \theta_1^{-1} \circ [(\delta_j/\delta_k)\theta_1]$. As a result, for all $p \geq 1$ and $(j_1, k_1, ..., j_p, k_p)$,

$$H^{l_1}_{j_1 k_1} \circ ... \circ H^{l_p}_{j_p k_p}(y) = \theta_1^{-1} \left[\prod_{m=1}^{p} \left(\frac{\delta_{j_m}}{\delta_{k_m}} \right)^{l_m} \times \theta_1(y) \right]. \tag{3.6}$$

Such functions are either strictly above, equal to, or strictly below the identity function, depending on the position of the constant inside the brackets with respect to one. Thus, they are equal to the identity function as soon as they admit a fixed point, and Assumption 7 holds.

Of course, under this condition, one cannot rely on the previous fixed point strategy anymore. Instead, the idea is to start from y_0, where $\theta_1(.)$ is identified, and describes the set of points $\mathcal{S}$ which can be reached by applying successively the horizontal transforms. At this stage, it is helpful to state this problem in terms of group theory, in order to use powerful results on the topological nature of orbits. Let G denote the group generated by $(H_{ij})_{(i,j)\in\{1,...,K\}^2}$ and the composition operator.[13] Then $\mathcal{S}$ is equal to the *orbit* of y_0, $\mathcal{O}_{y_0}$, defined by

$$\mathcal{O}_{y_0} = \{g(y_0), g \in G\}.$$

Identification then depends on how large is $\mathcal{O}_{y_0}$.

When $K = 2$, G simply consists of the iterated functions H^n_{12}, $n \in \mathbb{Z}$, so that $\mathcal{O}_{y_0}$ corresponds to the sequence $(y_n)_{n\in\mathbb{Z}}$ defined by $y_n = H^n_{12}(y_0)$ (the black points in Figure 3). Because it is impossible to recover exactly $\theta_1(.)$ between two points, the model is partially identified.

[13]For definitions related to group theory, see the proof of Theorem 3.4.

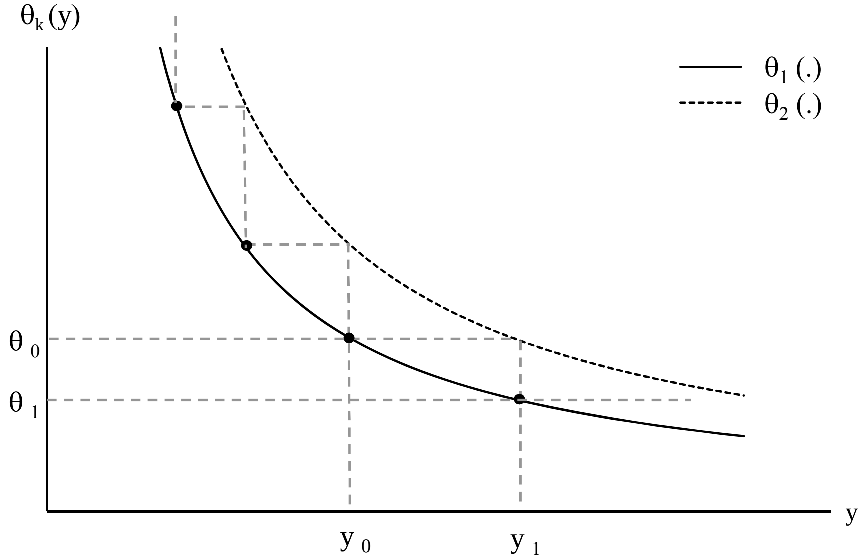

Figure 3: Identification in the noncrossing case, with $K = 2$.

Theorem 3.3 *If $K = 2$ and Assumptions 1, 2, 4 and 7 hold, $C'(.)$ and $F_\theta(.)$ are point identified respectively on $(y_n)_{n\in\mathbb{Z}}$ and $(\theta_n)_{n\in\mathbb{Z}}$, with $\theta_n = \theta_1(y_n)$. Elsewhere, they are partially identified by:*

$$\frac{t_1'(y)}{\inf_{n\in\mathbb{Z}:\, y_n\leq y}\theta_n} \leq \quad C'(y) \quad \leq \frac{t_1'(y)}{\sup_{n\in\mathbb{Z}:\, y_n\geq y}\theta_n}, \tag{3.7}$$

$$1 - F_{y_1}\left(\inf_{n\in\mathbb{Z}:\, \theta_n\leq\theta} y_n\right) \leq \quad F_\theta(\theta) \quad \leq 1 - F_{y_1}\left(\sup_{n\in\mathbb{Z}:\, \theta_n\geq\theta} y_n\right). \tag{3.8}$$

Theorem 3.3 extends Theorem 4.1 of D'Haultfoeuille and Février (2010) to any kinds of transfer functions, provided that they do not cross. It shows that $F_\theta(.)$ and $C'(.)$ are point identified on an infinite sequence, and by monotonicity, can be bounded elsewhere.[14] Even if we do not obtain full identification in this case, Theorem 3.3 implies that standard parametric models on $C'(.)$ and $F_\theta(.)$ are identified with an exogenous change. For instance, the parameters of a lognormal, Weibull or gamma distribution are identified thanks to the knowledge of $F_\theta(.)$ on the sequence $(\theta_1(y_n))_{n\in\mathbb{Z}}$. Actually, because we retrieve an infinite sequence of points on $C'(.)$ and $F_\theta(.)$, such standard parametric models are overidentified. The sequences $(C'(y_n))_{n\in\mathbb{Z}}$ and $(F_\theta(\theta_1(y_n)))_{n\in\mathbb{Z}}$ may thus serve as a guidance for

[14] The bounds (3.7) and (3.8) are not sharp in general, because we do not use the restrictions implied by second-order conditions. However, it is possible to show that these restrictions are not informative when transfer functions are convex. In such cases, our bounds are sharp.

choosing appropriate parametric restrictions (see D'Haultfoeuille and Février, 2010, for an application).

When $K \geq 3$, we may use H_{12} but also H_{13}. As shown in Figure 4, this allows us to define nonmonotonic sequences where $\theta_1(.)$ is identified. In particular, and contrary to previously, we are able to recover some information between y_0 and y_1. We may thus expect $\mathcal{S}$ to be large. We show below that under Assumption 7 and a mild restriction, this set is actually dense in $\mathbb{R}^+$. This implies that $\theta_1(.)$ is identified on a dense subset of $\mathbb{R}^+$, and thus, by continuity, on $\mathbb{R}^+$. As a result, the model is fully identified with two changes.

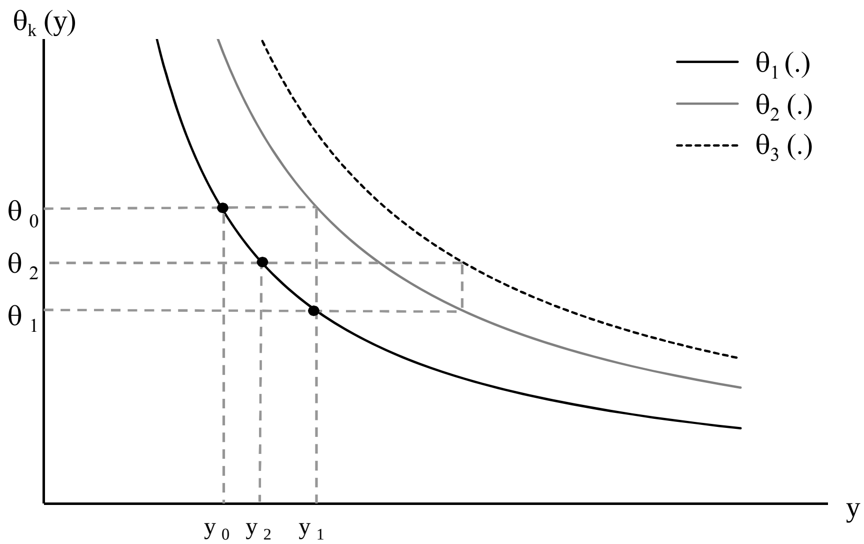

Figure 4: Identification in the noncrossing case, with $K \geq 3$.

To get some intuition on this result, reconsider the case of separable transfer functions $t'_k(y) = \delta_k m(y)$, with $K = 3$. Equation (3.6) shows that

$$\mathcal{O}_{y_0} = \left\{ \left(\frac{\delta_2}{\delta_1}\right)^m \left(\frac{\delta_3}{\delta_1}\right)^n,\ (m,n) \in \mathbb{Z}^2 \right\}.$$

This implies that $\theta_1 \circ \exp(.)$ is identified on

$$A = \left\{ mE_2 + nE_3,\ (m,n) \in \mathbb{Z}^2 \right\},$$

with $E_i = \ln(\delta_i/\delta_1)$ for $i \in \{2,3\}$. A is an additive subgroup of $\mathbb{R}$. By a classical result on these additive subgroups (see, e.g., Stillwell, 1992, p.33), A is either discrete or dense. Density is achieved if and only if $E_2/E_3 \notin \mathbb{Q}$, a very mild condition since the Lebesgue measure of $\mathbb{Q}$ is zero.

Qualitatively, the same phenomenon arises in the general case where Assumption 7 holds but transfer functions are not separable. The set $\mathcal{O}_{y_0}$ is either discrete or dense, the discrete case being the exception. Assumption 8 rules out this particular case.

Assumption 8 *(non periodicity) There exists* $(i,j,k) \in \{1,...,K\}^3$ *such that for all* $(m,n) \in \mathbb{Z}^2$, $(m,n) \neq (0,0)$, $H_{ij}^m \neq H_{ik}^n$.

Theorem 3.4 *If* $K \geq 3$ *and Assumptions 1, 2, 4, 7 and 8 hold,* $C'(.)$ *and* $F_\theta(.)$ *are identified.*

The proof relies on Hölder and Denjoy theorems, two deep results in group and dynamical systems theories. Denjoy theorem, in particular, ensures that $\mathcal{O}_{y_0}$ is either discrete or dense in $\mathbb{R}^+$, depending on whether a scalar called the *rotation number* (which corresponds to the ratio E_2/E_3 in the example above) is rational or not. Assumption 8 ensures that this number is irrational, establishing the density of $\mathcal{O}_{y_0}$. Interestingly, the proof shows that only three different contracts (i.e. two exogenous changes) are needed to achieve point identification of $C'(.)$ and $F_\theta(.)$. If $K \geq 4$ and Assumption 8 holds for four indices or more, the model is overidentified. Indeed, we can use different subsets of contracts to recover $C'(.)$ and $F_\theta(.)$. If the different corresponding functions do not coincide, then the model is rejected.

Assumption 7 is abstract, so one may wonder how it can be tested in practice. Actually, by Hölder theorem, the horizontal transforms commute under this condition. Hence, if $K = 3$ for instance, we can test for the much simpler condition $H_{12} \circ H_{13} = H_{13} \circ H_{12}$.

3.5 Implication for the principal's model

The identification results obtained so far only rely on the agent's model. This is convenient since the principal's model is often questionable. However, if one is willing to assume optimality of contracts, the first-order condition of the principal (2.3) can be used to recover some information about the objective function of the principal. More precisely, if $F_\theta(.)$ and $C'(.)$ are point identified for instance, as in Theorems 3.2 and 3.4, we identify the function $y \mapsto \frac{\partial S_k}{\partial y}(y, \theta_k(y))$. In general, this is not sufficient to recover the principal's objective function $S_k(.)$.[15] However, it is useful for testing the optimality of contracts,

[15] If one assumes that $S_k(.)$ does not not depend on θ, as is supposed in the standard price discrimination model for instance, this function is identified up to an additive constant.

provided that economic theory provides some restriction on $S_k(.)$. For instance, in the case of regulation, the principal objective function considered by the theory satisfies Equation (2.4). In this case,

$$\alpha_k = \frac{\frac{\partial S_k}{\partial y}(y, \theta_k(y)) - p_k(y)}{t'_k(y)}.$$

Thus, α_k is overidentified by this equation, and we can test for the optimality of the contracts by checking that the right-hande side is constant.

4 Discussion and extensions

4.1 The cost separability assumption

The vertical transform relies strongly on the cost separability assumption $C(\theta, y) = \theta C(y)$. This condition is nevertheless not as important as one may think in the first place. As mentioned above, what really matters is to reduce the dimensionality of the cost function $C(y, \theta)$ to secure identification. In some settings, other restrictions than the cost separability may be more natural. One example is the delegation of a task to an agent, as in, e.g., Ferrall and Shearer (1999) or Paarsch and Shearer (2000). Suppose that his production depends on an heterogeneity term θ that he observes ex ante, and on his effort e, so that $y = g(\theta, e)$. θ may represent the agent's productivity or the difficulty of the task itself. g is supposed to be increasing in e and known (or specified) by the econometrician. The cost $C(e)$ only depends on the effort e. Because there is no uncertainty for the agent, this model is not a moral hazard model but a truly adverse selection one.[16] We can reformulate it in our framework by replacing e by $g_2^{-1}(\theta, y)$, where $g_2^{-1}(\theta, .)$ denotes the inverse function of $g(\theta, .)$. In this case, the cost function satisfies the restriction $C(y, \theta) = C(g_2^{-1}(\theta, y))$.

In this setting, the horizontal and vertical transforms still apply, but on variables (e, θ) instead of (y, θ). Letting $e_k(\theta)$ denote the effort chosen by agent θ when facing menu k, we have

$$e_j(\theta) = g_2^{-1}\left[\theta, F_{y_j}^{-1} \circ F_{y_i}\left(g(\theta, e_i(\theta))\right)\right],$$

which defines the horizontal transform in this context. Besides, by the agents' first-order condition,

$$t'_i\left[g\left(\theta_i(e), e\right)\right]\frac{\partial g}{\partial e}(\theta_i(e), e) = t'_j\left[g\left(\theta_j(e), e\right)\right]\frac{\partial g}{\partial e}(\theta_j(e), e),$$

[16] It is sometimes referred to as a "false moral hazard" model (see e.g. Laffont and Martimort, 2002).

where $\theta_k(.)$ denotes the inverse of $e_k(.)$. Solving this equation in $\theta_j(e)$ defines the vertical transform.[17] As a result, the previous results also apply in this setting.

Another example of a restriction different from the cost separability condition arises in the first price auction model with n risk averse bidders. In this model, the expected utility of the bidders satisfies $U(y,\theta) = F_y^{n-1}(y)u(\theta - y)$ if the player with valuation θ bids y, $u(.)$ denoting the vNM utility of the player. The restriction, implied by the model, is that the function $U(y,\theta)/F_y^{n-1}(y)$ only depends on $\theta - y$. The first-order condition of the agent satisfies (see Guerre et al., 2008)

$$\theta_n(y) = y + \left(\frac{u}{u'}\right)^{-1}(L(y)),$$

where $L(y) = \frac{F_y(y)}{(n-1)f_y(y)}$. With exogenous variation in the number n of bidders, our results can be adapted. The horizontal transform is defined as usual but the vertical transform is replaced by a "diagonal" transform $D(.)$ defined by

$$\theta_{n_2}(D(y)) = \theta_{n_1}(y) + D(y) - y,$$

where n_1 and n_2 are two different number of bidders.[18]

4.2 Endogenous changes in the menus of contracts

Up to now, we have discussed cases where menus of contracts change exogenously, according to an observable instrument Z.[19] In practice, it may happen that a random term ε observed by the principal but not by the econometrician affects the cost function of the agents or their distribution of types. In this case, contracts change endogenously and the transfer function depends not only on Z but also on ε. Similarly to Guerre, Perrigne and Vuong (2008) when considering endogenous participation to auctions, our method can still be applied, under the exclusion restriction that $C_{Z,\varepsilon} = C_\varepsilon$ and $F_{\theta,\varepsilon,Z} = F_{\theta,\varepsilon}$. Suppose that

$$t'_{Z,\varepsilon}(y) = \psi(y, Z, \varepsilon),$$

where ψ is strictly monotonic in ε and $\varepsilon \perp\!\!\!\perp Z$. This latter condition is usual in instrumental variable models (see e.g. Imbens and Newey, 2009). It is also natural in our framework for

[17] Of course, assumptions on the primitives are needed to ensure the unicity of the solution, as well as the validity of the first-order approach.

[18] Guerre et al. (2009) show that this model is identified. This result is similar to Theorem 3.1, because θ_{n_1} and θ_{n_2} cross (see Guerre et al., 2009).

[19] As previously, we omit covariates X for the ease of exposition.

which Z is a determinant of the principal's objective function whereas ε only affects the agent's side.

If ε is continuously distributed, and under the strict monotonicity of $\psi(y,z,.)$, we can suppose without loss of generality (up to redefining ψ) that ε is uniformly distributed. Then $\Pr(t'_{Z,\varepsilon}(y) \leq t | Z = z) = \Pr(\varepsilon \leq \psi^{-1}(y,Z,t)) = \psi^{-1}(y,Z,t)$, where $\psi^{-1}(y,Z,.)$ is the inverse of $\psi(y,Z,.)$. Hence, $\varepsilon = \psi^{-1}(y(\theta), Z, t'_{Z,\varepsilon}(y(\theta)))$ is identified. Then we can control for ε, and our method applies conditional on ε, just as it applies conditional on covariates X. This idea is close to the one of control variables in identification of nonparametric models with endogenous variables (see e.g. Imbens and Newey, 2009). In this framework, the causal effect of the endogenous variable (the marginal transfer function) on a dependent variable is identified by adding a control variable (ε) resulting from a first step regression on instruments (Z).

As an illustration, consider the case where ε is an heterogeneity term on the cost function only, and $\varepsilon \mapsto C'_\varepsilon$ is increasing. Then, by, the principal's first-order condition and the fact that S_Z does not depend on θ here,

$$\theta_{Z,\varepsilon}(y) = G^{-1}\left(\frac{S'_Z(y)}{C'_\varepsilon(y)}\right),$$

where $G(\theta) = \theta + F_\theta / f_\theta(\theta)$. Thus, the optimal contract $t^*_{Z,\varepsilon}$ satisfies

$$t^*_{Z,\varepsilon}{}'(y) = G^{-1}\left(\frac{S'_Z(y)}{C'_\varepsilon(y)}\right) C'_\varepsilon(y).$$

If $\theta \mapsto F_\theta / \theta f_\theta(\theta)$ is increasing,[20] we can show that $t^*_{Z,\varepsilon}{}'$ is strictly increasing in ε, and our previous result applies.

4.3 Continuous instruments

It may happen that the econometrician has a continuous instrument at his disposal. In the price discrimination example, the price of an input of the monopoly may take any value in an interval, implying that the value function of the principal changes continuously. In this case, Theorem 4.1 shows that full identification can be obtained without the cost separability assumption. In other terms, no restriction is needed on $C(.,.)$, defined as a function of the two variables (y,θ). In this case, we can always normalize θ to be uniform.

[20] This condition is satisfied for instance by all Fréchet, Weibull and Pareto distributions.

Theorem 4.1 *Suppose that Z takes values in $\mathcal{Z} = [\underline{z}, \overline{z}]$, θ is uniformly distributed on $[0, 1]$, $(y, z) \mapsto t(y, z)$ is twice differentiable and $\frac{\partial^2 t}{\partial y \partial z}(y, z) > 0$.*[21] *Then $\frac{\partial C}{\partial y}(., .)$ is identified on $\{(y, \theta) : \exists z \in \mathcal{Z} : \theta(y, z) = \theta\}$.*

4.4 Selection effects

We have supposed until now that variation in the transfer functions does not yield any changes in $F_\theta(.)$. However, selection effects can be important. Lazear (2000), for instance, showed that half of the productivity increase observed in a car glass company after moving from constant wages to piece rates could be explained by the arrival of more productive workers. More generally, these effects may arise in competitive environments where agents can choose between several menus of contracts proposed by different principals. In this case indeed, a change in one principal's menu may induce some agents with particular θ to choose the new menu of contracts. Such effects are not taken into account in our model where all types of agent participate in all menus of contracts.[22] Hence, our analysis is not valid in general when selection occurs.

However, selection effects are not problematic if panel data are available. First, such effects can be detected by comparing the distributions of the stayers' and entrants' type, as in Lazear (2000). Moreover, our method still applies even in the presence of selection effects, provided that the distribution $\widetilde{F}_\theta(.)$ of the stayers (i.e., those who participate in all menus of contracts) remains the same for the different menus. If the exclusion restrictions hold on the population of stayers, $\widetilde{F}_\theta(.)$ can be recovered as well as the marginal cost function $C'(.)$. Then, once $C'(.)$ has been identified, we can use data on the movers to recover their own distribution of types, thanks to Lemma 3.1. At the end, the distribution of the types of the whole population of agents (i.e., movers and stayers) is identified, showing that selection effects can be handled in this framework.

[21] We also assume that the first-order condition of the agent is necessary and sufficient for optimality. For this to be satisfied, Assumption 2 could be modified in order to take into account the non-separability of the cost function.

[22] Selection effects could be modeled by letting the participation constraint of agent θ depends on θ.

5 Conclusion

This work contributes to the recent structural analysis of incentive problems.[23] We show in particular that when contracts vary exogenously, full identification of the model is achieved with at two most two changes. Our results are based on a new induction method that we apply to derive our identification results. In our companion paper (see D'Haultfoeuille and Février, 2010), we also rely on this induction technique to derive a consistent estimator with one exogenous change and no crossing. It would be interesting to extend this estimator to the more general setting considered here. When point identification is achieved, another possibility would be to develop a sieve estimator, following the suggestion of Guerre et al. (2009) on the estimation of auction models with risk averse bidders.

[23]For a structural analysis of moral hazard, see Ke (2008).

Appendix: proofs

Theorem 3.1

By the normalization and the fact that $y_0 > 0$, we can always fix $0 < \theta_0 < \infty$ such that $\theta_0 = \theta_i(y_0)$. We also suppose, without loss of generality that $t'_j(.) > t'_i(.)$ for all $y < y_0$.

Let $y_\alpha < y_0$, and define the increasing sequence $(y^n)_{n\in\mathbb{N}}$ by $y^0 = y_\alpha$ and, for all $n \geq 1$, $y^n = H^n_{ij}(y_\alpha)$. We have $y^n < y_0$ for all $n \in \mathbb{N}$. Indeed, the result is true for $n = 0$. Moreover, if it holds for $n-1$, then $y^n = H_{ij}(y^{n-1}) < H_{ij}(y_0) = y_0$ since H_{ij} is strictly increasing. The sequence is increasing and bounded above by y_0, so that it admits a limit y^∞ which satisfies $y^\infty = H_{ij}(y^\infty)$ and $y^\infty \leq y_0$. Hence $y^\infty = y_0$.

Now, by the first-order condition,

$$\theta_i(y^{n+1}) = V_{21}(\theta_i(y^n), y^{n+1}) = \frac{t'_i(y^{n+1})}{t'_j(y^{n+1})}\theta_i(y^n).$$

Thus, by a straightforward induction,

$$\theta_i(y_\alpha) = \theta_i(y^n)\prod_{i=1}^{n}\left[\frac{t'_j(y^i)}{t'_i(y^i)}\right].$$

Because $(y^n)_{n\in\mathbb{N}}$ converges to y_0 and $\theta_i(.)$ is continuous, the sequence $(\theta_i(y^n))_{n\in\mathbb{N}}$ converges to θ_0. Because $\theta_0 \in (0,\infty)$, the product on the right-hand side also admits a finite and positive limit as $n \to \infty$, and

$$\theta_i(y_\alpha) = \theta_0\prod_{i=1}^{\infty}\left[\frac{t'_i(y^i)}{t'_j(y^i)}\right].$$

The right-hand side can be recovered from the data, which proves that $\theta_i(y_\alpha)$ is identified. $y_\alpha < y_0$ was arbitrary, so that $\theta_i(.)$ is identified on $]0, y_0]$.

If $M = 0$, i.e. there is a unique crossing point, we can identify $\theta_i(y_\alpha)$ for all $y_\alpha > y_0$ as previously, using a decreasing sequence instead of an increasing one. Thus $\theta_i(.)$ is actually identified on $\mathbb{R}^{*+}$ in this case. If $M \geq 1$, we can identify similarly $\theta_i(y_\alpha)$ for all $y_1 > y_\alpha > y_0$. $\theta_i(y_1)$ is then identified by continuity. Hence, $\theta_i(.)$ is identified on $]0, y_1]$. A straightforward induction on M then shows that $\theta_i(.)$ is identified on $(0,\infty)$. The result then follows from Lemma 3.1. ■

Theorem 3.2

The proof is similar although a bit more involved than the one of Theorem 3.1. Let $0 < y_0 < ... < y_M$ be the fixed points of $G = H^{l_1}_{j_1k_1} \circ ... \circ H^{l_p}_{j_pk_p}$. By the normalization, we can always fix $0 < \theta_0 < \infty$ such that $\theta_0 = \theta_{j_p}(y_0)$. By the intermediate value theorem, for all $y \in (0, y_0)$ we have either $G(y) < y$ or $G(y) > y$. We suppose, without loss of generality, that $G(y) > y$ on this interval.

Let $y_\alpha \in (0, y_0)$, and define the sequence $(y^n)_{n\in\mathbb{N}}$ by $y^n = G^n(y_\alpha)$. As previously, $(y^n)_{n\in\mathbb{N}}$ is increasing and bounded above by y_0, so that it converges to y_0.

Now, applying the horizontal and vertical transforms, we get

$$\begin{aligned}\theta_{j_p}(y^n) &= \theta_{k_p}(H_{j_pk_p}(y^n)) \\ &= V_{j_pk_p}(\theta_{j_p}(H_{j_pk_p}(y^n)), H_{j_pk_p}(y^n)) \\ &= \frac{t'_{k_p}(H_{j_pk_p}(y^n))}{t'_{j_p}(H_{j_pk_p}(y^n))}\theta_{j_p}(H_{j_pk_p}(y^n)).\end{aligned}$$

Thus, by a straightforward induction,

$$\theta_{j_p}(y^n) = \frac{t'_{k_p}(H_{j_pk_p}(y^n))}{t'_{j_p}(H_{j_pk_p}(y^n))} \times ... \times \frac{t'_{k_p}\left(H^{l_p}_{j_pk_p}(y^n)\right)}{t'_{j_p}\left(H^{l_p}_{j_pk_p}(y^n)\right)}\theta_{j_p}\left(H^{l_p}_{j_pk_p}(y^n)\right). \tag{5.1}$$

By the vertical transform once more,

$$\theta_{j_p}\left(H^{l_p}_{j_pk_p}(y^n)\right) = \frac{t'_{j_p}\left(H^{l_p}_{j_pk_p}(y^n)\right)}{t'_{j_{p-1}}\left(H^{l_p}_{j_pk_p}(y^n)\right)}\theta_{j_{p-1}}\left(H^{l_p}_{j_pk_p}(y^n)\right). \tag{5.2}$$

Equations (5.1) and (5.2) imply that there exists a function $Q_{j_pk_pj_{p-1}}$, identified in the data, such that

$$\theta_{j_p}(y^n) = Q_{j_pk_pj_{p-1}}(y^n)\theta_{j_{p-1}}\left(H^{l_p}_{j_pk_p}(y^n)\right).$$

Applying the same reasoning to $\theta_{j_{p-1}}\left(H^{l_p}_{j_pk_p}(y^n)\right)$, $\theta_{j_{p-2}}\left(H^{l_{p-1}}_{j_{p-1}k_{p-1}} \circ H^{l_p}_{j_pk_p}(y^n)\right)$,... shows that there exists a function $\widetilde{Q}$ which is identified and such that $\theta_{j_p}(y^n) = \widetilde{Q}(y^n)\theta_{j_1}(G(y^n)) = \widetilde{Q}(y^n)\theta_{j_1}(y^{n+1})$. Finally, by an application of the vertical transform to $\theta_{j_1}(y^{n+1})$ and $\theta_{j_p}(y^{n+1})$, there is a function Q identified in the data satisfying

$$\theta_{j_p}(y^n) = Q(y^n)\theta_{j_p}(y^{n+1}).$$

As a result,

$$\theta_{j_p}(y_\alpha) = \theta_{j_p}(y^n)\left[\prod_{i=1}^{n} Q(y^{i-1})\right].$$

Because $(y^n)_{n\in\mathbb{N}}$ converges to y_0 and $\theta_{j_p}(.)$ is continuous, the sequence $(\theta_{j_p}(y^n))_{n\in\mathbb{N}}$ converges to θ_0. Because $\theta_0 \in (0,\infty)$, the product into bracket also admits a finite and positive limit as $n \to \infty$, and

$$\theta_{j_p}(y_\alpha) = \theta_0 \prod_{i=1}^{\infty} Q(y^{i-1}).$$

The right-hand side can be recovered from the data, proving that $\theta_{j_p}(y_\alpha)$ is identified. As $y_\alpha \in (0, y_0)$ was arbitrary, $\theta_{j_p}(.)$ is identified on $(0, y_0]$. The rest of the proof is similar to the one of Theorem 3.1. ■

Theorem 3.3

Because $H_{12}(.)$ and $H_{21}(.)$ are identified, it follows from the discussion before Theorem 3.3 that $\theta_1(.)$ is point identified on $(y_n)_{n\in\mathbb{Z}}$. For other y, we get, by monotonicity of $\theta_1(.)$,

$$\sup_{n:\, y_n \geq y} \theta_n \leq \theta_1(y) \leq \inf_{n:\, y_n \leq y} \theta_n,$$

where the supremum (resp. the infimum) is set to zero (resp. infinity) when the set is empty. Similarly,

$$\sup_{n\in\mathbb{Z}:\, \theta_n \geq \theta} y_n \leq y_1(\theta) \leq \inf_{n\in\mathbb{Z}:\, \theta_n \leq \theta} y_n.$$

Then Equations (3.1) and (3.3) imply Inequalities (3.8) and (3.7). ■

Theorem 3.4

Before proving the results, let us recall some definitions and results on groups. A group G is a set endowed with an operator $*$ such that for all $(a,b,c) \in G^3$, $a*b \in G$, $(a*b)*c = a*(b*c)$ and such that an identity element e satisfying $a * e = e * a = a$ for all $a \in G$ exists. Moreover, every element $a \in G$ admits an element (called the inverse of a) b which satisfies $a * b = b * a = e$. G is abelian if, for all a, b, $a * b = b * a$. A subgroup H of G is a subset of G which is itself a group for $*$. The group generated by a subset I of G is the smallest subgroup of G containing I. For any set X and a group G, a group action . is a function from $G \times X$ to X (denoted by $g.x$) satisfying, for every $(g,h) \in G^2$ and $x \in X$, $(g * h).x = g.(h.x)$ and $e.x = x$. A group action is free if, for any $x \in X$, $g.x = x$ implies $g = e$. The orbit $\mathcal{O}_x$ of x is then defined by

$$\mathcal{O}_x = \{g.x, g \in G\}.$$

We also consider functions on the unit circle $[0, 1)$.[24] For any real x, let $\pi(x)$ denote the fractional part of x. A map q on $[0, 1)$ is orientation-preserving if there exists an increasing function Q such that $q \circ \pi = \pi \circ Q$ and $Q(x + 1) = Q(x) + 1$. Then q is continuous (resp. a C^k diffeomorphism) on the unit circle if Q is continuous (resp. a C^k diffeomorphism) on the real line. If q is continuous, its rotation number $\rho(q)$ is defined by

$$\rho(q) = \lim_{n \to \infty} \frac{Q^n(x) - x}{n}.$$

Poincaré (1885) showed that this limit exists and is independent of x and Q. Finally, we use subsequently the following lemma.

Lemma 5.1 *For any increasing C^2 diffeomorphism r on $\mathbb{R}^{*+}$ satisfying $r(x) > x$, there exists an increasing C^2 diffeomorphism h from $\mathbb{R}$ to $\mathbb{R}^{*+}$ such that $r = h \circ \varphi \circ h^{-1}$, where φ is the translation $\varphi(x) = x + 1$ on the real line.*

Proof: let us consider an increasing C^2 diffeomorphism $\widetilde{h}$ defined on the interval $(0, 1)$ with a positive limit at 0 and such that $\lim_{x \to 1} \widetilde{h}(x) = \lim_{x \to 0} r \circ \widetilde{h}(x)$, $\lim_{x \to 1} \widetilde{h}'(x) = \lim_{x \to 0} \left[r \circ \widetilde{h}\right]'(x)$ and $\lim_{x \to 1} \widetilde{h}''(x) = \lim_{x \to 0} \left[r \circ \widetilde{h}\right]''(x)$. Such a $\widetilde{h}$ exists. Then define the function h by $h = \widetilde{h}$ on $(0, 1)$ and extend it on the real line, using $h(x + 1) = r \circ h(x)$ or $h(x) = r^{-1} \circ h(x + 1)$. By construction, h is strictly increasing and C^2. Hence, it admits a limit at $-\infty$ and $+\infty$. Suppose that $\lim_{x \to -\infty} h(x) = M > 0$. Then, because $h(x + 1) = r \circ h(x)$, we would have $r(M) = M$, a contradiction. Thus, $\lim_{x \to -\infty} h(x) = 0$. Similarly, $\lim_{x \to +\infty} h(x) = +\infty$. Consequently, h is a C^2 diffeomorphism from $\mathbb{R}$ to $\mathbb{R}^{*+}$. $\square$

Now, let us prove Theorem 3.4. Without loss of generality, we set the indices (i, j, k) defined in Assumption 8 to $(1, 2, 3)$. By Assumptions 1 and 2 and the first-order condition, θ_1 and θ_2 are C^2 diffeomorphisms on $\mathbb{R}^{*+}$. Then $H_{12} = \theta_2^{-1} \circ \theta_1$ and $H_{13} = \theta_3^{-1} \circ \theta_1$ are also C^2 diffeomorphisms on $\mathbb{R}^{*+}$. By Assumptions 7 and 8, H_{12} does not admit any fixed point. Suppose without loss of generality that $H_{12}(x) > x$. By Lemma 5.1, there exists an increasing C^2 diffeomorphism h such that $H_{12} = h \circ \varphi \circ h^{-1}$. Let $f = h^{-1} \circ H_{13} \circ h$, so that f is a real, increasing C^2 diffeomorphism. Let us denote by G the subgroup of real diffeomorphisms on the real line (endowed with the composition operator $\circ$) generated by φ and f. Consider the group action of G on $\mathbb{R}$ defined by $g.x = g(x)$. By Assumption 7, this group action is free. Then, by a theorem of Hölder (see, e.g., Ghys, 2001, Theorem

[24]Formally, the unit circle corresponds to classes of equivalence for the equivalence relationship $\mathcal{R}$ defined on $\mathbb{R}$ by $x\mathcal{R}y \Leftrightarrow x - y \in \mathbb{Z}$, but this can be ignored here.

6.10), G is abelian. As a result, $f(x+1) = f \circ \varphi(x) = \varphi \circ f(x) = f(x) + 1$ for all $x \in \mathbb{R}$. Define $\widetilde{f}$ on the unit circle by $\widetilde{f} \circ \pi = \pi \circ f$. This defines properly $\widetilde{f}$ because

$$\begin{aligned} \pi(x) = \pi(y) \quad &\Leftrightarrow \quad \exists\, k \in \mathbb{Z} \;/\; x = y + k = \varphi^{(k)}(y) \\ &\Rightarrow \quad f(x) = \varphi^{(k)} \circ f(y) \\ &\Rightarrow \quad \pi \circ f(x) = \pi \circ f(y). \end{aligned}$$

By construction, $\widetilde{f}$ is an orientation-preserving C^2 diffeomorphism on the unit circle. Then, by Denjoy's theorem (see, e.g., Navas, 2009, Theorem 3.1.1), any orbit for the group generated by $\widetilde{f}$ is finite if $\rho(\widetilde{f}) \in \mathbb{Q}$, and dense otherwise.

Suppose that the orbits are finite. Then there exists $n \in \mathbb{Z}$ such that $\widetilde{f}^n(x) = x$. It is easy to see that this implies that there exists $m \in \mathbb{Z}$ such that $f^n(x) = \varphi^m(x)$. Hence, by definition of f and φ, $H_{13}^n(x) = H_{12}^m(x)$, contradicting Assumption 8. We thus conclude that any orbit for the group generated by $\widetilde{f}$ is dense in the unit circle. Now, fix $(x, y) \in \mathbb{R}^2$ and consider a neighbourhood $\mathcal{V}_y$ of y. By definition, $\pi(\mathcal{V}_y)$ is a neighbourhood of $\pi(y)$ in the unit circle. Thus, there exists $n \in \mathbb{Z}$ such that $\widetilde{f}^n \circ \pi(x) \in \pi(\mathcal{V}_y)$. Because $\widetilde{f}^n \circ \pi = \pi \circ f^n$, $\pi \circ f^n(x) \in \pi(\mathcal{V}_y)$. Hence, there exists $m \in \mathbb{Z}$ such that

$$\varphi^m \circ f^n(x) \in \mathcal{V}_y.$$

This proves that any orbit $\mathcal{O}_x$ for the group generated by f and φ is dense in $\mathbb{R}^{*+}$. Now, the orbit $\mathcal{O}'_x$ for the group generated by H_{12} and H_{13} satisfies $\mathcal{O}'_x = h\left(\mathcal{O}_{h^{-1}(x)}\right)$. Thus $\mathcal{O}'_x$ is dense in $\mathbb{R}^{*+}$. In other words, starting from a given $y_0 > 0$, we can identify $\theta_1(.)$ on a subset which is dense in $\mathbb{R}^+$. $\theta_1(.)$ is thus identified everywhere by continuity. As a result, $C'(.)$ and $F_\theta(.)$ are also identified. ■

Theorem 4.1

By uniformity of θ, (3.1) now writes $\theta(y, z) = 1 - F_{y|z}(y|z)$. Thus, $\theta(., z)$ is identified on $\mathcal{Y}_z = \{y : \exists \theta \in \Theta : \theta(y, z) = \theta\}$. By (3.3), $\frac{\partial C}{\partial y}(., .)$ is also identified on $\{(y, \theta(y, z)), y \in \mathcal{Y}_z\}$, for all $z \in \mathcal{Z}$. Now, for all $y \in \mathcal{Y}_z$,

$$\frac{\partial^2 C}{\partial y \partial z}(y, \theta(y, z)) = \frac{\partial^2 C}{\partial y \partial \theta}(y, \theta(y, z)) \frac{\partial \theta}{\partial z}(y, z).$$

Moreover, $\frac{\partial^2 t}{\partial y \partial z}(y, z) > 0$ implies that $\frac{\partial \theta}{\partial z}(y, z) > 0$. Thus, for all $y \in \mathcal{Y}_z$, $\frac{\partial^2 C}{\partial y \partial \theta}(y, \theta(y, z))$ is identified by

$$\frac{\partial^2 C}{\partial y \partial \theta}(y, \theta(y, z)) = \frac{\frac{\partial^2 C}{\partial y \partial z}(y, \theta(y, z))}{\frac{\partial \theta}{\partial z}(y, z)} \;■$$

References

Aryal, G., Perrigne, I. and Vuong, Q. (2010), Identification of insurance models with mutlidimensional screening. Mimeo.

Ausubel, L. M. (1999), Adverse selection in the credit card market. Mimeo.

Baron, D. P. and Myerson, R. B. (1982), 'Regulating a monopolist with unknown costs', *Econometrica* **50**, 911–930.

Crawford, G. S. and Shum, M. (2007), 'Monopoly quality degradation and regulation in cable television', *Journal of Law and Economics* **50**, 181–219.

D'Haultfoeuille, X. and Février, P. (2010), The provision of wage incentives: A structural estimation using contracts variation. Mimeo.

Ferrall, C. and Shearer, B. (1999), 'Incentives and transaction costs within the firm: Estimating an agency model using payroll records', *Review of Economic Studies* **99**, 309–338.

Freixas, X. and Laffont, J. J. (1990), Optimal banking contracts, *in* 'Essays in Honor of Edmond Malinvaud, vol. 2: Macroeconomics', MIT Press, pp. 33–61.

Gagnepain, P. and Ivaldi, M. (2002), 'Incentive regulatory policies: The case of public transit systems in france', *The RAND Journal of Economics* **33**, 605–629.

Gagnepain, P. and Ivaldi, M. (2007), Contract choice, incentives, and political capture in public transit. Mimeo.

Ghys, E. (2001), 'Groups acting on the cirle', *L'enseignement mathématique* **47**, 329–407.

Guerre, E., Perrigne, I. and Vuong, Q. (2000), 'Optimal nonparametric estimation of first-price auctions', *Econometrica* **68**, 525–574.

Guerre, E., Perrigne, I. and Vuong, Q. (2009), 'Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions', *Econometrica* **77**, 1193–1227.

Hahn, J., Todd, P. and van der Klauw, W. (2001), 'Identification and estimation of treatment effects with a regression-discontinuity design', *Econometrica* **69**, 201–209.

Imbens, G. W. and Newey, W. K. (2009), 'Identification and estimation of triangular simultaneous equations models without additivity', *Econometrica* **77**, 1481–1512.

Ivaldi, M. and Martimort, D. (1994), 'Competition under nonlinear pricing', *Annales d'Economie et de Statistique* **34**, 71–114.

Karlan, D. and Zinman, J. (2009), 'Observing unobservables: identifying information asymmetries with a consumer credit field experiment', *Econometrica* **77**, 1993–2009.

Ke, R. (2008), Identifying contract optimality non-parametrically with moral hazard: First order approach and statistical inference. Mimeo.

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

Laffont, J. J. and Tirole, J. (1986), 'Using cost observation to regulate firms', *The Journal of Political Economy* **94**, 614–641.

Lavergne, P. and Thomas, A. (2005), 'Semiparametric estimation and testing in a model of environmental regulation with adverse selection', *Empirical Economics* **30**, 171–192.

Lazear, E. (2000), 'Performance pay and productivity', *American Economic Review* **90**, 1346–1361.

Leslie, P. (2004), 'Price discrimination in broadway theater', *The RAND Journal of Economics* **35**, 520–541.

Lewis, T. and Sapppington, D. (1995), 'Optimal capital structure in agency relationships', *The RAND Journal of Economics* **26**, 343–361.

Manning, W., Newhouse, J., Duan, N., Keeler, E. and Leibowitz, A. (1987), 'Health insurance and the demand for medical care: Evidence from a randomized experiment', *American Economic Review* **77**, 251–277.

Maskin, E. and Riley, J. (1984), 'Monopoly with incomplete information', *The RAND Journal of Economics* **15**, 171–196.

Miravete, E. J. (2002), 'Estimating demand for local telephone service with asymmetric information and optional calling plans', *Review of Economic Studies* **69**, 943–971.

Miravete, E. J. and Roller, L.-H. (2005), 'Estimating markups under nonlinear pricing competition', *Journal of the European Economic Association* **2**, 526–535.

Mussa, M. and Rosen, S. (1978), 'Monopoly and product quality', *Journal of Economic Theory* **18**, 301–317.

Myerson, R. B. (1979), 'Incentive compatibility and the bargaining problem', *Econometrica* **47**, 61–73.

Myerson, R. B. (1981), 'Optimal auction design', *Mathematics of Operation Research* **6**, 58–73.

Navas, A. (2009), Groups of circle diffeomorphisms. Mimeo.

Paarsch, L. (1992), 'Deciding between the common and private value paradigms in empirical models of auctions', *Journal of Econometrics* **51**, 191–215.

Paarsch, L. and Shearer, B. (2000), 'Piece rates, fixed wages and incentive effects: Statistical evidence from payroll records', *International Economic Review* **41**, 59–92.

Perrigne, I. (2002), Incentive regulatory contracts in public transportation: An empirical study. Working Paper, Pennsylvania State University.

Perrigne, I. and Vuong, Q. (2008), Nonparametric identification of incentive regulation models. Mimeo.

Perrigne, I. and Vuong, Q. (2010), Nonlinear pricing in yellow pages. Mimeo.

Pioner, H. M. (2009), Semiparametric identification of multidimensional screening models. Mimeo.

Poincaré, H. (1885), 'Sur les courbes définies par les équations différentielles', *Journal de Mathématiques* **4**.

Stillwell, J. (1992), *Geometry of Surfaces*, Springer Verlag.

Wolak, F. (1994), 'An econometric analysis of the asymmetric information, regulator-utility interaction', *Annales d'Economie et de Statistique* **34**, 13–69.