

## ESTIMATION OF DYNAMIC PROGRAMMING MODELS WITH CENSORED DEPENDENT VARIABLES

VICTOR AGUIRREGABIRIA  
*The University of Western Ontario*

*This paper considers the estimation of dynamic structural models where the decision variables are censored. We present and discuss several econometric issues and estimation methods under alternative stochastic structures of the unobservables, different potential sources of censoring, and different characteristics of the dataset (e.g., temporal dimension, frequency of corner solutions, or distribution of duration spells between two consecutive interior solutions). A labor demand model with kinked and lump-sum hiring and firing costs is used to illustrate the econometric problems and estimation methods. (JEL C34, C15, J23)*

### 1. Introduction

Longitudinal datasets with information about firms investment, employment or price decisions, among others, show how frequently firms do not respond to observed changes in costs or in the demand, and they prefer “doing nothing”. Retail firms may not change their prices for several months even if wholesale prices are changing. Firms wait to renovate their capital stock even when technological change is rapidly depreciating this stock, or when significant reductions in real interest rates occur. Employment in a plant may not be reduced even

This paper is a revised version of Chapter 3 of my thesis. I would like to thank comments from Namkee Ahn, César Alonso, Emilio Cerdá, Zvi Eckstein, Albert Marcet, Costas Meghir, Bob Miller, Pedro Mira, and Alfonso Novales, as well as from two anonymous referees and the editor of this journal. I am specially grateful to my thesis advisor Manuel Arellano.

if the plant is suffering significant and persistent reductions in its demand. This lack of response to sizable changes in relevant variables can be also observed in individuals' purchasing decisions of durable goods. From a theoretical point of view there are several potential characteristics of the decision problem that can explain the existence of this censoring in observed decision variables: non-negativity constraints, partial irreversibility of the decision, kinked adjustment costs, indivisibilities, or lump-sum adjustment costs, among others. Each of these sources of censoring can have different economic interpretations for each particular decision problem, e.g., regulatory or technological restrictions, or market conditions in which the agents operate.

In this paper we analyze the identification and estimation of the sources of censoring in dynamic structural models. The paper discusses different econometric issues associated to the estimation of these models and presents several approaches to overcome some of these problems. Special emphasis is placed on the identification and estimation of the parameters of interest under different assumptions on the stochastic structure of the unobservables and under different characteristics of the longitudinal dataset, like its temporal dimension, the frequency of corner solutions, or the distribution of duration spells between two consecutive interior solutions.

There are several reasons that motivate the use of a dynamic structural approach to analyze the economic implications of infrequent adjustments in individuals' decisions. First, in the presence of any of the different potential sources of censoring, an individual's decision problem becomes dynamic, even if in the absence of these factors the problem is static. Second, in the context of an axiomatic approach to individual behavior, both identification assumptions and parameters of interest have clear economic or behavioral interpretations. This makes it easier to discuss the plausibility of the assumptions or of the parameter estimates. Finally, but not less important, the estimation of a structural model allows one to implement counterfactual experiments to evaluate the effects of changes in institutional or behavioral parameters entering in the model. Section 2 characterizes the type of models that will be considered in this paper and presents

the notation. In that section we also present a model of labor demand that will be used to illustrate methods and results.

There are at least two types of empirical questions that can be answered from the estimation of the sources of censoring in a dynamic decision model. First, different sources of censoring in a decision variable are associated to different institutional characteristics of the market under study. To illustrate this, consider an individual's purchasing decision of an automobile (see Bar-Ilan and Blinder, 1988, or Eberly, 1994, for examples of these models). A simple potential explanation for individuals' purchasing infrequency is the existence of indivisibilities. If that is the unique source of censoring, there are not borrowing constraints, and there are perfect competitive markets for purchasing and renting cars, an individual would be indifferent between purchasing a new car or renting it during its operative life. In such a case, purchasing infrequency would not have any implication on individuals' consumption behavior. Alternatively, consider that there exist significant informational asymmetries in the market of second-hand cars. These asymmetries imply that the price of a car in the second-hand market is lower than the price of exactly the same car in the market of new cars. This introduces an additional source of infrequency in individuals' purchasing behavior (i.e., partial irreversibility). Another possible source of censoring would be the existence of search costs or administrative costs that could persist even in the absence of informational asymmetries or indivisibilities. It is clear that the identification of the quantitative importance of these different factors affecting infrequent purchasing behavior is of significant interest both for firms operating in the market and for regulators.

A second aspect that motivates the interest in the identification of the sources of censoring is that these sources have different implications on individual behavior and on aggregate cross-sectional dynamics of the dependent variable. For example, kinked adjustment costs reduce the time variability of the dependent variable, but lump-sum adjustment costs contribute to increase this volatility. Therefore, the implications of these two factors on the propagation of the business cycle can be very different (see Caballero and Engel, 1991, and Bertola and

Caballero, 1990, for the implications of kinked adjustment costs on aggregate dynamics).

Since the seminal paper by Hansen and Singleton (1982), the most common approach to estimate dynamic decision models has been to construct sample counterparts of the orthogonality conditions provided by the stochastic Euler equations, and use them to estimate the parameters of interest by the Generalized Method of Moments (GMM). The main advantage of this estimation strategy is that it permits to identify and estimate parameters of interest without having to solve the model. However, the econometric approach in Hansen and Singleton has at least three limitations in the context of censored dynamic decision models. First, marginal conditions of optimality hold only at interior solutions, and using the subsample of interior solutions introduces a selection bias. Second, not all the structural parameters can be identified exploiting only the structure in the marginal conditions of optimality (i.e., lump-sum adjustment costs). Finally, if corner solutions are relatively frequent in the sample, the discrete choice “corner solution versus interior solution” can contain more information than the marginal conditions of optimality about most of the structural parameters.

An alternative approach to estimate dynamic structural models is to employ a solution estimation method. This method uses an outer algorithm to maximize a criterion function (e.g., likelihood) and an inner algorithm to solve the dynamic programming model at each iteration in the search for the parameter estimates. This econometric approach was first used by Miller (1984), Pakes (1986) and Rust (1987) in the context of discrete choice dynamic structural models, and it has been employed since then in several discrete choice applications (see Eckstein and Wolpin, 1989, and Rust, 1992 and 1994, for excellent surveys). This solution estimation approach has also been used in applications with relatively simple continuous decision models (Deaton and Laroque, 1996). The main advantage of this approach is that it incorporates simultaneously all the restrictions or the structure of the model in the estimation process. Its main disadvantage is the large computational cost when the number of state variables is relatively large (i.e., larger than 2 or 3). In spite of the notorious

advances during the last years in the techniques for the solution and estimation of these models (as well as the improvements in computer equipment), there are still many interesting dynamic decision models in economics that cannot be estimated using this solution estimation method.

In this paper we concentrate on those econometric techniques that do not require one to obtain an explicit solution of the model, and on their application to censored dynamic structural models. In Section 3 we obtain the form of the Euler equations in censored dynamic decision models, and discuss how to obtain moment conditions from these Euler equations in order to estimate some of the parameters of interest. There are two main econometric problems that should be taken into account to obtain these moment conditions. The first one is to control for the sample selection bias that results from the fact that marginal conditions of optimality only hold for those observations at interior solutions. The second problem is that, when unobservables are autocorrelated, current and previous values of predetermined explanatory variables will be correlated with the unobservables. In that case the problem is how to obtain a transformation of the Euler equations that guarantees orthogonality between some observable explanatory variables and the transformed unobservables. Although this is a standard problem in time-series and panel data econometric models, both the existence of censoring and the particular way in which unobservables enter in these Euler equations introduce some additional considerations.

Section 3 builds on Pakes (1994) and extends his results in two directions. First, we consider a general censored dynamic programming model. And second, we present relatively simple ways to control for selection bias and autocorrelated unobservable state variables in these models that do not integration over the space of unobservables. We finish Section 3 with a discussion of several limitations associated to the estimation of censored dynamic decision models using only Euler equations. Some of these limitations motivate the importance of exploiting the discrete choice *corner solution versus interior solution* to estimate the parameters of these models.

Section 4 discusses the econometric problems associated to the esti-

mation of the optimal discrete choice *corner solution versus interior solution* using methods that do not require the explicit solution of the model. This section is based on Aguirregabiria (1996), and shows how a method in the spirit of Hotz and Miller (1993) can be applied to the estimation of this discrete choice in a censored dynamic programming model. This approach consists of using individuals' decisions to obtain nonparametric or semiparametric estimates of the unknown functions that are needed to have a solution of the model, i.e., value function or conditional choice value functions. Given these initial estimates, the structural parameters can be estimated in a second stage without having an explicit solution of the decision model. This approach has been used by Hotz and Miller (1993), Manski (1991 and 1993), Hotz et al. (1994), and Ahn (1995) in the context of structural dynamic discrete choice models, by Aguirregabiria (1996) in censored dynamic decision models, and by Park (1996) in dynamic games.

We conclude in Section 5 summarizing the paper and with a brief discussion of several unsolved problems in this literature.

## 2. Continuous Markov decision process with censored decision variables

This section presents the notation that will be used for the rest of the paper, and describes some general characteristics of censored continuous Markov decision processes (hereinafter, censored CMDP). We start with a brief description of a general CMDP, and then we discuss several sources of censoring in these models.

### 2.1. Continuous Markov decision processes

There are two types of variables in a MDP: the vector of state variables,  $s$ , and a control variable,  $d$ .<sup>1</sup> In a CMDP the control variable has a continuous range of variation. Let  $D$  be the set of feasible choices, where  $D \subseteq \mathfrak{R}$  and  $D$  contains a compact set in  $\mathfrak{R}$ , e.g.,

<sup>1</sup>In this paper we consider a model with only one decision variable. However, all the results below can be extended to the case of multiple decision variables. For instance, Pakes (1994) considers the case of one binary choice variable and one continuous variable, i.e., indicator of firms' liquidation and investment, respectively. Aguirregabiria (1996) presents a model with two censored decision variables, i.e., prices and orders.

$D = \Re$  or  $D = \{0\} \cup [1, +\infty)$ . The vector of state variables belongs to the state space  $S$  that, for notational simplicity, we assume is a compact subset of the  $\Re^{|S|}$ , where  $|S|$  is the dimension of  $s$  (i.e., we assume that the state variables are also continuous).<sup>2</sup> Time is discrete and indexed by  $t$ . At each period  $t$  a decision-maker or agent observes  $s_t$  and decides  $d_t$  in order to maximize the expected value of current and future returns. Future values of some state variables are uncertain for the decision-maker, but she has subjective beliefs about uncertain future states. These beliefs can be represented by a Markov transition density function  $p(s_{t+1}|s_t, d_t)$ . Preferences are time separable and  $u(s_t, d_t)$  represents the one-period return or utility function. The time horizon of the decision problem is infinite. The parameter  $\beta$  represents the rate at which the agent discounts utility at future periods, and it belongs to the interval  $(0, 1)$ .

In this context, an agent can be represented by the set of primitives  $\{u, p, D, \beta\}$ . We assume that these primitives are known functions of a vector of parameters  $\theta$ . The decision problem of an individual at period  $t$  is:

$$\begin{aligned} \text{Max}_{\{d_t \in D\}} \quad & u(s_t, d_t; \theta) \\ & + \sum_{j=1}^{\infty} \beta^j \int u(s_{t+j}, d_{t+j}; \theta) \\ & \prod_{k=1}^j p(ds_{t+k}|s_{t+k-1}, d_{t+k-1}; \theta), \end{aligned} \quad [1]$$

given that all future decisions,  $\{d_{t+1}, d_{t+2}, \dots\}$ , will be optimally taken, and where the integral is over all future state variables. Let  $\delta_t(s_t; \theta)$  be the optimal decision rule at period  $t$  (i.e., the argmax of the expression in [1]), and define the value function at period  $t$ ,  $V_t(s_t; \theta)$ , as the discounted expected value of current and future utilities when  $d_{t+j} = \delta_{t+j}(s_{t+j}; \theta)$  for every  $j > 0$  (i.e., the value of expression [1]). The Markov structure of the transition probabilities, the time-separability of preferences, and the infinite horizon of the problem

<sup>2</sup>We might consider both continuous and discrete state variables. However, we have preferred to consider only continuous state variables to avoid having to define both transition density functions and transition probability functions, as well as to avoid the combination of integrals and sums.

imply that both value functions and optimal decision rules are time-invariant (i.e., Blackwell's theorem). Using the definition of value function we can write the following recursive functional equation for  $V(s_t; \theta)$  (i.e., Bellman's equation):

$$V(s_t; \theta) = \max_{\{d_t \in D\}} \left\{ u(s_t, d_t; \theta) + \beta \int V(s_{t+1}; \theta) p(ds_{t+1} | s_t, d_t; \theta) \right\}. \quad [2]$$

And the optimal decision rule is:

$$\delta(s_t; \theta) = \arg \max_{\{d_t \in D\}} \left\{ u(s_t, d_t; \theta) + \beta \int V(s_{t+1}; \theta) p(ds_{t+1} | s_t, d_t; \theta) \right\}. \quad [3]$$

Define the *conditional value function* as the expected value of next period value function conditional to current decision and current state variables:

$$EV(s, d; \theta) \equiv \int V(s'; \theta) p(ds' | s, d; \theta). \quad [4]$$

And define  $G(s, d; \theta) \equiv u(s, d; \theta) + \beta EV(s, d; \theta)$ . Then, we can also define  $V(s; \theta)$  and  $\delta(s; \theta)$  as the maximum value and the argmax, respectively, of  $G(s, d; \theta)$  with respect to  $d \in D$ .

We can distinguish two types of state variables in a CMDP: endogenous and exogenous state variables. Endogenous state variables are those with transition rules that depend on the decision variable. On the contrary, the transition rules of the exogenous state variables do not depend on previous decisions. Let  $s_t$  be equal to  $(k_t, z_t)$ , where  $k_t$  is an endogenous state variable, and  $z_t$  is a vector of exogenous state variables. By definition  $p(z_{t+1} | s_t, d_t)$  does not depend on  $k_t$  and  $d_t$ , and we denote this function by  $q_z(dz_{t+1} | z_t)$ . For simplicity we will consider the following transition rule for the endogenous state variable:

$$k_{t+1} = \lambda k_t + d_t + h(z_{t+1}), \quad [5]$$

where  $|\lambda| < 1$ . Notice that, conditional to the information at period  $t$ , this transition is stochastic because it depends on  $z_{t+1}$ .

## 2.2. Sources of censored decision variables in CMDP

DEFINITION 1: Given a pair  $(s, \theta)$ , the optimal decision associated to this pair,  $\delta(s; \theta)$ , is a corner solution if:

$$G_{d[\rightarrow]}(s, \delta[s; \theta]; \theta) \neq 0 \quad \text{and} \quad G_{d[\leftarrow]}(s, \delta[s; \theta]; \theta) \neq 0, \quad [6]$$

where  $G_{d[\rightarrow]}(\cdot)$  and  $G_{d[\leftarrow]}(\cdot)$  are the partial derivatives of  $G(\cdot)$  with respect to  $d$  from the left and from the right, respectively.<sup>3</sup>

DEFINITION 2: Given a CMDP with vector of structural parameters  $\theta$ , we say that this process has corner solutions if:

$$\int I \left( G_{d[\rightarrow]}(s, \delta[s; \theta]; \theta) \neq 0 \quad \text{and} \right. \\ \left. G_{d[\leftarrow]}(s, \delta[s; \theta]; \theta) \neq 0 \right) P(ds; \theta) > 0, \quad [7]$$

where  $I(\cdot)$  is the indicator function, and  $P(ds; \theta)$  is the unconditional or steady-state distribution of  $s$  induced by the optimal decision rule  $\delta[s; \theta]$  and the conditional transition probabilities of the state variables.

In this paper we concentrate on corner solutions that imply a decision of inaction, that is we concentrate on *censored decision models*. Without loss of generality it is possible to define *inaction* as  $d = 0$ . Therefore, I consider a class of CMDP where the decision variable is censored at zero.

DEFINITION 3: A CMDP with a vector of structural parameters  $\theta$  has censored decision variable at zero if:

$$\int I \left( \delta(s; \theta) = 0 ; G_{d[\rightarrow]}(s, \delta[s; \theta]; \theta) \neq 0 ; \right. \\ \left. G_{d[\leftarrow]}(s, \delta[s; \theta]; \theta) \neq 0 \right) P(ds; \theta) > 0. \quad [8]$$

The rest of this subsection analyzes four sources of censoring which appear in several CMDP in economics: (1) non-negativity constraints;

<sup>3</sup>Notice that this definition accounts for both standard corner solutions and solutions at values  $d$  where  $G(\cdot)$  is not differentiable.

(2) partial irreversibility and kinked adjustment costs; (3) lump-sum adjustment costs; and (4) indivisibilities.

### *Non-negativity constraints*

Non-negativity restrictions, i.e.,  $D = [0, \infty)$ , are the result of institutional or physical constraints. However, non-negativity restrictions do not always imply corner solutions. If, for any value of the vector of state variables, the marginal utility with respect to  $d$  is equal to infinity at  $d = 0$ , there will not be a positive probability of corner solution. That is the case in lifetime consumption models (or in models of firms' capital or employment decisions) where consumption (capital or employment) cannot be negative, but marginal utility (profit) is infinite at zero consumption (capital or employment). However, if the marginal utility with respect to  $d$  is not always equal to infinity at  $d = 0$ , the non-negativity restriction will imply corner solutions. That is the case in models of capital investment with total irreversibility, like in Pindyck (1988). The following Lemma characterizes the form of the optimal decision rule in this type of models.

LEMMA 1: *If:*

- (i)  $D = [0, \infty)$ ,
- (ii)  $u(\cdot)$  and  $p(\cdot)$  are continuous and twice differentiable in all their arguments,
- (iii)  $u_d(s, 0; \theta) < \infty$ , for any value of  $s$ ,
- (iv)  $u(\cdot)$  is strictly concave in  $d$  and in the endogenous state variables.

the optimal decision rule of the CMDP has the following form:

$$\delta(s; \theta) = \begin{cases} \delta^*(s; \theta) & \text{if } \delta^*(s; \theta) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad [9]$$

where  $\delta^*(s; \theta)$  is the optimal interior solution that is implicitly defined by:

$$G_d(s, \delta^*[s; \theta]; \theta) = 0, \quad [10]$$

and  $G_d \equiv \partial G / \partial d$ .

PROOF: It is simple to verify that if  $G(s, d)$  is continuous and strictly concave with respect to  $d$ , Lemma 1 holds. Under conditions (ii) to (iv) and  $\beta \in (0, 1)$ , Theorem 9.10 in Stokey and Lucas (1989) applies, and the value function is continuously differentiable. Therefore, the proof of Lemma 1 reduces to prove that  $G_{dd}(s, d) < 0$  for any pair  $(s, d)$ . Now:

$$G_{dd}(s, d) = u_{dd}(s, d) + \beta \int V_{kk}(\lambda k + d + h[z'], z') q_z(dz'|z), \quad [11]$$

where  $V_{kk} \equiv \partial^2 V / \partial k^2$ . Under conditions (i) to (iv) Theorem 9.8 in Stokey and Lucas applies, and the value function of this problem is strictly concave. Therefore, given equation [11],  $G(\cdot)$  is also strictly concave with respect to  $d$ .  $\square$

The optimal decision rule in Lemma 1 contains the main characteristics of the decision rules in censored CMDP. There is a first order condition of optimality for the interior solution (equation [10]), and there is a discrete choice between corner solution and interior solution. For most dynamic decision models there are not closed form expressions for equations [9] and [10] in terms of variables and structural parameters. The estimation methods that we will discuss in this paper are characterized by the fact that, without solving explicitly the model, they exploit its structure to obtain closed form expressions, in terms of variables and structural parameters, which are alternative representations of the marginal conditions of optimality and of the optimal discrete choice, and use these expressions to construct moment conditions to estimate the parameters of the model.

#### *Partial irreversibility and kinked adjustment costs*

We have partial irreversibility of a decision when there is a cost associated to negative values of the decision variable but to positive values. This cost creates a kink in the utility function at  $d = 0$ . This kink implies the optimality of inaction, i.e.,  $d = 0$ , for a subset of values of  $s$  that has a positive mass of probability. Partial irreversibility or kinked adjustment costs can arise in many economic problems. For example, in models of capital investment or consumption of durable goods, the existence of informational asymmetries in the second-hand market and the fact that capital equipment and durable goods are, in

a certain degree, specific to the firm and the consumer, respectively, make the selling price of the good lower than the purchasing price. The implicit cost of a negative value of  $d$  is the difference between the purchasing price and the selling price. Examples of this partial irreversibility models are Eberly (1994) for consumption of durable goods, and Abel and Eberly (1996) for capital investment. Another type of models are those with an explicit cost for negative values of  $d$ . That is the case in labor demand models with firing costs (Hamermesh, 1989 and 1993, Bentolila and Bertola, 1990, or Hopenhayn and Rogerson, 1993).

Here I consider a simple representation of this type of kinked adjustment costs.

$$u(s, d; \theta) = u^*(s, d; \theta) + c^L I(d < 0)d - c^H I(d > 0)d, \quad [12]$$

where  $u^*(\cdot)$  is continuous and twice differentiable in all its arguments;  $I(\cdot)$  is the indicator function; and  $c^L$  and  $c^H$  are positive parameters representing the cost of negative and positive adjustments, respectively. Lemma 2 characterizes the form of the optimal decision rule in this model.

LEMMA 2: *If:*

- (i)  $u(s, d; \theta) = u^*(s, d; \theta) + c^L I(d < 0)d - c^H I(d > 0)d$ ,
  - (ii)  $u^*(\cdot)$  and  $p(\cdot)$  are continuous and twice differentiable in all their arguments,
  - (iii)  $u^*(\cdot)$  is strictly concave in  $d$  and in the endogenous state variables,
- the optimal decision rule of the CMDP has the following form:

$$\delta(s; \theta) = \begin{cases} \delta^H(s; \theta) & \text{if } \delta^H(s; \theta) > 0 \\ \delta^L(s; \theta) & \text{if } \delta^L(s; \theta) < 0 \\ 0 & \text{otherwise,} \end{cases} \quad [13]$$

where  $\delta^H(s; \theta)$  is the optimal interior solution when  $d > 0$  and  $\delta^L(s; \theta)$  is the optimal interior solution when  $d < 0$ , and they are implicitly defined by:

$$u_d^*(s, \delta^H[s; \theta]; \theta) + \beta EV_d(s, \delta^H[s; \theta]; \theta) = c^H, \quad [14]$$

and

$$u_d^*(s, \delta^L[s; \theta]; \theta) + \beta EV_d(s, \delta^L[s; \theta]; \theta) = -c^L, \quad [15]$$

with  $u_d^* \equiv \partial u^* / \partial d$ , and  $EV_d \equiv \partial EV / \partial d$ .

The proof of this Lemma is very similar to the previous proof of Lemma 1. In particular, under conditions (ii) and (iii), the non differentiability of  $u(\cdot)$  at  $d = 0$  does not affect the continuity and strict concavity of  $G$  (i.e., Theorems 9.8 and 9.10 in Stokey and Lucas still hold).

#### *Lump-sum adjustment costs*

Lump-sum adjustment costs are costs which do not depend on the amount of the adjustment. They introduce a discontinuity in the one-period utility function that can generate censoring in the decision variable. These costs may arise in different economic problems. When a retailer places an order to a manufacturer, she faces some costs that are not proportional to the amount of the order; e.g., transportation costs and costs of organizing the new deliveries into the store (see Blinder, 1980). Other examples are models of price adjustment with menu costs (Sheshinski and Weiss, 1983), and models of capital replacement (Rust, 1987, or Cooper, Haltinwanger and Power, 1995). Here we consider the following specification with symmetric lump-sum adjustment costs (the extension to asymmetric adjustment costs is straightforward):

$$u(s, d; \theta) = u^*(s, d; \theta) - F I(d \neq 0), \quad [16]$$

where  $F$  is a positive parameter representing the lump-sum cost. Lemma 3 characterizes the form of the optimal decision rule in this model.

LEMMA 3: *If:*

- (i)  $u(s, d; \theta) = u^*(s, d; \theta) - F I(d \neq 0)$ ,
- (ii)  $u^*(\cdot)$  and  $p(\cdot)$  are continuous and twice differentiable in all their arguments,
- (iii)  $u^*(\cdot)$  is strictly concave in  $d$  and in the endogenous state variables,

the optimal decision rule of the CMDP has the following form:

$$\delta(s; \theta) = \begin{cases} \delta^*(s; \theta) & \text{if } \gamma(s; \theta) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad [17]$$

where  $\delta^*(s; \theta)$  is the optimal interior solution implicitly defined by:

$$u_d^*(s, \delta^*[s; \theta]; \theta) + \beta EV_d(s, \delta^*[s; \theta]; \theta) = 0, \quad [18]$$

and  $\gamma(s; \theta)$  is a threshold function that indicates when it is optimal to choose the interior solution:

$$\begin{aligned} \gamma(s; \theta) = & \{u^*(s, \delta^*[s; \theta]; \theta) - u^*(s, 0; \theta)\} \\ & + \beta \{EV(s, \delta^*[s; \theta]; \theta) - EV(s, 0; \theta)\} - F. \end{aligned} \quad [19]$$

The proof of this Lemma is a little more complicated than the proofs of Lemmas 1 and 2. In particular, the discontinuity of the one-period utility function implies that the value function is not concave. The proof of this Lemma exploits some properties of  $K$  - concave functions. See Scarf (1959) or Bertsekas (1976) for proofs of this Lemma.

#### *Indivisibilities*

Consider a firm who has to decide each year how much to spend in computer equipment, and let  $\bar{d}$  be the minimum price of a computer in the market. Assume that, in the absence of a minimum price, the optimal decision is to spend an amount lower than  $\bar{d}$ . If that is the case, once the restriction of the minimum price  $\bar{d}$  is taken into account, the optimal decision of the firm might be not to spend any positive amount in computers during that year. That is, the discontinuity in the choice set can generate optimal decision rules where inaction is optimal for some values of the state variables. The choice set of the decision maker is:

$$D = \{0\} \cup [\bar{d}, +\infty), \quad [20]$$

where  $\bar{d} > 0$  is the minimum feasible value of a positive adjustment. Lemma 4 characterizes the optimal decision rule in this decision problem.

LEMMA 4: *If:*

- (i)  $D = \{0\} \cup [\bar{d}, +\infty)$  with  $\bar{d} > 0$ ,
  - (ii)  $u(\cdot)$  and  $p(\cdot)$  are continuous and twice differentiable in all their arguments,
  - (iii)  $u(\cdot)$  is strictly concave in  $d$  and in the endogenous state variables,
- the optimal decision rule of the CMDP has the following form:

$$\delta(s; \theta) = \begin{cases} \delta^*(s; \theta) & \text{if } \delta^*(s; \theta) > \bar{d} \\ \bar{d} & \text{if } 0 < \delta^*(s; \theta) < \bar{d} \text{ and } \gamma(s; \theta) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad [21]$$

where  $\delta^*(s; \theta)$  is the optimal interior solution implicitly defined by:

$$u_d^*(s, \delta^*[s; \theta]; \theta) + \beta EV_d(s, \delta^*[s; \theta]; \theta) = 0, \quad [22]$$

and  $\gamma(s; \theta)$  is a threshold function that indicates when it is optimal to choose the minimum value  $\bar{d}$ , and it is defined by:

$$\begin{aligned} \gamma(s; \theta) = & \{u^*(s, \bar{d}; \theta) - u^*(s, 0; \theta)\} \\ & + \beta\{EV(s, \bar{d}; \theta) - EV(s, 0; \theta)\}. \end{aligned} \quad [23]$$

The proof of Lemma 4 is very similar to the proofs of Lemmas 1 and 2.

### 2.3. An example: A model of labor demand

Consider a risk neutral firm that at each period  $t$  decides the amount of employment in order to maximize the expected discounted stream of current and future profits. Current profits are equal to cash-flow net of adjustment costs:

$$u(s_t, d_t; \theta) = f(l_t, d_t, \epsilon_t; \theta_f) - (l_t + d_t)w_t - AC(d_t; \theta_{AC}), \quad [24]$$

where  $f(\cdot)$  is the production function;  $AC(\cdot)$  is the adjustment costs function;  $l_t$  is the number of employees at the beginning of period  $t$ ;  $d_t$  is the amount of hired workers (if positive) or fired workers (if negative) during period  $t$ ;  $\epsilon_t$  is a productivity shock observable to the firm at the beginning of period  $t$ ;  $w_t$  is the real wage; and  $\theta_f$  and  $\theta_{AC}$

are vectors of parameters in the production function and in the cost function, respectively.

The real wage is determined in a competitive labor market, and it is exogenous for the firm. Both  $w_t$  and  $\epsilon_t$  follow strictly exogenous first order Markov processes with transition densities  $p_w(w_{t+1}|w_t; \theta_w)$  and  $p_\epsilon(\epsilon_{t+1}|\epsilon_t; \theta_\epsilon)$ , respectively. The transition rule for the number of employees is:

$$l_{t+1} = \lambda l_t + d_t + \xi_{t+1}, \quad [25]$$

where  $(1 - \lambda)$  represents the average rate of voluntary quits of workers from the firm, and  $\xi_{t+1}$  is a shock in the number of retirements, that is observable to the firm at the beginning of period  $t + 1$  but not at period  $t$  (i.e., stochastic transition rule). We assume that  $\{\xi_t\}$  is strictly exogenous and follows a first order Markov process. In this context, the vector of state variables at period  $t$  is  $s_t = (l_t, w_t, \epsilon_t, \xi_t)'$ .

The production function is continuous, twice differentiable and increasing in all the arguments, and strictly concave with respect to  $l$  and  $d$ . Adjustment costs have the following structure:

$$AC(d_t; \theta_{AC}) = I\{d_t < 0\}(F^L - c^L d_t) + I\{d_t > 0\}(F^H + c^H d_t), \quad [26]$$

where  $F^L$  and  $F^H$  represent lump-sum firing and hiring costs, respectively, and  $c^L$  and  $c^H$  represent linear firing and hiring costs, respectively. In this model, the optimal decision rule has the following form:

$$\delta(s; \theta) = \begin{cases} \delta^L(s; \theta) & \text{if } \delta^L(s; \theta) < 0 \text{ and } \gamma^L(s; \theta) > 0 \\ \delta^H(s; \theta) & \text{if } \delta^H(s; \theta) > 0 \text{ and } \gamma^H(s; \theta) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad [27]$$

where  $\theta$  is the vector of structural parameters  $(\beta, \theta_f, \theta_w, \theta_\epsilon, F^L, F^H, c^L, c^H)'$ . We will use this model to illustrate several results and estimation methods.

### 3. Estimating Euler equations in censored CMDP

Consider a censored CMDP and a longitudinal dataset containing the following information:  $\{d_{it}, x_{it} : i = 1, \dots, N; t = 1, \dots, T_i\}$ , where  $x$  is the subvector of  $s$  that is observable to the econometrician.

That is,  $s = (x, \epsilon)$ , where  $x$  are state variables observable to the econometrician and  $\epsilon$  are unobservable from the point of view of the econometrician.<sup>4</sup> We are interested in using this information to estimate the vector of structural parameters  $\theta$ .

Since the seminal paper by Hansen and Singleton (1982) the most common approach to estimate dynamic decision models with continuous decision variables has been to construct sample counterparts of the orthogonality conditions provided by the stochastic Euler equations and use them to estimate the parameters of interest by the Generalized Method of Moments (GMM). Although standard Euler equations do not hold when there is a positive probability of corner solution, we follow Pakes (1994) and show that it is still possible to obtain a type of Euler equations that account for the existence of censoring in the decision variable. We describe how to exploit these Euler equations to obtain estimates of some parameters of the model. Finally, we discuss several limitations of this approach. These limitations motivate the use of the method described in Section 4 that exploits the information in the discrete choice *corner solution versus interior solution*.

### 3.1. Euler equations in censored CMDP

Standard Euler equations are the result of combining marginal conditions of optimality at two consecutive periods. They are obtained under the assumption that, conditional to the information at period  $t$ , marginal conditions of optimality (i.e., interior solutions) will hold with probability one at period  $t + 1$ . Nevertheless, for the class of models that we consider in this paper, interior solutions do not occur at each period with probability one, and when an agent makes her decision at period  $t$  she assigns a non-zero probability to the event

<sup>4</sup>The consideration of unobservable state variables seems a natural and realistic way of introducing unobservables in these models. However, there are at least another two potential sources of unobservables or “error terms” in most econometric models: measurement errors in variables and approximation errors. The rest of this paper assumes that the postulated model is the “true model” and there are not specification or approximation errors. However, I will discuss below some econometric issues related to the existence of measurement errors in some of the observed state variables (see Rust, 1994, p. 3100-01, for a discussion of alternative models of the “error term” in dynamic decision models).

“existence of corner solution at period  $t + 1$ ”. Therefore, standard Euler equations do not hold in censored CMDP.

However, it is still possible to obtain Euler equations in censored CMDP. We follow the approach in Pakes (1994) who obtains the Euler equation in a model with both a continuous and a discrete control variable. Here we obtain the expression of the Euler equation for the case of non-negativity constraints. Euler equations for the other three cases can be obtained using the same procedure, and we present an example using the labor demand model that we have described in Subsection 2.3.

Consider that the optimal decision at period  $t$  is an interior solution, i.e.,  $d_t = \delta^*(s_t)$ , and define  $\tau_t$  as the number of periods until the next interior solution. The variable  $\tau_t$  is unknown to the decision-maker at period  $t$ . Now, consider the following decision rule, where  $\alpha$  is a real value:

$$\Lambda(s_{t+j}, \alpha) = \begin{cases} \Lambda^*(s_{t+j}, \alpha) & \text{if } \Lambda^*(s_{t+j}, \alpha) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad [28]$$

where:

$$\Lambda^*(s_{t+j}, \alpha) = \begin{cases} \delta^*(s_{t+j}) - \alpha & \text{for } j = 0 \\ \delta^*(k_{t+j} + \lambda^{\tau_t-1}\alpha, z_{t+j}) + \lambda^{\tau_t-1}\alpha & \text{for } j = \tau_t \\ \delta^*(s_{t+j}) & \text{otherwise,} \end{cases} \quad [29]$$

where  $\lambda$  is the parameter in the transition rule of the endogenous state variable (equation [5]).

Notice that, for  $j \in \{1, \dots, \tau_t-1\}$ ,  $\Lambda^*(s_{t+j}, \alpha) = \delta^*(k_{t+j} - \lambda^{j-1}\alpha, z_{t+j})$ . Since  $|\lambda| < 1$  and  $\delta^*(\cdot)$  is a continuous function, for  $\alpha$  close to zero we have that  $\delta^*(k_{t+j} - \lambda^{j-1}\alpha, z_{t+j}) < 0$  (because  $\delta^*(k_{t+j}, z_{t+j}) < 0$ ) and thus  $\Lambda(s_{t+j}, \alpha) = 0$  for  $j = 1, \dots, \tau_t - 1$ . For the same reasons we have that  $\Lambda(s_{t+j}, \alpha) = \Lambda^*(s_{t+j}, \alpha)$  at  $j = \tau_t$ . It is also simple to verify that, for  $j > \tau_t$ , the endogenous state variable  $\{k_{t+j}\}$  and the decision variable are the same under the optimal decision rule and under the alternative policy  $\Lambda(\cdot, \alpha)$ . Therefore, the decisions under the optimal rule  $\delta(\cdot)$  and under  $\Lambda(\cdot)$  are only different at periods  $t$  and  $t + \tau_t$ . The difference between the discounted expected stream of utilities under  $\delta(\cdot)$  and under  $\Lambda(\cdot)$  is equal to:

$$\begin{aligned}
D(\alpha, s_t) = E_t \left\{ & u(s_t, \delta_t^*) - u(s_t, \delta_t^* - \alpha) \right. \\
& + \sum_{j=1}^{\tau_t-1} \beta^j \lambda^{j-1} [u(k_{t+j}, z_{t+j}, 0) - u(k_{t+j} - \lambda^{j-1} \alpha, z_{t+j}, 0)] \\
& + \beta^{\tau_t} \lambda^{\tau_t-1} [u(s_{t+\tau_t}, \delta_{t+\tau_t}^*) - u(k_{t+\tau_t} - \lambda^{\tau_t-1} \alpha, z_{t+\tau_t}, \\
& \left. \delta^*[k_{t+\tau_t} + \lambda^{\tau_t-1} \alpha, z_{t+\tau_t}] + \lambda^{\tau_t-1} \alpha)] \right\}. \quad [30]
\end{aligned}$$

where  $\delta_{t+j}^* \equiv \delta^*[s_{t+j}]$ . This function is continuous and differentiable in  $\alpha$  and it has its minimum at  $\alpha = 0$ . Therefore, the following marginal condition of optimality should hold:

$$\begin{aligned}
E_t \left\{ & u_d(s_t, \delta_t^*) + \sum_{j=1}^{\tau_t-1} \beta^j \lambda^{j-1} u_k(s_{t+j}, 0) \right. \\
& \left. + \beta^{\tau_t} \lambda^{\tau_t-1} (u_k[s_{t+\tau_t}, \delta_{t+\tau_t}^*] - u_d[s_{t+\tau_t}, \delta_{t+\tau_t}^*]) \right\} = 0. \quad [31]
\end{aligned}$$

Notice that, if  $\tau_t = 1$  with probability one (i.e., no censored decision variable) equation [31] becomes the standard Euler equation,

$$E_t \{ u_d(s_t, \delta_t^*) + \beta (u_k[s_{t+1}, \delta_{t+1}^*] - u_d[s_{t+1}, \delta_{t+1}^*]) \} = 0. \quad [32]$$

The expression of the Euler equations for the other three sources of censoring can be derived in the same way: (1) postulate an alternative decision rule that implies the same decisions that the optimal rule except at periods  $t$  and  $t + \tau_t$ , and that is identical to the decision rule when  $\alpha = 0$ ; and (2) obtain the marginal conditions of optimality at  $\alpha = 0$ . Here we present an example for the labor demand model with kinked and lump-sum adjustment costs.

**EXAMPLE 1:** Consider the labor demand model presented in Subsection 2.3. The Euler equation for this model is:

$$\begin{aligned}
E_t \left( f_d(l_t, d_t, \epsilon_t) - w_t - c^H I(d_t > 0) + c^L I(d_t < 0) \right. \\
+ \sum_{j=1}^{\tau_t-1} \beta^j \lambda^{j-1} [f_d(l_{t+j}, 0, \epsilon_t) - w_{t+j}] \\
\left. + \beta^{\tau_t} \lambda^{\tau_t-1} [c^H I(d_{t+\tau_t} > 0) - c^L I(d_{t+\tau_t} < 0)] \right) = 0, \quad [33]
\end{aligned}$$

where  $d_t$  and  $d_{t+\tau_t}$  can be either  $\delta^H(\cdot)$  or  $\delta^L(\cdot)$ , i.e., this Euler equation holds for the two possible interior solutions, hiring or firing.

### 3.2. Inference using the Euler equations

We are interested in using the longitudinal dataset  $\{d_{it}, x_{it} : i = 1, \dots, N; t = 1, \dots, T_i\}$  to estimate the vector of structural parameters  $\theta$ . First of all, we define:

$$\begin{aligned}
\Psi(X_i^{t,t+\tau_{it}}, \epsilon_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) \equiv \\
u_d(s_{it}, d_{it}) + \sum_{j=1}^{\tau_{it}-1} \beta^j \lambda^{j-1} u_k(s_{i,t+j}, 0) \\
+ \beta^{\tau_{it}} \lambda^{\tau_{it}-1} (u_k[s_{i,t+\tau_{it}}, d_{i,t+\tau_{it}}] - u_d[s_{i,t+\tau_{it}}, d_{i,t+\tau_{it}}]), \quad [34]
\end{aligned}$$

where  $X_i^{t,t+\tau_{it}} = (x_{it}, \dots, x_{i,t+\tau_{it}})$ , and  $\epsilon_i^{t,t+\tau_{it}} = (\epsilon_{it}, \dots, \epsilon_{i,t+\tau_{it}})$ . Let  $Z_{it}$  be a vector of functions of observable state variables dated at periods  $t$  and before. Based on the Euler equation in expression [31], we can construct the following vector of orthogonality conditions:

$$E \left( Z_{it} I(d_{it} \neq 0) \Psi(X_i^{t,t+\tau_{it}}, \epsilon_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) \right) = 0, \quad [35]$$

where the indicator of interior solution, i.e.,  $I(d_{it} \neq 0)$ , appears in these conditions because equation [31] holds only for interior solutions.

However, without further assumptions, the conditions in expression [35] cannot be used to estimate  $\theta$  because they are not moment conditions in terms only of observables and structural parameters. That is, there are unobservables in  $\Psi(\cdot)$ , and these unobservables can be correlated with the conditioning variables  $Z_{it}$  and with the conditioning event " $d_{it} \neq 0$ ". Therefore, one of the econometric issues related

to the estimation of  $\theta$  using equation [35] is to obtain moment conditions in terms only of observables and parameters of interest. Some of the problems to obtain these moment conditions are common to the estimation of any type of CMDP. However, we will see below that the existence of censoring introduces some additional considerations.

We start with a simple assumption that permits to overcome the problem of unobservables in Euler equations. This assumption has been considered in several applications that estimate Euler equations, e.g., Bond and Meghir (1994) in the context of investment models, or Pfann and Palm (1993) and Alonso-Borrego (1994) in labor demand models. Let  $y_{it}$  be an observable variable that is a function of state and decision variables. This variable is neither a decision nor a state variable of the problem, and we can define it as an *outcome variable*. That could be the case of the variable *output* in some models of firms' inputs demand, *sales* in a model of price decisions, or market value of the firm in an investment model. We assume that the marginal utilities  $u_d(\cdot)$  and  $u_k(\cdot)$  depend on  $\epsilon$  only through  $y$ :

$$u_j(x, \epsilon, d; \theta) = U_j(x, d, y; \theta), \quad [36]$$

for  $j = k, d$ .

It is clear that under this assumption we can write equation [35] in terms only of observables and parameters of interest. That is:

$$E \left( Z_{it} I(d_{it} \neq 0) \tilde{\Psi}(X_i^{t,t+\tau_{it}}, Y_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) \right) = 0, \quad [37]$$

where  $\tilde{\Psi}(\cdot)$  results from substituting  $u_j(x, \epsilon, d; \theta)$  by  $U_j(x, d, y; \theta)$ ; and  $Y_i^{t,t+\tau_{it}} \equiv (y_{it}, \dots, y_{i,t+\tau_{it}})$ . The moment conditions in equation [37] can be used to estimate the structural parameters that enter in  $\tilde{\Psi}(\cdot)$ . Notice that this assumption "solves" the problem of correlation between unobservables and instrumental variables  $Z_{it}$ , as well as the selection bias that results from the correlation between the unobservables and the event " $d_{it} \neq 0$ ". The observable  $y_{it}$  is implicitly "controlling" for these problems.

Although this assumption may be plausible in some circumstances, it is important to be aware of the strong predictions associated to it. This assumption implies that  $\epsilon$  is unobservable to the econometrician only because the function that relates  $y_{it}$  with  $(x_{it}, \epsilon_{it}, d_{it})$  is

unknown, or it depends on some unknown parameters. If this function is monotonic in  $\epsilon_{it}$ , by the implicit function theorem there is a function  $\epsilon_{it} = e(x_{it}, d_{it}, y_{it})$ , and solving this function in the optimal decision rule we obtain a deterministic function that relates the observable decision variable  $d_{it}$  with the observables  $x_{it}$  and  $y_{it}$ . It is clear that this prediction will be easily rejected in most datasets because we can find different values  $d$  associated with the same value of the pair  $(x, y)$ .

Since there are many applications in which the predictions of the previous assumption are not plausible, or in which there is not an observable outcome variable, the rest of this paper considers the estimation of censored CMDP where some unobservables do not enter in the way postulated by the previous assumption. Instead, we consider the following assumption.

*ASSUMPTION 1: There are two unobservables in the model:  $\epsilon$ , that affects marginal utilities, and  $\omega$ , that affects the level of utility but not marginal utilities. The utility function is additive separable in  $(x, d, \epsilon)$  and  $\omega$ ; and marginal utilities are additive separable in  $(x, d)$  and  $\epsilon$ .*

$$u(x, \epsilon, \omega, d; \theta) = U(x, d; \theta) + d \epsilon + \theta_x x \epsilon + (F + \omega) I(d \neq 0), \quad [38]$$

where  $\theta_x$  is a subvector of parameters in  $\theta$ .

Additive separability between observables and unobservables in the utility function was first considered, in the context of structural discrete choice models, by McFadden (1973). Although this additive separability could be relaxed, it requires one to postulate a probability distribution of  $\epsilon$  conditional to  $x$ , and to integrate  $\Psi(\cdot)$  over  $\epsilon$  to obtain moment conditions in terms only of observables and structural parameters. In order to concentrate the discussion in what we think are more important econometric issues, we consider Assumption 1. This assumption is relatively general and it can be plausible in many economic applications. Furthermore, we may combine Assumption 1 with the existence of an additional unobservable that enters in  $U(\cdot)$  only through an observable outcome variable  $y$ . All the results below can be extended to that case, and we will present an example using our labor demand model. But first we discuss the previous

assumptions in the context of that model.

EXAMPLE 2: Consider the labor demand model in Subsection 2.3. If the technological shock is the only unobservable, we observe output, and the production function is Cobb-Douglass, we will be in the case of equation [36]. In that case the unobservable shock enters in the marginal profits only through the observable ratio  $y_{it}/l_{it}$ , where  $y_{it}$  is output of firm  $i$  at period  $t$ . However, there may be additional unobservables associated to labor costs and adjustment costs. Real wage can be measured with error, and adjustment costs may change over time or/and over individuals. For instance, linear adjustment costs can have the following structure:

$$c_{it}^H = c^H + \epsilon_{it}^H \quad \text{and} \quad c_{it}^L = c^L + \epsilon_{it}^L. \quad [39]$$

And we can consider a similar specification for lump-sum adjustment costs:

$$F_{it}^H = F^H + \omega_{it}^H \quad \text{and} \quad F_{it}^L = F^L + \omega_{it}^L. \quad [40]$$

It is simple to verify that this specification is consistent with Assumption 1.

We are interested in using equation [35] to obtain moment conditions in terms only of observable variables and parameters of interest. Under Assumption 1, we can decompose  $\Psi$  in two additive terms:

$$\begin{aligned} & \Psi(X_i^{t,t+\tau_{it}}, \epsilon_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) \\ &= \Psi^O(X_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) + v_{it}(\tau_{it}), \end{aligned} \quad [41]$$

where

$$\begin{aligned} \Psi^O(X_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) &= \\ & U_d(x_{it}, d_{it}; \theta) + \sum_{j=1}^{\tau_{it}-1} \beta^j \lambda^{j-1} U_k(x_{i,t+j}, 0; \theta) \\ & + \beta^{\tau_{it}} \lambda^{\tau_{it}-1} (U_k[x_{i,t+\tau_{it}}, d_{i,t+\tau_{it}}; \theta] - U_d[s_{i,t+\tau_{it}}, d_{i,t+\tau_{it}}; \theta]), \end{aligned} \quad [42]$$

and

$$v_{it}(\tau_{it}) = \epsilon_{it} + \theta_x \left( \sum_{j=1}^{\tau_{it}-1} \beta^j \lambda^{j-1} \epsilon_{i,t+j} \right) + \beta^{\tau_{it}} \lambda^{\tau_{it}-1} [\theta_x - 1] \epsilon_{i,t+\tau_{it}}. \quad [43]$$

We use the notation  $v_{it}(\tau_{it})$  to emphasize that this random variable depends on  $\tau_{it}$ . Taking into account expression [41], the moment conditions in [35] become:

$$E \left( Z_{it} I(d_{it} \neq 0) \Psi^O(X_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) \right) + E(Z_{it} I(d_{it} \neq 0) v_{it}(\tau_{it})) = 0. \quad [44]$$

Expression [44] shows that we will be able to construct moment conditions that depend only on observables and parameters of interest if the second term in the left hand side of this equation is zero, or if there is a transformation of equation [44] where this second term disappears. The solution to this econometric problem depends crucially on whether  $\epsilon_{it}$  is autocorrelated or not. Therefore, we distinguish between these two cases.

#### *Non autocorrelated unobservables*

ASSUMPTION 2: *The unobservable  $\epsilon_{it}$  is a strictly exogenous state variable, it is independently distributed over  $t$  with zero mean, and  $\{\epsilon_{it}, x_{it}\}$  are independently distributed.*

First of all, notice that, given the definition of  $v_{it}(\tau_{it})$  in equation [43], Assumption 2 implies that  $E(Z_{it} I(d_{it} \neq 0) v_{it}(\tau_{it})) = E(Z_{it} I(d_{it} \neq 0) \epsilon_{it})$ .<sup>5</sup> Second, under Assumptions 1 and 2, the construction of moment conditions in CMDP without censoring is straightforward. In these models the event “ $d_{it} \neq 0$ ” holds with probability one, and Assumption 2 implies that  $E(Z_{it} \epsilon_{it}) = 0$ . However, in the context of censored CMDP, Assumption 2 guarantees orthogonality between  $\epsilon_{it}$  and any function of  $x_{it}$ , but it is not true that  $\epsilon_{it}$  is orthogonal to  $I(d_{it} \neq 0)$  because  $d_{it}$  depends on the unobservable  $\epsilon_{it}$ . In other words, the second term in equation [44] is not zero due to the existence of selection bias. For instance, in the context of our labor demand model, if  $\epsilon_{it}$  represents an unobservable component of linear firing costs, the sample selection based on  $d_{it} \neq 0$  implies that we will tend to oversample observations from those firms with relatively small firing costs (e.g., new firms with young workers). If we do not control for this selection bias, it will probably introduce a downward bias in our estimates of firing costs.

<sup>5</sup>All the unobservables  $\epsilon_{i,t+1}, \dots, \epsilon_{i,t+\tau_{it}}$  that appear in  $v_{it}(\tau_{it})$  are independent of  $Z_{it}$  and  $d_{it}$

Therefore, under Assumptions 1 and 2 the only problem to obtain moment conditions in censored CMDP is the selection bias. In the context of non autocorrelated unobservables, controlling for selection bias in this model is relatively standard. We can write the second term in equation [44] as follows:

$$E(Z_{it} I(d_{it} \neq 0) \epsilon_{it}) = E(Z_{it} \Pr[d_{it} \neq 0 | Z_{it}] E[\epsilon_{it} | d_{it} \neq 0, Z_{it}]). \quad [45]$$

Thus, controlling for selection bias implies to obtain the expression of the function  $\Pr[d_{it} \neq 0 | Z_{it}] E[\epsilon_{it} | d_{it} \neq 0, Z_{it}]$ . Assumptions 1 and 2 guarantee the monotonicity of the decision rule with respect to  $\epsilon$ . Therefore, the optimal discrete choice can be represented in terms of threshold functions for  $\epsilon$ . For the case of partial irreversibility or kinked adjustment costs, individual  $i$  chooses an interior solution at period  $t$  if  $\epsilon_{it} < e^L(x_{it}; \theta)$  or if  $\epsilon_{it} > e^H(x_{it}; \theta)$ , where  $e^L(\cdot)$  and  $e^H(\cdot)$  are two threshold functions that result from inverting  $\delta^L(\cdot)$  and  $\delta^H(\cdot)$  in the optimal decision rule at equation [13]. Therefore, the conditional choice probabilities for the events “ $d_{it} < 0$ ” and “ $d_{it} > 0$ ” are:

$$P^L(x_{it}) \equiv \Pr(d_{it} < 0 | x_{it}) = \Pr(\epsilon_{it} < e^L[x_{it}; \theta] | x_{it}); \quad [46]$$

$$P^H(x_{it}) \equiv \Pr(d_{it} > 0 | x_{it}) = \Pr(\epsilon_{it} > e^H[x_{it}; \theta] | x_{it}). \quad [47]$$

Furthermore, the independence between  $\epsilon_{it}$  and  $x_{it}$  and the monotonicity of the discrete choice with respect to  $\epsilon_{it}$  imply that the conditional expectation  $E(\epsilon_{it} | d_{it} \neq 0, Z_{it})$  is a function of the choice probabilities  $P^L(x_{it})$  and  $P^H(x_{it})$  (see Stoker, 1991). Therefore, we can write:

$$\Pr[d_{it} \neq 0 | Z_{it}] E[\epsilon_{it} | d_{it} \neq 0, Z_{it}] = H(P^L[x_{it}], P^H[x_{it}]), \quad [48]$$

where the form of the function  $H(\cdot)$  depends on the probability distribution of  $\epsilon_{it}$ . For example, if  $\epsilon_{it}$  is normally distributed:

$$H(P^L[x_{it}], P^H[x_{it}]) = \sigma_\epsilon \phi(\Phi^{-1}(1 - P^H[x_{it}])) - \phi(\Phi^{-1}(P^L[x_{it}])), \quad [49]$$

where  $\phi(\cdot)$  and  $\Phi^{-1}(\cdot)$  are the pdf and the inverse-cdf of the standard normal, respectively.

Taking into account the previous results we can write the following moment conditions:

$$E(Z_{it} I(d_{it} \neq 0) [\Psi^O(X_i^{t,t+\tau_{it}}, d_{it}, d_{i,t+\tau_{it}}; \theta) + H(P_{it}^L, P_{it}^H)]) = 0. \quad [50]$$

Given these moment conditions, it is possible to estimate  $\theta$  (or the subvector of  $\theta$  that enters in  $\Psi^O[\cdot]$ ) using a two-stage procedure. In a first stage the conditional choice probabilities  $P^L[x_{it}]$  and  $P^H[x_{it}]$  are estimated, for each value  $x_{it}$  in the data. In a second stage, the estimated probabilities are solved in equation [50] and the resulting moment conditions are used to obtain GMM estimates of  $\theta$ . Notice that, since the threshold functions that appear in the discrete choice probabilities, i.e.,  $e^{\Psi}(\cdot)$  and  $e^H(\cdot)$ , are unknown (unless the dynamic programming model is solved), we should approximate these functions using a polynomial in the vector of observable state variable  $x_{it}$ .

This two-stage method can be semiparametric or parametric, depending on whether a parametric assumption about the probability distribution of  $\epsilon_{it}$  is made or not. If we assume that  $\epsilon_{it}$  is normally distributed, we estimate the probabilities  $P^L(x_{it})$  and  $P^H(x_{it})$  using two probit models (or an ordered probit model), we substitute these estimates in equation [49], and obtain GMM estimates of  $\theta$  and  $\sigma_\epsilon$  using the moment conditions in [50]. If there is not a parametric assumption about the distribution of  $\epsilon$ , the probabilities  $P^L[x_{it}]$  and  $P^H[x_{it}]$  can be estimated using a kernel method. In the second stage the function  $H(\cdot)$  can be approximated by a polynomial in the estimated discrete choice probabilities, and the parameters of this polynomial are jointly estimated with  $\theta$  using a GMM (see Stoker, 1991, and Ahn and Powell, 1993, for a description of similar semiparametric methods to control for selection bias). This two-stage semiparametric method provides root-n-consistent and asymptotically normal estimates of  $\theta$ , and its asymptotic covariance matrix can be estimated using the procedure proposed by Newey (1994).

The previous approach (either parametric or semiparametric) can be slightly modified to account for time-heteroscedasticity in  $\epsilon_{it}$ . This is an important consideration because the well-known result that when  $\epsilon$  is heteroscedastic, and it is not taken into account, the estimates of the choice probabilities will be inconsistent. In that case the conditional choice probabilities should be estimated separately for each

period  $t$ , and in the second stage the parameters associated to  $H(\cdot)$  should be different for each period  $t$ .

**EXAMPLE 3:** Consider our labor demand model where the production function is Cobb-Douglass and the specification of the unobservables is the one presented in Example 2. Under Assumption 2, and considering that the maximum value of  $\tau_{it}$  in the sample is 3, we can represent the moment conditions in [50] using the following linear (in variables) model:

$$\begin{aligned}
 \tilde{y}_{it} = & \gamma_{w0} w_{it} \\
 & + \gamma_{y1} I(\tau_{it} \geq 2) \tilde{y}_{i,t+1} + \gamma_{w1} I(\tau_{it} \geq 2) w_{i,t+1} \\
 & + \gamma_{y2} I(\tau_{it} \geq 3) \tilde{y}_{i,t+2} + \gamma_{w2} I(\tau_{it} \geq 3) w_{i,t+2} \\
 & + \gamma_{L0} I(d_{it} < 0) + \gamma_{H0} I(d_{it} > 0) \\
 & + \gamma_{L1} I(d_{i,t+1} < 0) + \gamma_{H1} I(d_{i,t+1} > 0) \\
 & + \gamma_{L2} I(d_{i,t+2} < 0) + \gamma_{H2} I(d_{i,t+2} > 0) \\
 & + \gamma_{L3} I(d_{i,t+3} < 0) + \gamma_{H3} I(d_{i,t+3} > 0) \\
 & + H(P^L[x_{it}], P^H[x_{it}]) + \zeta_{it} \tag{51}
 \end{aligned}$$

and the orthogonality conditions,  $E(Z_{it} I[d_{it} \neq 0] \zeta_{it}) = 0$ . Where  $\tilde{y}_{i,t+j} = y_{i,t+j}/l_{i,t+j}$ , and the  $\gamma$  parameters are the following functions of the structural parameters:

$$\begin{aligned}
 \gamma_{w0} = \frac{1}{\alpha} \quad ; \quad \gamma_{L0} = \frac{-c^L}{\alpha} \quad ; \quad \gamma_{H0} = \frac{c^H}{\alpha} \\
 \gamma_{y1} = -\beta \quad ; \quad \gamma_{w1} = \frac{\beta}{\alpha} \quad ; \quad \gamma_{L1} = \frac{\beta c^L}{\alpha} \quad ; \quad \gamma_{H1} = \frac{-\beta c^H}{\alpha} \\
 \gamma_{y3} = -\beta^2 \lambda \quad ; \quad \gamma_{w2} = \frac{\beta^2 \lambda}{\alpha} \quad ; \quad \gamma_{L2} = \frac{\beta^2 \lambda c^L}{\alpha} \quad ; \quad \gamma_{H2} = \frac{-\beta^2 \lambda c^H}{\alpha} \\
 \quad \quad \quad ; \quad \gamma_{L3} = \frac{-\beta^3 \lambda^2 c^L}{\alpha} \quad ; \quad \gamma_{H3} = \frac{-\beta^3 \lambda^2 c^H}{\alpha}; \tag{52}
 \end{aligned}$$

where  $\alpha$  is the parameter associated to the labor input in the Cobb-Douglas production function.

If we assume normality of  $\epsilon_{it}$  we can estimate the choice probabilities  $P^L[x_{it}]$  and  $P^H[x_{it}]$  using two probit models where the explanatory variables are the terms in a second order polynomial in  $x_{it} = (l_{it}, w_{it})$ . Alternatively, if we do not make any parametric assumption about  $\epsilon_{it}$  we can estimate the discrete choice probabilities using, for instance,

Nadaraya-Watson kernel estimators where the regressors are  $l_{it}$  and  $w_{it}$  and the dependent variables are the indicators  $I(d_{it} < 0)$  and  $I(d_{it} > 0)$ .

Given the estimates of the choice probabilities, the  $\gamma$  parameters and the parameters in the  $H(\cdot)$  function can be jointly estimated using a linear GMM, like the one in Arellano and Bond (1991). Finally, given the estimates of  $\gamma$  and the estimated covariance matrix, we can obtain estimates of the structural parameters  $(\alpha, \beta, \lambda, c^H, c^L)$  using a Minimum Distance method.

The GMM estimation of  $\gamma$  could include not only the moment conditions from the Euler equations but also moment conditions for the production function and for the transition rule of employment. In principle, this simultaneous equations GMM estimator will provide more precise estimates than a GMM estimator that only exploits moment conditions from the Euler equation.

#### *Autocorrelated unobservables*

The assumption of no autocorrelation in the unobservables is not realistic in most applications using micro datasets, where individual heterogeneity and persistent idiosyncratic shocks are usually important. In the context of a solution estimation method, dealing with autocorrelated unobservables is a particularly difficult issue that requires one to use simulation based estimation techniques (see Keane and Wolpin, 1996, for an example in a dynamic discrete choice model). Autocorrelated unobservables introduce also some additional complications in the estimation of Euler equations. In particular, if  $\epsilon$  is autocorrelated all previous values of the endogenous state variables are correlated with current and future values of  $\epsilon$ . Therefore, in the context of autocorrelated unobservables, endogenous state variables cannot be used as instrumental variables in the moment conditions of equation [44]. If there is enough variability in the observable exogenous state variables, and we can assume that these variables are orthogonal to unobservables, it would be possible to estimate the model using the moment conditions in [44], where now  $Z_{it}$  includes only functions of these exogenous state variables. However, in most micro applications, it is commonly the case that either these exogenous state variables are not available or their variability is not enough

to guarantee the identification of the parameters of interest (e.g., exogenous state variables with variability at the level of industry or region but not at the level of individuals).

However, given an assumption about the structure of the stochastic process  $\{\epsilon_{it}\}$  it is possible to obtain a transformation of the Euler equation that permits one to construct moment conditions where previous values of endogenous state variables are valid instruments. In this context, the econometric problem is to obtain the appropriate transformation of the Euler equations (e.g., time-differences) that guarantees the orthogonality between observable state variables and transformed unobservables. First, we relax Assumption 2 to allow for autocorrelation in the unobservables.

**ASSUMPTION 3:** *The stochastic structure of the unobservables is:*

$$\epsilon_{it} = \eta_i + \epsilon_{it}^*, \quad \text{where} \quad \epsilon_{it}^* = \rho \epsilon_{i,t-1}^* + a_{it}, \quad a_{it} \text{ is iid over } t, \quad [53]$$

where  $\rho \in (0, 1)$  and  $a_{it}$  is iid with zero mean and it is independently distributed of current and previous values of  $x_{it}$ .

In order to distinguish the specific problems associated to censored CMDP, we consider first the estimation of a CMDP without censoring under Assumptions 1 and 3. Notice that in the model without censoring  $\tau_{it} = 1$  with probability one, and thus the unobservable  $v_{it}(\tau_{it})$  is equal to  $\epsilon_{it} + \beta(\theta_x - 1) \epsilon_{i,t+1}$  (see equation [43]). Therefore,

$$\begin{aligned} \Delta v_{it} - \rho \Delta v_{i,t-1} &= \{\Delta \epsilon_{it} - \rho \Delta \epsilon_{i,t-1}\} + \beta(\theta_x - 1) \{\Delta \epsilon_{i,t+1} - \rho \Delta \epsilon_{it}\} \\ &= \Delta a_{it} + \beta(\theta_x - 1) \Delta a_{i,t+1}, \end{aligned} \quad [54]$$

where  $\Delta$  is the first-differences operator. Under Assumption 2,  $E(Z_{i,t-1} [\Delta a_{it} + \beta(\theta_x - 1) \Delta a_{i,t+1}]) = 0$ , and thus the following moment conditions hold:

$$E\left(Z_{i,t-1} \left[\Delta \Psi_{it}^O(\theta) - \rho \Delta \Psi_{i,t-1}^O(\theta)\right]\right) = 0, \quad [55]$$

where  $Z_{i,t-1}$  is a vector of functions of observable state variables (both endogenous and exogenous) dated at  $t - 1$  or before; and  $\Psi_{it}^O(\theta) \equiv \Psi^O(X_i^{t,t+1}, d_{it}, d_{i,t+1}; \theta)$ . These moment conditions depend only on observable variables and structural parameters, and then they can be used to estimate by GMM the subvector of  $\theta$  that enters in  $\Psi^O(\cdot)$ . Notice that equation [55] simplifies if the stochastic structure

of the unobservables is such that there is not time-invariant individual heterogeneity, or if there is not an AR(1) process. In the first case the appropriate transformation would not require to take first differences, and in the second case the term  $-\rho\Delta\Psi_{i,t-1}^O(\theta)$  would not enter in these moment conditions.

The transformation in the case of a censored CMDP is a bit more complicated. The first problem is that Euler equations do not hold at each period and, therefore, it is not always possible to obtain first-differences between Euler equations. The second problem is that the stochastic structure of the unobservable  $v_{it}(\tau_{it})$  (e.g., autocorrelation structure, expression of the time invariant individual component) depends  $\tau_{it}$ .

For the sake of simplicity assume that  $\rho = 0$ , that is,  $\epsilon_{it} = \eta_i + a_{it}$ . First, consider that we transform equation [44] taking time differences between the Euler equation at period  $t$  and the Euler equation at the next interior solution, i.e., at  $t + \tau_{it}$ . The unobservable component in the transformed equation is  $\nabla^{\tau_{it}} v_{it} \equiv v_{i,t+\tau_{it}} - v_{it}$ . This unobservable is not orthogonal to  $x_{it}$  unless  $\tau_{i,t+\tau_{it}}$  is equal to  $\tau_{it}$ . To see this consider that  $\tau_{it} = 1$  and  $\tau_{i,t+\tau_{it}} = 2$ . In that case:

$$\nabla^{\tau_{it}} v_{it} = v_{i,t+2}(2) - v_{it}(1) = \xi_{it} + \beta(1 + \beta\lambda[\theta_x - 1]) \eta_i, \quad [56]$$

where  $\xi_{it}$  represents the sum of values of  $a$  between periods  $t$  and  $t + 3$ , and it is orthogonal to  $x_{it}$ . However,  $\beta(1 + \beta\lambda[\theta_x - 1]) \eta_i$  is not orthogonal to  $x_{it}$ .

Nevertheless, there exists a transformation of the Euler equation that guarantees the orthogonality between observable state variables and transformed unobservables. Define, for each interior solution, say  $(i, t)$ , the following variable:

$$\mu_{it} = \min \{ \mu > 0 : \delta(s_{i,t+\mu}) \neq 0 \text{ and } \tau_{i,t+\mu} = \tau_{it} \}. \quad [57]$$

That is,  $\mu_{it}$  is the number of periods between period  $t$  and the next interior solution with the same duration spell  $\tau$ . Under Assumption 3 and  $\rho = 0$  the following orthogonality condition holds:

$$E_t(v_{i,t+\mu_{it}}(\tau_{it}) - v_{it}(\tau_{it})) = 0. \quad [58]$$

Therefore, it is possible to construct the following moment conditions.

$$\begin{aligned} E \left( Z_{it} I(d_{it} \neq 0) \left[ \Psi_{i,t+\mu_{it}}^O(\theta) - \Psi_{it}^O(\theta) \right. \right. \\ \left. \left. + H(\mu_{it}, \tau_{it}, P^L[x_{it}], P^H[x_{it}]) \right] \right) = 0, \end{aligned} \quad [59]$$

where

$$\begin{aligned} & H(\mu_{it}, \tau_{it}, P^L[x_{it}], P^H[x_{it}]) \\ &= \Pr[d_{it} \neq 0 | Z_{it}] E(v_{i,t+\mu_{it}}(\tau_{it}) - v_{it}(\tau_{it}) | x_{it}, d_{it} \neq 0). \end{aligned} \quad [60]$$

Notice that now the selection function  $H(\cdot)$  depends on  $\mu_{it}$  and  $\tau_{it}$  because the variance of  $\{v_{i,t+\mu_{it}}(\tau_{it}) - v_{it}(\tau_{it})\}$  and the covariance of this variable with  $\epsilon_{it}$  (i.e., the unobservable in the discrete choice) depend on  $\mu_{it}$  and  $\tau_{it}$ . The approach to control for sample selection is very similar to the one described for the case with non autocorrelated unobservables. The only difference is that in the GMM estimation of equation [59] the parameters associated to the  $H(\cdot)$  function should be different for different values of  $\mu_{it}$  and  $\tau_{it}$ .

For  $\rho \neq 0$  and no time-invariant individual heterogeneity the appropriate transformation is  $(v_{i,t+\mu_{it}}(\tau_{it}) - \rho^{\tau_{it}} v_{it}(\tau_{it}))$ .

**EXAMPLE 4:** Consider our labor demand model where now there is time-invariant individual heterogeneity in the unobservable components associated to adjustment costs:  $\epsilon_{it} = \eta_i + a_{it}$ . Assuming that the maximum value of  $\tau_{it}$  in the sample is 3, we can represent the moment conditions in equation [59] using the linear model in expression [51], where now all the variables are transformed using the operator  $\nabla^{\mu_{it}}$ , and the  $\gamma$  parameters have the same definition that in [52].

However, there is an important practical limitation in the construction of the moment conditions in [59]. This limitation is associated to the truncated distribution of  $\tau_{it}$  and  $\mu_{it}$  in the sample. We discuss this issue in the following subsection.

### 3.3. The limitations of the Euler equations approach

In the context of censored CMDP there are several potential limitations associated to the estimation of the model using only Euler

equations. These limitations are model and data specific, and, therefore, they should be evaluated at each specific application.

The first limitation is that, in some particular models, not all the parameters of interest enter in the marginal conditions of optimality, and thus not all the parameters can be identified exploiting moment conditions from the Euler equations. That is the case of lump-sum adjustment costs parameters, and other parameters of interest that affect the discrete choice interior solution/corner solution but do not enter explicitly in the marginal conditions of optimality. To estimate these parameters it will be necessary to estimate that discrete choice. Of course, this is not a limitation *per se* of the estimation of Euler equations, but of the estimation of some censored CMDP exploiting only the structure in Euler equations.

The second problem is associated to our ability to construct the moment conditions in equations [50] or [59] given a particular sample. In particular, the problem is that in some datasets there may be many interior solution observations for which the variables  $\tau_{it}$  and  $\mu_{it}$  are not observed, that is, they are truncated. For the last interior solution of each individual  $\tau_{it}$  is unobservable, i.e., it is truncated because we only know that  $\tau_{it} > T_i - t$ . Furthermore, for some interior solutions (not necessarily the last one of each individual)  $\mu_{it}$  is observable only if  $\mu_{it} \leq T_i - t$ . This introduces two practical problems. First, the sample counterparts of the moments in [50] and [59] are in fact conditional to  $\tau_{it} \leq T_i - t$  and to  $\mu_{it} \leq T_i - t$ . That is not a problem if the temporal dimension of the panel is “relatively large” because, though we eliminate one interior solution per individual, we still have almost a random sample of interior solutions for each individual. However, if the temporal dimension of the panel is “relatively small”, this selection of interior solutions will be under-representing those individuals with relatively large adjustment costs (i.e., large  $\eta_i$ ) or with relatively large initial conditions (i.e., large  $\epsilon_{i0}$ ). In that case there will be a sample selection bias in our estimates. Notice that the concepts “large” and “small” temporal dimension of the panel depend on the distribution of  $\tau_{it}$ . If the duration between two adjustments tends to be low (e.g.,  $\tau_{it}$  between 1 and 3) values of  $T_i$  like 8 could be enough to avoid the selection bias. But larger average durations

between two adjustments will require a larger temporal dimension of the panel to avoid this selection bias.

Finally, in the context of some censored CMDP, an estimator based only on moment conditions from Euler equations can be very imprecise. The reason is that, if there is “lumpiness” in the interior solutions, the variability of the decision variable in the subsample of interior solutions can be very small. Lumpiness of the decision variable tends to be a common characteristic in many datasets with censored decision variables. For instance, a firm’s orders to manufacturers, the nominal price change of a retailer selling price, the number of workers that a firm hires, or the amount of a durable good that a consumer purchases, when they are positive, tend to have very small variability over time. In that case, the discrete choice *corner solution versus interior solution* represents most of the within-individuals variability of the decision variable and, therefore, it contains most of the sample information that may allow us to identify the parameters of interest (see Aguirregabiria, 1997, for an example of this issue). In those cases, the estimation of the optimal discrete choice will be crucial to identify the structural parameters.

#### 4. Estimation of the optimal discrete choice

We have seen at the end of previous section that in some empirical applications the estimation of a censored CMDP based only on Euler equations can provide very imprecise estimates of the parameters of interest. In this section we present an econometric approach to estimate the optimal discrete choice. This method maintains the main advantage of the Euler equations approach: lower computational cost than solution estimation methods. The estimation of the optimal discrete choice can be combined with the estimation of the Euler equation to obtain more efficient estimates of the structural parameters. We begin presenting the basic idea behind the estimation of the optimal discrete choice without solving explicitly the dynamic decision model.

The optimal discrete choice can be represented in terms of inequalities between *conditional choice value functions*. Before solving explicitly the model these value functions are unknown functions of state vari-

ables and structural parameters. However, each conditional value function can be represented as the expected value of a *known* function of future paths of state and decision variables and structural parameters, where the expectation is conditional to current state variables and to current hypothetical decisions. Under some assumptions about the joint distribution of observables and unobservables, this alternative representation of the conditional value functions can be exploited to estimate the optimal discrete choice without solving explicitly the model.

To be more specific, assume that we can represent the conditional value functions as known functions of the structural parameters and conditional expectations that incorporate only observable variables. In such a case we can estimate the optimal discrete choice using a two-stage method. In a first stage the conditional expectations entering in the value functions are estimated (nonparametrically). In a second stage these estimates are solved in the known expressions for the conditional value functions and the discrete choice is estimated. In the context of dynamic discrete choice models different versions of this approach have been proposed and implemented by Manski (1991 and 1993), Hotz and Miller (1993), Hotz et al. (1994), and Ahn (1995). For censored CMDP, a procedure in the spirit of Hotz-Miller method has been implemented by Aguirregabiria (1996) in a model of firms' price and inventory decisions. Park (1996) has used a similar method to estimate a dynamic game.

A key aspect of this approach is to know under what assumptions about the joint distribution of observables and unobservables we can express the unknown value functions as known functions of structural parameters and conditional expectations that incorporate only observables. There are several possible sufficient conditions. We consider the assumption that has been most commonly used in the literature, the so called *Conditional Independence Assumption* (Rust, 1987). Here we will follow Aguirregabiria (1997) to obtain the optimal discrete choice of the labor demand model in Subsection 2.3.

Consider the characterization of the optimal decision rule of this model in equation [27]. We can distinguish three regimes, hiring, firing, and no (gross) change in employment. Let  $j \in \{L, H, 0\}$  be

the index for this discrete choice. We can represent the optimal decision rule in equation [27] using the following compact expression:

$$\delta(s_{it}; \theta) = I(j^*[s_{it}; \theta] = L) \delta^L(s_{it}; \theta) + I(j^*[s_{it}; \theta] = H) \delta^H(s_{it}; \theta), \quad [61]$$

where  $j^*(s_{it}; \theta)$  is the optimal discrete decision. Let  $V^j(s_{it}; \theta)$  be the value function conditional to the hypothetical choice of discrete alternative  $j$ . By definition,  $V(s_{it}; \theta) = \max_{\{j\}} \{V^j(s_{it}; \theta)\}$ , and therefore:

$$V^j(s_{it}; \theta) = u(s_{it}, \delta^j[s_{it}; \theta]; \theta) + \beta E \left( \max_{h \in \{L, H, 0\}} \{V^h(s_{i, t+1}; \theta)\} \mid s_{it}, j; \theta \right), \quad [62]$$

where  $\delta^0(s_{it}; \theta) = 0$ . Using the previous definitions, the optimal discrete choice can be represented using the following expression:

$$j^*(s_{it}; \theta) = j \quad \Leftrightarrow \quad j = \arg \max_{h \in \{L, H, 0\}} \{V^h(s_{it}; \theta)\}. \quad [63]$$

Let  $Ew^j(x_{it}; \theta)$  be the expected value of the one-period profit conditional to  $x_{it}$  and to the hypothetical choice of discrete alternative  $j$ .

$$Ew^j(x_{it}; \theta) \equiv E \left( u(s_{it}, \delta^j[s_{it}; \theta]; \theta) \mid x_{it}, j; \theta \right). \quad [64]$$

Using this definition we can write:

$$u(s_{it}, \delta^j[s_{it}; \theta]; \theta) = Ew^j(x_{it}; \theta) + \mu_{it}^j, \quad [65]$$

where, by construction, the unobservable  $\mu_{it}^j$  is orthogonal to  $x_{it}$ . Given our specification of the one-period profit function in Subsection 2.3, we have that:

$$Ew^j(x_{it}; \theta) = \Pi^j(x_{it})' \gamma(\theta), \quad [66]$$

where

$$\begin{aligned} \Pi^L(x_{it}) &= \{ E[y_{it} - w_{it}(l_{it} + d_{it}) \mid x_{it}, j = L], \\ &\quad -E[d_{it} \mid x_{it}, j = L], 0, -1, 0 \}' \\ \Pi^H(x_{it}) &= \{ E[y_{it} - w_{it}(l_{it} + d_{it}) \mid x_{it}, j = H], \\ &\quad 0, -E[d_{it} \mid x_{it}, j = H], 0, -1 \}' \\ \Pi^0(x) &= \{ E[y_{it} - w_{it}(l_{it} + d_{it}) \mid x_{it}, j = 0], \\ &\quad 0, 0, 0, 0 \}' \end{aligned} \quad [67]$$

where  $y_{it}$  is real output. And

$$\gamma(\theta) = \left(1, c^L, c^H, F, H\right)'. \quad [68]$$

The unobservables  $\{\mu_{it}^j\}$  represent the uncertainty of the researcher about the actual expected profit that is observable to the firm. By construction they are mean independent of  $x$ . We consider the following assumption about the joint distribution of  $x_{it}$  and  $\mu_{it} = (\mu_{it}^L, \mu_{it}^H, \mu_{it}^0)'$ .

**ASSUMPTION 4: Conditional Independence Assumption (Rust [1987, 1994]):** The transition probability of the state variables factors as:

$$p(x_{i,t+1}, \mu_{i,t+1} | x_{it}, \mu_{it}, j) = p_\mu(\mu_{i,t+1} | x_{i,t+1}) p_x(x_{i,t+1} | x_{it}, j). \quad [69]$$

The conditional independence assumption is in fact the combination of two different assumptions. The first one says that, conditional on contemporaneous observable state variables, the unobservables do not depend on any previous decision or state variable. This condition is weaker than assuming that  $\mu_{it}$  is *iid* over time, but it is stronger than Assumption 3 because all the autocorrelation in the unobservables should be captured by the autocorrelation in the observables. The second assumption imbedded in Assumption 4 is that, conditional on observable decision and state variables at period  $t$ , the observable state variables at  $t + 1$  do not depend on  $\mu_{it}$ .

Assumption 4 has some useful implications on the form of the discrete choice model in equations [62] and [63]. Let  $EV^j(s_{it}; \theta)$  be the second component of  $V^j(s_{it}; \theta)$  in expression [62], i.e., the conditional choice expectation of next period value function. Under Assumption 4,  $EV^j(s_{it}; \theta)$  does not depend on the vector of unobservables  $\mu_{it}$ :

$$EV^j(s_{it}; \theta) \equiv E \left( \max_{j \in \{L, H, 0\}} \{V^j(s_{i,t+1}; \theta)\} \mid s_{it}, j; \theta \right) = EV^j(x_{it}; \theta). \quad [70]$$

Therefore, we can write the optimal discrete choice as:

$$j^*(s_{it}; \theta) = j \quad \Leftrightarrow \\ j = \arg \max_{h \in \{L, H, 0\}} \left\{ \Pi^h(x_{it})' \gamma(\theta) + \mu_{it}^h + \beta EV^h(x_{it}; \theta) \right\}. \quad [71]$$

Another implication of the conditional independence assumption is that, for a certain firm  $i$ , the probability of the history of discrete choices conditional to observable state variables is equal to the product of the conditional choice probabilities at each period.

Lemma 5 presents the main result that will be used to define an estimator of  $\theta$  that exploits the structure of the optimal discrete choice.

LEMMA 5: *Under Assumption 4 and the multiplicative separability between  $\theta$  and  $x$  in  $Eu^d(x; \theta)$ , the optimal discrete choice can be represented using the following expression:*

$$j^*(s_{it}; \theta) = j \Leftrightarrow j = \arg \max_{h \in \{L, H, 0\}} \left\{ \Pi^h(x_{it})' \gamma(\theta) + W_{it}^{h'} \varphi(\theta) + \mu_{it}^h \right\}, \quad [72]$$

where  $\varphi(\theta) = (\gamma(\theta)', 1)'$ . And, given a discretization of the space of observable state variables, the  $W_{it}^j$  vectors have the following expression:

$$W_{it}^j = F^j(x_{it}) (I_M - \beta F)^{-1} \left( \sum_{h \in \{L, H, 0\}} P^h * \Pi^h ; \sum_{h \in \{L, H, 0\}} P^h * e^h(P) \right), \quad [73]$$

where  $*$  means element-by-element product;  $M$  is the number of cells in the discretized space;  $F^j(x_{it})$  is the  $M \times 1$  vector of transition probabilities of  $x$  conditional to  $j$  and to  $x_{it}$ ;  $F$  is the  $M \times M$  matrix of unconditional transition probabilities of  $x$ ;  $P^j$  is the  $M \times 1$  vector of conditional choice probabilities for alternative  $j$ ;  $\Pi^j$  is the matrix with the  $M$  row vectors  $\Pi^j(x)'$ ; and  $e^j(P)$  is a known function of the choice probabilities. The form of  $e^j(P)$  depends on the distribution of  $\mu_{it}$  (e.g., if  $\mu_{it}$  is extreme value type 1,  $e^j(P[x]) = const - \ln P^j(x)$ , where  $const$  is the Euler's constant).

The proof of this Lemma is in Aguirregabiria (1997).

Based on Lemma 5 we can obtain a root- $n$ -consistent estimator of  $\theta$  using a sequential procedure. In the first stage we obtain nonparametric kernel estimates of  $P^j(x)$  for each discrete alternative and each value of  $x$  in the sample. We also estimate the transition probabilities of the observable state variables. In the second stage we use the

previous estimates to construct the values  $\hat{W}_{it}^j$ , and estimate  $\theta$  using a GMM that exploits the following moment conditions:

$$E\left(Z_{it}\left[I(j_{it} = j) - p^j(\Pi_{it}, \hat{W}_{it}; \theta)\right]\right) = 0, \quad \text{for } j = L, H, \quad [74]$$

where  $Z_{it}$  is a vector of instrumental variables (i.e., current and previous values of  $x_{it}$ ); and, given an extreme value distribution for  $\mu_{it}$ :

$$p^j(\Pi_{it}, \hat{W}_{it}; \theta) = \frac{\exp\{\Pi_{it}^j \gamma(\theta) + \hat{W}_{it}^j \varphi(\theta)\}}{\sum_{h \in \{L, H, 0\}} \exp\{\Pi_{it}^h \gamma(\theta) + \hat{W}_{it}^h \varphi(\theta)\}}. \quad [75]$$

Hotz and Miller (1993) prove the consistency and asymptotic normality of a general class of estimators that includes this one. They also obtain the expression of the asymptotic covariance matrix of this estimator, that accounts for its sequential nature (Hotz and Miller, 1993, Equation 5.11).

## 5. Conclusion

In this paper we have analyzed the estimation of dynamic structural models with censored decision variables. We have discussed the advantages and limitations of several estimation methods under different specifications of the structure of the unobservables and under different characteristics of the working sample. Given our current state of knowledge the results are limited. In particular, the incorporation of autocorrelated unobservables in discrete choice dynamic structural models is a problem that has not been completely solved. To a certain extent, that is also the case for the estimation of Euler equations with autocorrelated unobservables.

However, the methods that we have described above can be currently applied to many problems for which there have been a relative lack of structural applications, i.e., machine replacement and irreversible investment, labor demand with non convex adjustment cost, consumption of durable goods, models of price competition with menu costs, or firms' entry and exit in a market, among others.

## References

- Abel, A. and J. Eberly (1996): "Optimal investment with costly reversibility", *Review of Economic Studies* 63, pp. 581-594.
- Aguirregabiria, V. (1996): "Estimation of dynamic decision models with corner solutions: A model of price and inventory decisions", Working Paper 9602, The University of Western Ontario.
- Aguirregabiria, V. (1997): "The dynamics of markups and inventories in retailing firms", manuscript, The University of Western Ontario.
- Ahn, H. (1995): "Nonparametric two stage estimation of conditional choice probabilities in a binary choice model under uncertainty", *Journal of Econometrics* 67, pp. 337-378.
- Ahn, H. and J. L. Powell (1993): "Semiparametric estimation of censored selection models with a nonparametric selection mechanism", *Journal of Econometrics* 58, pp. 3-30.
- Alonso-Borrego, C. (1994): "Demand for labour inputs with asymmetric adjustment costs", CEMFI Working Paper 9417.
- Arellano, M. and S. Bond (1993): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies* 58, pp. 277-297.
- Bar-Ilan, A. and A. S. Blinder (1988): "Consumer durables and the optimality of usually doing nothing", NBER Working Paper 2488.
- Bentolila, S. and G. Bertola (1990): "Titulo Firing costs and labor demand: How bad is Euroesclerosis", *Review of Economic Studies* 57, pp. 381-402.
- Bertola, G. and R. Caballero (1990): "Kinked adjustment costs and aggregate dynamics", *NBER Macroeconomics Annual 1990* MIT Press.
- Bertsekas, D. (1976): "Dynamic programming and stochastic control", New York, Academic Press.
- Blinder, A. (1980): "Retail inventory investment and business fluctuations", *Brookings Papers of Economic Activity* 2, pp. 443-505.
- Bond, S. and C. Meghir (1994): "Dynamic investment models and the firm's financial policy", 61, pp. 197-222.
- Caballero, R. and E. Engel (1991): "Dynamic (S,s) economies", *Econometrica* 59, pp. 1659-1686.
- Cooper, R. J. Haltiwanger, and L. Power (1995): "Machine replacement and the business cycle: Lumps and bumps", NBER Working Paper 5260.
- Deaton, A. and G. Laroque (1996): "Competitive storage and commodity price dynamics", *Journal of Political Economy* 104, pp. 896-923.

- Eberly, J. (1994): "Adjustment of consumers' durables stocks: Evidence from automobile purchases", 102, pp. 403-436.
- Eckstein, Z. and K. Wolpin (1989): "The specification and estimation of dynamic stochastic discrete choice models", *Journal of Human Resources* 24, pp. 562-598.
- Hamermesh, D. (1989): "Labor demand and the structure of adjustment costs", 79, pp. 674-689.
- Hamermesh, D. (1993): "Labor demand and the source of adjustment costs", NBER Working Paper 4394.
- Hansen, L.P. and K. Singleton (1982): "Generalized instrumental variables estimation of nonlinear rational expectations models", *Econometrica* 50, pp. 1269-1281.
- Hopenhayn, H. and R. Rogerson (1993): "Job turnover and policy evaluation: A general equilibrium analysis", *Journal of Political Economy* 101, pp. 915-938.
- Hotz, J. and R.A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models", *Review of Economic Studies* 60, pp. 497-529.
- Hotz, J., R.A. Miller, S. Sanders and J. Smith (1994): "A simulation estimator for dynamic models of discrete choice", *Review of Economic Studies* 61, pp. 265-289.
- Keane, M. and K. I. Wolpin (1994): "The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence", *The Review of Economics and Statistics* 76, pp. 648-672.
- Manski, C. (1991): "Nonparametric estimation of expectations in the analysis of discrete choice under uncertainty", in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, W. Barnett, J. Powell, and G. Tauchen (eds.). Cambridge University Press.
- Manski, C. (1993): "Dynamic choice in social settings: Learning from the experience of others", *Journal of Econometrics* 58, pp. 121-136.
- McFadden, D. (1973): "Conditional logit analysis of qualitative choice behavior", in P. Zarembka (ed.) *Frontiers of Econometrics*, New York, Academic Press, pp. 105-142.
- Miller, R. A. (1984): "Job matching and occupational choice", *Journal of Political Economy* 92, pp. 1086-1120.
- Newey, W. (1995): "The asymptotic variance of semiparametric estimators", *Econometrica* 62, pp. 1349-1382.

- Pakes, A (1986): "Patents as options: Some estimates of the value of holding European patent stocks", *Econometrica* 54, pp. 755-784.
- Pakes, A. (1994): "Dynamic structural models: Problems and prospects. Mixed continuous discrete controls and market interactions", in C. Sims (ed.) *Advances in Econometrics. Sixth World Congress*, Volume 2, Cambridge University Press.
- Pakes, A. and S. Olley (1995): "A limit theorem for a smooth class of semiparametric estimators", *Journal of Econometrics* 65, pp. 295-332.
- Park, S. (1996): "Semiparametric estimation of some dynamic structural models with strategic interdependence", manuscript, SUNY at Stony-Brook.
- Pfann, G. and F. Palm (1993): "Asymmetric adjustment costs in non-linear labour demand models for the Netherlands and the UK manufacturing sectors", *Review of Economic Studies* 60, pp. 397-412.
- Pindyck, R.S. (1988): "Irreversible investment, capacity choice, and the value of the firm", *American Economic Review* 78, pp. 969-985.
- Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher", *Econometrica* 55, pp. 999-1033.
- Rust, J. (1992): "Do People Behave According to Bellman's Principle of Optimality?", manuscript, Hoover Institution, Stanford University.
- Rust, J. (1994): "Structural estimation of Markov decision processes", in R. E. Engle and McFadden (eds.) *Handbook of Econometrics*, Volume 4, North-Holland.
- Scarf, H. (1959): "The optimality of (S,s) policies in the dynamic inventory problem", in K. J. Arrow, S. Karlin, and P. Suppes (eds.) *Mathematical Methods in the Social Sciences*. Stanford University Press.
- Sheshinski, E., and Y. Weiss (1983): "Optimum pricing policy under stochastic inflation", *Review of Economic Studies* 50, pp. 513-529.
- Stoker, T. (1991): *Lecture on semiparametric econometrics*, CORE Lecture Series, Universite Catholique de Louvain, Louvain-La-Neuve.
- Stokey, N. and R. Lucas (1989): *Recursive methods in economic dynamics*, Harvard University Press.

**Resumen**

*En este artículo se analiza la estimación de modelos estructurales dinámicos en los que las variables de decisión están censuradas. El artículo discute los problemas que aparecen en la estimación de estos modelos bajo distintos supuestos sobre la distribución de las variables no observables, diferentes fuentes de censura, así como distintas características de los datos utilizados (su dimensión temporal, la frecuencia empírica de las soluciones esquina, o la distribución de las duraciones entre dos soluciones interiores). Los métodos y problemas econométricos son ilustrados utilizando un modelo de demanda de trabajo con costes de ajuste lineales y de suma fija.*