IDEI Working Paper Series n° 624

April, 2010

"Endogeneity and Instrumental Variables in Dynamic Models"

Jean-Pierre FLORENS and Guillaume SIMON

Institut d'Economie Industrielle (IDEI) – Manufacture des Tabacs
Aile Jean-Jacques Laffont – 21, allée de Brienne – 31000 TOULOUSE – FRANCE
Tél. + 33(0)5 61 12 85 89 – Fax + 33(0)5 61 12 86 37 – **www.idei.fr** – contact@cict.fr

INSTITUT D'ECONOMIE INDUSTRIELLE

Endogeneity And Instrumental Variables In Dynamic Models*

Jean-Pierre Florens[†], Guillaume Simon[‡]

23rd April 2010

Abstract

The objective of the paper is to draw the theory of endogeneity in dynamic models in discrete and continuous time, in particular for diffusions and counting processes. We first provide an extension of the separable set-up to a separable dynamic framework given in term of semi-martingale decomposition. Then we define our function of interest as a stopping time for an additional noise process, whose role is played by a Brownian motion for diffusions, and a Poisson process for counting processes.

JEL Codes: C14; C32; C51.

Keywords: Endogeneity; Instrumental Variables; Dynamic Models; Duration Models.

*We are grateful to Marc Hoffmann, Idris Kharroubi, Nour Meddahi, Mathieu Rosenbaum, Stéphane Villeneuve, and the participants and organizers of the conferences "First French Econometrics Conference" (Toulouse, 2009) and "Stats in the Chateau" (Paris, 2009). The usual disclaimer nonetheless applies, all errors remain ours.

[†]Toulouse School of Economics (GREMAQ, University of Toulouse Capitole). Email: `florens@cict.fr`.

[‡]Toulouse School of Economics (GREMAQ, University of Toulouse Capitole).. Email: `guillaume.simon@ensae.org`.

1 INTRODUCTION

1.1 Motivations

An econometric model has often the form of a relation where a random element Y depends on a set of random elements Z and a random noise U. If Z is exogenous (see for precise definition of this concept [Engle et al., 1983] or [Florens and Mouchart, 1982]) some independence or non correlation property is assumed between the Z and the U in order to characterize uniquely the relation. For example, if the relation has the form $Y = \phi(Z) + U$ the condition $\mathbb{E}[U|Z] = 0$ characterizes ϕ as the conditional expectation and if $Y = \phi(Z, U)$ with ϕ monotonous in U, U uniform, the condition that Z and U are independent characterizes ϕ as the conditional quantile function. This exogeneity condition is usually not satisfied (as for instance in market models, treatment effect models, selection models...) and the relation should be characterized by other assumptions.

The instrumental variables approach replaces the independence between Z and U by an independence condition between U and another set of variables W called the instruments. For example, in the separable case $Y = \phi(Z) + U$ the assumption becomes $\mathbb{E}[U|W] = 0$ (see for a recent literature [Florens, 2003], [Newey and Powell, 2003], [Hall and Horowitz, 2005]). In the nonseparable model, it is assumed that $U \perp\!\!\!\perp W$ (see contributions of [Horowitz and Lee, 2007], [Chernozhukov et al., 2007a], or [Chernozhukov et al., 2007b]). In these cases the characterization of the relation is not fully determined by the independence condition but also by a dependence condition between the Z and the W. This dependence determines the identifiability of the relation: in a nonparametric framework, this impacts the speed of convergence of the estimators.

The objective of this paper is to analyze dynamic models with endogenous elements. The goal is concentrated on the specification of the models in such a way that the functional parameter of interest appears as the solution of a functional equation (essentially linear or nonlinear integral equation). Using this equation, identification or local identification condition may be discussed. This paper is not concerned by statistical inference but shows how the functional parameter may be derived from objects which may be estimable using data. The theory of nonparametric estimation in these cases belongs to the theory of ill-posed inverse problems (see [Darolles et al., 2010], [Carrasco et al., 2003], [Carrasco, 2008]) and will be treated in specific cases in other papers.

We address the question of endogeneity in dynamic models in two ways. First we consider a separable case which extends the usual model $Y = \phi(Z) + U$ with $\mathbb{E}[U|W]$. However, this case is not sufficient to cover the endogeneity question in models where the structure of the process generating Y is given (counting processes or diffusions for instance). In this case, we analyze the impact of endogenous variables through a change of time depending on the endogenous variables. This approach covers the example of the duration models, the counting processes, the diffusion with a volatility depending on the endogenous for example. It will be shown that those change of time models give an interesting extension of non-separable models in the dynamic case. These two approaches will be treated in Section 2 and 3 of the paper and will be illustrated by examples. We first recall in the next paragraph the main mathematical tools we will use.

1.2 Mathematical framework

In this paper we essentially analyze a large class of stochastic processes verifying a decomposition property. Let $(X_t)_{t\geq 0}$ (t may be discrete or continuous) and $\mathcal{F}_t$ a filtration of σ-fields such that X_t is càdlàg (its trajectories are right-continuous and have a left-limit) and that $(\mathcal{F}_t)_t$ is right-continuous (that is to say that $\bigcap_{s>t}\mathcal{F}_s = \mathcal{F}_t$). In the usual terminology of the general theory of stochastic processes we will say that $\mathcal{F}_t$ satisfies the "*conditions habituelles*".

A process X_t is a special semi-martingale w.r.t. $(\mathcal{F}_t)_t$ if there exists two processes H_t and M_t such that:

$$X_t = X_0 + H_t + M_t; \tag{1}$$

- M_t is an $\mathcal{F}_t$-martingale;
- H_t is $\mathcal{F}_t$-predictable.

A more general definition only assumes that M_t is a local martingale but for sake of simplicity only the martingale case is treated in this paper. We also simplify the expressions by always assuming $X_0 = 0$. Extension to local martingales and to cases where $X_0 \neq 0$ requires more technicalities (in particular in Section 3). Let us note that the decomposition (1) is a.s. unique. These concepts are fundamental in the theory of stochastic processes (see in particular [Dellacherie and Meyer, 1971] - Vol II - Chap VII).

We may easily illustrate this definition in the case of discrete time models. In that case we have: $M_0 = 0$, $M_t = M_{t-1} + (X_t - \mathbb{E}[X_t|\mathcal{F}_{t-1}])$ and $H_t = H_{t-1} + [\mathbb{E}[X_t|\mathcal{F}_{t-1}] - X_{t-1}]$ (see [Protter, 2003] - Chap III). Equivalently $\Delta X_t = X_t - X_{t-1}$ may also be written:

$$\Delta X_t = X_t - X_{t-1} = (\mathbb{E}[X_t|\mathcal{F}_{t-1}] - X_{t-1}) + (X_t - \mathbb{E}[X_t|\mathcal{F}_{t-1}]).$$

In case of continuous time processes, we also restrict our study to cases where H_t is differentiable and we have the expression:

$$dX_t = h_t dt + dM_t$$

where $H_t = \int_0^t h_s ds$. Some particular cases will be analyzed in details. The first one is the single duration model with endogenous cofactors possibly time-dependent. More generally, we analyze counting processes and an example of Markovian transition model is also discussed. Finally, we also applied our approach to diffusion models.

2 THE ADDITIVELY SEPARABLE CASE : THE INSTRUMENTAL VARIABLES DECOMPOSITION OF SEMI-MARTINGALES

2.1 The framework

Let us consider a multivariate stochastic process $X_t = (Y_t, Z_t, W_t)$ (with $Y_t \in \mathbb{R}$, $Z_t \in \mathbb{R}^p$, $W_t \in \mathbb{R}^q$) and $\mathcal{X}_t$ the filtration generated by X_t i.e. $\mathcal{X}_t$ is the σ-field generated by $((Y_s, Z_s, W_s)_{s\leq t})$. We consider different subfiltrations of $\mathcal{X}_t$:

1. $\mathcal{Y}_t$, $\mathcal{Z}_t$, $\mathcal{W}_t$ are the filtrations generated by each subprocess;
2. we call the *endogenous filtration* the filtration generated by $\mathcal{Y}_t$ and $\mathcal{Z}_t$, and the *instrumental filtration* the filtration $\mathcal{Y}_t \vee \mathcal{W}_t$ generated by $\mathcal{Y}_t$ and $\mathcal{W}_t$.

We first extend the usual decomposition of semi-martingales in the following way.:

Definition 2.1. The process Y_t has a Doob-Meyer Instrumental Variable (DMIV) decomposition if:

$$Y_t = \Lambda_t + U_t$$

where:

1. **A1 -** Λ_t is $\mathcal{Y}_t \vee \mathcal{Z}_t$ predictable ;
2. **A2 -** $\mathbb{E}[U_t - U_s | \mathcal{Y}_t \vee \mathcal{W}_t] = 0$ for $0 \leq s < t$.

□

Equivalently we may say that Y_t is an IV semi-martingale w.r.t. $(\mathcal{Y}_t \vee \mathcal{Z}_t)_t$ and $(\mathcal{Y}_t \vee \mathcal{W}_t)_t$. First we can note that if $\mathcal{W}_t = \mathcal{Z}_t$ this definition reduces to the usual Doob-Meyer decomposition. If the filtration $(\mathcal{Y}_t \vee \mathcal{Z}_t)_t$ is included into $(\mathcal{Y}_t \vee \mathcal{W}_t)_t$ the problem becomes a problem of enlargement of filtrations and preservation of the martingale property. This question is central in the theory of non-causality treated e.g. by [Florens and Fougère, 1996].

We consider then the more general case where $(\mathcal{Y}_t \vee \mathcal{Z}_t)_t$ and $(\mathcal{Y}_t \vee \mathcal{W}_t)_t$ have no inclusion relation. Moreover, the two filtrations do not need to be generated by processes and $(\mathcal{Y}_t \vee \mathcal{Z}_t)_t$ and $(\mathcal{Y}_t \vee \mathcal{W}_t)_t$, and may be replaced by more general filtrations $\mathcal{F}_t$ and $\mathcal{G}_t$ under the condition that Y_t has to be adapted to each of them. Assumption **A1** means that the predictable process "only depends" on the past of Y_t and on the past of Z_t. Assumption **A2** is the independence condition between the "noise" U_t and the instruments W_t. Equality in $A2$ is a mean independence only (like in the static separable model $Y = \phi(Z) + U$) and looks like a martingale property. It's not strictly speaking a martingale property because U_t is not assumed to be adapted to $(\mathcal{Y}_t \vee \mathcal{W}_t)_t$. The usual decomposition when $(\mathcal{Y}_t \vee \mathcal{Z}_t)_t = (\mathcal{Y}_t \vee \mathcal{W}_t)_t$ is unique a.s. but in the general case, it should be noted that this unicity result is not true: this will be precisely the object of the identification condition analyzed below.

2.2 Identification

Let us first consider the characterization of the decomposition in term of conditional expectation.

Theorem 2.1. Let us assume that Y_t is a special semi-martingale w.r.t. $\mathcal{Y}_t \vee \mathcal{W}_t$ and that :

$$dY_t = h_t dt + dM_t$$

where $H_t = \int_0^t h_s ds$ is $\mathcal{Y}_t \vee \mathcal{W}_t$-predictable and M_t is a $\mathcal{Y}_t \vee \mathcal{W}_t$-martingale.

If the following family of integral equations:

$$h_t = \mathbb{E}[\lambda_t | \mathcal{Y}_t \vee \mathcal{W}_t] \quad t \geq 0 \tag{2}$$

with λ_t $\mathcal{Y}_t \vee \mathcal{Z}_t$-measurable and integrable

has a sequence of solutions λ_t, then Y_t is an IV semi-martingale and $\Lambda_t = \int_0^t \lambda_s ds$.

□

Roughly speaking, Equation (2) means that we have to solve:

$$\begin{aligned} h_t dt &= \mathbb{E}[dX_t|(Y_s, W_s)_{0\le s\le t}] \\ &= \left[\int \lambda_t((Y_s, Z_s)_{0\le s\le t}) f((Z_s)_{0\le s\le t}|(Y_s, W_s)_{0\le s\le t}) d(Z_s)_{0\le s\le t}\right] dt. \end{aligned}$$

This expression is not mathematically rigorous because the arguments of the functions are infinite dimensional but it shows how our definition extends the static separable case.

A DMIV decomposition exists if and only if h_t belongs to the range of the "instrumental" conditional expectation operator. If we restrict our attention to square integrable variables, this operator is defined on $L^2(\mathcal{Y}_t \vee \mathcal{W}_t)$. Note that the conditional expectation operator is compact under minor regularity conditions. Its range is then a strict subspace of $L^2(\mathcal{Y}_t \vee \mathcal{W}_t)$ and the existence assumption is an over-identification condition on the model. The main question concerns the unicity of the solution, which is equivalently the identifiability problem. Given the distribution of the process X_t, the function h_t, and the conditional expectation operator $\mathbb{E}[...|\mathcal{Y}_t \vee \mathcal{W}_t]$ defined on $L^2(\mathcal{Y}_t \vee \mathcal{Z}_t)$ are identifiable. The DMIV decomposition is then unique (or equivalently Λ_t is identifiable) if and only if the conditional expectation operator is one-to-one. The following concept extends the full known case of static models.

Definition 2.2. The filtration $(\mathcal{Y}_t \vee \mathcal{Z}_t)_t$ is strongly identified by the filtration $(\mathcal{Y}_t \vee \mathcal{W}_t)_t$ (or Z_t is strongly identified by $\mathcal{W}_t$ given $\mathcal{Y}_t$) if and only if for $t \geq 0$:

$$\forall \psi \in L^2(\mathcal{Y}_t \vee \mathcal{Z}_t), \mathbb{E}[\psi|\mathcal{Y}_t \vee \mathcal{W}_t] = 0 \Rightarrow \psi = 0 \quad a.s.$$

□

Corollary 2.1. The DMIV is unique is $(\mathcal{Y}_t \vee \mathcal{Z}_t)$ strongly identified by $(\mathcal{Y}_t \vee \mathcal{W}_t)$.

□

For a good treatment of conditional strong identification and its relation with the completeness concept in statistics, see [Florens et al., 1990] - Chap 5. Then if $\mathcal{Z}_t$ is strongly identified by $\mathcal{W}_t$ given $\mathcal{Y}_t$, the conditional expectation operator is one-to-one and Λ_t is identified. Several papers give more primary conditions which link this property to the conditional expectation operator (see a recent contribution of [d'Haultfoeuille, 2008]). We want to illustrate this concept in two examples : discrete-time models and diffusions.

2.3 Examples

2.3.1 Example 1 : discrete time model

Suppose that we have a discrete time model such as:

$$y_t = \lambda(\xi_t) + \epsilon_t$$

with $\mathbb{E}[\epsilon_t|y_{t-1}, \ldots, \xi_{t-1}, \ldots] = 0$. In our framework, we have then:

$$Y_t = y_0 + ... + y_t \qquad \Lambda_t = \lambda(\xi_0) + ... + \lambda(\xi_t) \qquad U_t = \epsilon_0 + ...\epsilon_t.$$

Moreover if we define:

$$\mathcal{Z}_t = \sigma\{\xi_{t+1}, \xi_t, \ldots\} \quad \mathcal{Y}_t = \sigma\{y_t, y_{t-1}, \ldots\} \quad \mathcal{W}_t = \sigma\{\xi_t, \xi_{t-1}, \ldots\}$$

then we have the following properties:

- Y_t is $\mathcal{Y}_t \vee \mathcal{Z}_t$-adapted and $\mathcal{Y}_t \vee \mathcal{W}_t$-adapted;
- Λ_t is $\mathcal{Y}_t \vee \mathcal{Z}_{t-1}$-measurable and $\mathcal{Y}_t \vee \mathcal{Z}_t$-predictable;
- $\mathbb{E}[U_t|W_{t-1}] = U_{t-1}$ as $\mathbb{E}[\epsilon_t|\mathcal{Y}_t \vee W_{t-1}] = 0$ and $\mathbb{E}[\epsilon_s|\mathcal{Y}_t \vee W_{t-1}] = \epsilon_s$ if $s \leq t-1$.

In that case, $\mathcal{Y}_t \vee \mathcal{W}_t \subset \mathcal{Y}_t \vee \mathcal{Z}_t$, $\lambda_t = \lambda(\xi_t)$, $h_t = \mathbb{E}[\lambda_t|\xi_{t-1}, ..., y_{t-1}, \ldots]$, i.e. :

$$h_t = \mathbb{E}[\lambda(\xi_t)|\xi_{t-1}, \ldots, y_{t-1}, \ldots, ...].$$

If (y_t, ξ_t) is Markovian, then we have moreover:

$$\mathbb{E}[y_t|\xi_{t-1}, y_{t-1}] = \mathbb{E}[\lambda(\xi_t)|\xi_{t-1}, y_{t-1}].$$

One can then proceed to nonparametric estimation:

- for weakly dependent stationary processes, we face inverse problems as in the usual i.i.d. case;
- when studying unit root processes, we can use ordinary kernel estimation but there is a second order bias. [Wang and Phillips, 2009] treats it with a control function, but this does not address the case of the second order bias of instrumental variables. This is therefore an argument for pure IV in non-stationary models.

More generally, we could consider:

$$y_t = \lambda_t(z_t, z_{t-1}, \ldots, y_{t-1}, \ldots) + \epsilon_t$$

with $\mathbb{E}[\epsilon_t|w_t, w_{t-1}, \ldots, y_{t-1}, \ldots] = 0$, $\mathcal{Z}_t = \sigma(z_t, z_{t-1}, \ldots, y_t, \ldots)$ and $\mathcal{W}_t = \sigma(w_{t+1}, \ldots, y_t, \ldots)$. The decomposition of $Y_t = y_0 + \ldots + y_t$ w.r.t. $\mathcal{W}_t$ writes:

$$Y_t = \sum_{j=1}^{t} \underbrace{\mathbb{E}[y_j|\mathcal{W}_{j-1}]}_{=h_j} + \sum_{j=1}^{t}(y_j - \mathbb{E}[y_j|\mathcal{W}_{j-1}])$$

with $h_t = \sum_{j=1}^{t} h_j$. We must then solve:

$$h_t = \mathbb{E}[\lambda_t(z_t, \ldots, y_{t-1}, \ldots)|\mathcal{W}_t].$$

2.3.2 Example 2 : diffusions

Let us assume that the structural model has the following form :

$$dY_t = \lambda_t(Y_t, Z_t)dt + \sigma_t(Y_t)dB_t \tag{3}$$

where B_t is a Brownian motion. This means that if Z_t is fixed (or randomized, and not generated by the distribution mechanism), Y_t follows a diffusion process with a drift equal to λ_t and a volatility equal to $\sigma_t(Y_t)$. Note that we assume that Z_t does not appear in the volatility term. Let us assume that:

$$\mathbb{E}[dB_t|\mathcal{Y}_t \vee \mathcal{W}_t] = 0.$$

In that case Equation (3) characterizes the DMIV decomposition of Y_t. In order to identify λ_t we need to construct the decomposition of Y_t w.r.t. the filtration $\mathcal{Y}_t \vee \mathcal{W}_t$ (that we write $dY_t = h_t dt + dM_t$) and to solve:

$$h_t = \mathbb{E}[\lambda_t | \mathcal{Y}_t \vee \mathcal{W}_t]. \tag{4}$$

Note that the "reduced form" model $dY_t = h_t dt + dM_t$ has no reason to be a diffusion. Conditionally on $\mathcal{W}_t$, the process may be non Markovian and M_t maybe different from a Brownian motion. This general framework may be applied to particular cases, and simplifies the estimation problem. For example, let's assume that the structural model has an Ornstein-Uhlenbeck form:

$$\lambda_t(Y_t, Z_t) = \theta(\mu(Z_t) - Y_t) \quad \text{and} \quad \sigma_t(Y_t) = \sigma^2$$

where θ is a constant. In that case the model becomes a semi-parametric problem:

$$h_t = \theta(\mathbb{E}[\mu(Z_t)|\mathcal{Y}_t \vee \mathcal{Z}_t] - Y_t).$$

We may project this equation under the σ-field generated by Y_t and W_t, only having then to solve:

$$\mathbb{E}[h_t|\mathcal{Y}_t \vee \mathcal{Z}_t] = \theta(\mathbb{E}[\mu(Z_t)|\mathcal{Y}_t \vee \mathcal{Z}_t] - Y_t).$$

This construction may not be generalized if the volatility depends on Z_t because $\mathbb{E}[\sigma(Y_t, Z_t)dB_t|\mathcal{Y}_t \vee \mathcal{W}_t]$ does not cancel. The change of time models we will present in the next section, will solve this problem as it will be shown in paragraph 3.4.5. An other approach may be however adapted in the same direction as the DMIV decomposition. We briefly introduce this approach which will be treated in an other paper.

Let us start with a structural model (Z fixed or assigned).

$$dY_t = \lambda(Y_t, z)dt + \sigma(Y_t, Z)dB_t$$

which is assumed to be stationary and Z, the endogenous element is assumed to be not time-dependent. Following a method[1] presented by [Aït-Sahalia, 2002], let us introduce the transformation $\tilde{Y}_t = \gamma(Y_t, z) = \int_0^{Y_t} \frac{du}{\sigma(u,z)}$ which lead to the equation:

$$d\tilde{Y}_t = \mu(\tilde{Y}_t, z)dt + db_t$$

where:

$$\mu(\eta, z) = \frac{\lambda(\gamma^{-1}(\eta, z), z)}{\sigma(\gamma^{-1}(\eta, z), z)} - \frac{1}{2}\frac{\partial \sigma}{\partial u}(\gamma^{-1}(\eta, z), z)$$

The model may be completed by two assumptions:

$$\mathbb{E}[dB_t|\mathcal{Y}_t \vee \mathcal{W}_t] = 0$$

$$\mathbb{E}[(dB_t)^2|\mathcal{Y}_t \vee \mathcal{W}_t] = 1$$

which are satisfied in particular if dB_t is independent of $\mathcal{Y}_t \vee \mathcal{W}_t$. This two equations may be used to characterize λ and σ. The main difficulty is coming from the fact that $\tilde{Y}_t$ depends on the parameters and this type of model may be viewed as a dynamic extension to transformation models.

[1]We thank Nour Meddahi for helpful discussions on this topic.

3 The Non-Separable Case : The Time-Change Models

The DMIV decomposition is not sufficient to cover models like counting processes models, or diffusion models with volatility dependent on endogenous variables. We need to propose an other concept for instrumental variables analysis, which we will extend to dynamic models in the non-separable case (treated in the static case by e.g. [Horowitz and Lee, 2007]). In order to motivate our presentation, we start by the basic example of duration models.

3.1 Duration models: a motivating example

Let τ a be duration, i.e. a positive random variable. The distribution of τ is characterized by its survivor function $S(t) = \mathbb{P}(t \leq \tau)$ assumed to be differentiable. Let $\lambda(t)$ denotes the hazard function i.e. (that is $\lambda = -S'/S$) and $\Lambda(t)$ the integrated hazard function ($\Lambda(t) = \int_0^t \lambda(s)ds = -ln(S(t))$). We assume that Λ is strictly increasing. Such a duration model has a counting process representation through the process $N_t = \mathbf{1}\{t \geq \tau\}$. This process is a sub-martingale and then a semi-martingale that may be represented w.r.t. the filtration generated by the history of N_t through:

$$N_t = \int_0^t \lambda(s)\mathbf{1}(s < \tau)ds + M_t. \tag{5}$$

The intensity of N_t (relatively to its history) is equal to $\lambda_t \mathbf{1}\{\tau > t\} = \lambda_t(1 - N_{t^-})$ (see e.g. [Karr, 1991]). A fundamental property we will use in the following is that $\Lambda(\tau)$ has an exponential distribution with parameter 1. Then, if U_t is the counting process $\mathbf{1}\{t \geq \Lambda(\tau)\}$ we have:

$$U_t = \int_0^t \mathbf{1}\{s < \Lambda(\tau)\}ds + M_t^U \tag{6}$$

because the hazard function of the exponential is constant equal to 1. Equivalently, these relations imply that:

$$N_{\Lambda^{-1}(t)} = U_t \tag{7}$$

and the given N becomes the process U via a change of time.

We want now to introduce a random endogenous factor Z in the duration model and an instrument W. For sake of simplicity, both Z and W are not time-dependent in this paragraph. An important literature analyzes endogenous variables in duration models (see [VanDenBerg, 2008]) and is in particular motivated by treatment models where outcomes are durations (see [Abbring and VanDenBerg, 2003]). Our approach does not depend on any specific statistical models and extends the instrumental variable analysis to this problem. It is natural to assume that the integrated hazard function Λ becomes a function $\Lambda(t, Z)$ of Z (also noted $\Lambda_t(Z)$); the "noise" of the model, equal to $\Lambda_\tau(Z)$, is assumed to be independent of the instruments W, and has an exponential distribution with parameter 1. The model may be written in the usual way:

$$\tau = \Phi(U, Z) = \Lambda^{-1}(U, Z) \tag{8}$$

where $\Lambda(., Z)$ is strictly increasing, $U \perp\!\!\!\perp W$ and the distribution of U is given. This model becomes an example of non-separable IV model and generates a non-linear integral equation which characterizes $\Lambda(Z)$ or equivalently $\Phi_t(Z) = \Lambda_t^{-1}(Z)$. Let us consider the following function :

$$S(t, z|w) = \frac{\partial}{\partial z}\mathbb{P}(\tau \geq t, Z \leq z|W = w) \tag{9}$$

which may be seen as the joint survivor of τ and density of Z conditionally on $W = w$, identified by the joint observation of (τ, Z, W). Then the independence condition between U and W implies:

$$\int_Z S(\Phi_t(Z), z|w)dz = P(U \geq t) = e^{-t} \quad (10)$$

because U is exponential with parameter 1. We will discuss later the identification of Φ i.e. the unicity of the solution of this equation.

We may wish to apply the DMIV decomposition to the N_t process considering the two filtrations $\mathcal{N}_t \vee \mathcal{Z}$ and $\mathcal{N}_t \vee \mathcal{W}$, generated by the history of N_t and respectively the endogenous and the instrumental variable. We then obtain a decomposition $N_t = \tilde{\Lambda}_t + \tilde{U}_t$ where $\tilde{\Lambda}_t = \int_0^t \tilde{\lambda}(s,z)ds$ and $\tilde{\lambda}(s,z) = \tilde{\lambda}_0(s,z)\mathbf{1}\{\tau > t\}$. In this context, the function $\tilde{\lambda}_0(s,z)$ should then become the solution of:

$$\frac{f_\tau(t|W)}{S_\tau(t|W)} = \int \tilde{\lambda}_0(s,z) f_\tau(z|w, \tau \geq t)dz \quad (11)$$

where the left hand-side is the hazard function of τ given $W = w$ (in that case $h_t = \frac{f_\tau(t|W)}{S_\tau(t|W)}\mathbf{1}\{\tau > t\}$) and $f_\tau(z|w, \tau \geq t)$ is the conditional density of Z given $W = w$ and the event $\{\tau \geq t\}$. However the $\tilde{\lambda}_0(s,z)$ function is the derivative of $\tilde{\Lambda}_t$ but not the derivative of $\Lambda(t,z)$ we have introduced above, and is not in general the hazard rate of the counting process associated to the duration.

The counting process version of the non-separable model (8) follows from the previous remarks. We may consider $N_t = \mathbf{1}\{\tau \leq t\}$ and assume that there exists a time-change function $\Phi_t(Z)$ strictly increasing and depending on Z such that $N_{\Phi_t(Z)} = U_t$ where U_t is a counting process associated to an exponential distribution of parameter 1 and such that U_t is independent of W. We will see later that these assumptions generates a non-linear integral equation deriving from semi-martingale decompositions which is equivalent in this particular case to Equation (10).

3.2 Time-change models

We use the notations introduced at the beginning of Section 2. We consider a stochastic process Y_t and two filtrations $\mathcal{F}_t = \mathcal{Y}_t \vee \mathcal{Z}_t$ (the "endogenous filtration") and $\mathcal{G}_t = \mathcal{Y}_t \vee \mathcal{W}_t$ (the "instrumental filtration") such that Y_t is adapted to both. We also introduce $\mathcal{H}_t = \mathcal{F}_t \vee \mathcal{G}_t$ generated by the three processes, Y_t, Z_t and W_t.

Definition 3.1. The process Y_t has an instrumental variable non-separable representation if there exists a stochastic process Φ_t such that:

1. $(\Phi_t)_t$ is an increasing sequence of stopping times relatively to the filtration $\mathcal{F}$;
2. $(Y_{\Phi_t})_t$ (the process Y stopped at time Φ_t) is equal to a process U_t independent of the W_t process;
3. U_t is a semi-martingale w.r.t. to its own history ($U_t = H_t^U + M_t^U$) with a given compensator H_t^U.

□

Remember that the property that for $t \geq 0$, Φ_t is a stopping time w.r.t. $\mathcal{F}$ means that $\forall s \geq 0$, $\{\Phi_t \leq s\} \in \mathcal{F}_s$. In the introducing example of the duration model of Section 3.1, Z is not time-dependent and this property only means that Φ_t is measurable w.r.t. Z for any t. The property that

Φ_t is a $\mathcal{F}$-stopping time formalizes the idea that Φ_t only depends on Z and on the past of Y but not on W. However, Assumption (1) of Definition 3.1 implies that Φ_t is also a stopping time for the filtration $\mathcal{H}_t$. An important literature exists in abstract probability theory about the increasing sequences of stopping times and about the properties of processes stopped at these stopping times and the authors usually look at the properties (martingale, local martingale, ...) preserved by the change of time. Examples of this (not very recent) literature are [Kazamaki, 1972], [El-Karoui and Weidenfeld, 1977], [El-Karoui and Meyer, 1977], [LeJan, 1979].

3.3 Identification

Our objective is now to characterize the function Φ_t (depending also on the Z_t process) from objects identified by the obtention of the joint process (Y_t, Z_t, W_t). We adopt a strategy based on the decomposition of the Y_t process w.r.t. the larger filtration $\mathcal{H}$.

Theorem 3.1. Let us assume that:

1. Y_t is a semi-martingale w.r.t. filtration $\mathcal{H}$ and that we have:
$$dY_t = k_t dt + dE_t$$
where $K_t = \int_0^t k_s ds$ is an $\mathcal{H}_t$-predictable process and E_t is an $\mathcal{H}_t$ martingale.
2. Y_t has an instrumental variable non-separable representation as defined in Definition 3.1 when Φ_t is assumed to be continuous and differentiable (possibly except at a discrete set of points).
3. The distribution of $(Z_t)_t$, conditionally on σ-fields $\mathcal{Y}_s \vee \mathcal{W}_s$ for any s, is dominated by a measure Q and has a density denoted $g(z|\mathcal{Y}_s \vee \mathcal{W}_s)$.

Then:
$$\int Q(dz) \int_0^{\Phi_t} k_s g(z|\mathcal{Y}_s \vee \mathcal{W}_s) ds = H_t^U \tag{12}$$
□

This equation shows that Φ_t is the solution of a non-linear integral equation where the right-hand side term is given and all the left-hand side (k and g) are identified by the distribution of the process (Y_t, Z_t, W_t). We assume that the model is well specified or equivalently that a solution exists to the Equation (12). The identification question is concerned with the unicity of the solution. As the problem is non-linear it is natural to look at local unicity of the solution. Let us assume that Φ_t is the true process and we compute the Gateaux-derivative of the left hand-side, taken in Φ_t, in direction of a function $\tilde{\Phi}_t$: $T'_{\Phi_t}(\tilde{\Phi}_t)$. We get obviously for any t:

$$T'_{\Phi_t}(\tilde{\Phi}_t) = \int \tilde{\Phi}_t k_{\Phi_t} g(z|\mathcal{Y}_{\Phi_t} \vee \mathcal{W}_{\Phi_t}) Q(dz).$$

We note that $T'_{\Phi_t}(\tilde{\Phi}_t)$ is linear and we assume that it is equal to the Frechet-derivative. Local unicity is then obtained through the condition :

$$T'_{\Phi_t}(\tilde{\Phi}_t) = 0 \quad \Rightarrow \quad \tilde{\Phi}_t = 0 \quad a.s. \tag{13}$$

If Φ'_t (the derivative w.r.t. t) does not cancel, this implication is true as soon as:

$$\int R_s k_s g(z|\mathcal{Y}_s \vee \mathcal{W}_s) Q(dz) = 0 \quad \Rightarrow \quad R_s = 0 \quad a.s.$$

where $R_s = \tilde{\Phi}_{\Phi^{-1}(s)}$.

3.4 Examples

3.4.1 Duration model with constant covariates

We take here the example of Section 3.1 in the case where variables $Z_t = Z$ and $W_t = W$ are fixed and known at time-origin. We have $Y_t = \mathbf{1}\{t \geq \tau\}$, and we suppose that there exists a sequence $\Phi_t(Z)$ of stopping-times such that $Y_{\Phi_t(Z)} = U_t$ with $U_t = \mathbf{1}\{t \geq U\}$ where U follows an exponential of parameter 1, $U \perp\!\!\!\perp W$. In this framework, we want to use Equation (12) of Theorem (3.1). In this context, k_s is the intensity of Y_t w.r.t. to $\mathcal{H}_t$ (with $\mathcal{H}_t$ equal here to $\sigma(Y_t, Z, W) = \sigma(\{\tau \geq t\}, Z, W)$). Equivalently, $g(z|\mathcal{Y}_s \vee \mathcal{W}_s) = g(z|\tau \geq t, W)$. As Z is a random variable and not a process, $\int Q(dz)$ will then be an integral over the support of Z relatively to the Lebesgue measure. As U is an exponential variable, the compensator of $U_t = \mathbf{1}\{t \geq U\}$ is trivially equal to $H_t^U = t \wedge U$. If we apply Theorem 3.1, we get:

$$\int dz \int_0^{\Phi_{U\wedge t}(z)} k(s|\tau \geq s, z, W) g(z|\tau \geq s, W) ds = U \wedge t. \tag{14}$$

As we will work with a fixed, arbitrary t, we can therefore conceptually eliminate U in all calculations and replace $U \wedge t$ with t. We already had the result of Equation (10) and we want to show that it leads to the same equation than Equation (14). Now, we write $f(t, Z|W)$ the joint law of (τ, Z) conditional on W. Having $\tau = \Phi_u(Z)$, if we note $g(U, Z|W)$ the joint law of (U, Z) conditional on W, we have:

$$g(U, Z|W) = \Phi'_U(Z) \times f(\Phi_U(Z), Z|W).$$

Our main assumption was that $U = \Lambda(\tau, z) \sim \mathcal{E}xp(1)$ conditionally on W. Then, this leads to :

$$e^{-U} = \int g(U, z|W) dz = \int \Phi'_U(z) f(\Phi_U(z), z|W) dz.$$

Then we have the two following expressions, holding $\forall u \geq 0$:

$$\begin{cases} \int \Phi'_u(z) f(\Phi_u(z), z|W) dz &= e^{-u} \\ \int S_\tau(\Phi_u(z), z|W) dz &= e^{-u}. \end{cases}$$

If we divide the first equation by the second, we get:

$$\begin{aligned} 1 &= \int \frac{\Phi'_u(z) f(\Phi_u(z), z|W)}{\int S_\tau(\Phi_u(z'), z'|W) dz'} dz \\ &= \int \Phi'_u(z) \underbrace{\frac{f(\Phi_u(z), z|W)}{S_\tau(\Phi_u(z), z|W)}}_{I_1} \times \underbrace{\frac{S_\tau(\Phi_u(z), z|W)}{\int S_\tau(\Phi_u(z'), z'|W) dz'}}_{I_2}. \end{aligned}$$

I_1 is the hazard function of the process $\{Y_t\}$ taken in $\Phi_u(z)$ conditional on $Z = z, W$. Indeed:

$$\frac{f(t, z|W)}{S(t, z|W)} = \frac{f(t|z, W) f(z|W)}{S(t|z, W) f(z|W)} = \frac{f(t|z, W)}{S(t|z, W)} = k(t|z, W).$$

I_2 is the law of Z conditional to W and $U \geq u$. Finally:

$$\int \Phi'_u(z) k(\Phi_u(z)|Z = z, W) g(z|U \geq u, W) dz = 1.$$

If we integrate in u for u varying from 0 to t we get:

$$\int_0^t du \int_z \Phi'_u(z) k(\Phi_u(z)|Z = z, W) g(z|W, U \geq u) dz = t.$$

If we commute the integral terms and make the change of variable $s = \Phi_u(z)$, and remark that $\{U \geq u\}$ is equivalent to $\{\Phi_U(z) \geq \Phi_u(z)\} = \{\tau \geq s\}$, then we recover Equation (14):

$$\int dz \int_0^{\Phi_t(z)} k(s|\tau \geq s, z, W) g(z|\tau \geq s, W) ds = t.$$

3.4.2 An example of duration model with process covariate

Let $N_t = \mathbf{1}(t \geq \tau)$ the explained process associated to the duration τ and Z_t be an endogenous covariate process assumed to be a jump process: $Z_t = \mathbf{1}\{t \geq \epsilon\}$. The process Z_t may be a treatment equal to 0 up to a random time ϵ and to 1 after. The structural model may be interpreted in the following way: if Z_t is "fixed" or assigned we assume that N_t has a structural hazard function equal to $\lambda_t = \alpha + \beta Z_t$ with $\alpha, \beta > 0$ and its compensator is $\Lambda_t = \int_0^t \lambda_s(Z) ds = \alpha t + \beta(t - \epsilon)\mathbf{1}\{t \geq \epsilon\}$ and then:

$$\Phi_t(Z) = \Lambda_t^{-1}(Z) = \frac{t}{\alpha}\mathbf{1}\{t < \alpha\epsilon\} + \frac{t + \beta\epsilon}{\alpha + \beta}\mathbf{1}\{t \geq \alpha\epsilon\}$$

which is an increasing sequence of stopping times adapted[2] to $\mathcal{Z}_t$.

In that case the model is parametric and the structural parameters are α and β. Let us now consider an instrument constant in time W and we assume that $N_{\Phi_t} = \mathbf{1}\{t \geq u\}$ with $u \sim Exp(1)$ and $U \perp\!\!\!\perp W$. Now consider ρ the hazard rate of τ given the Z_t process and the W variable. We have:

$$\rho(t) = \rho_1(t|\epsilon, W)\mathbf{1}\{\epsilon \leq t\} + \rho_2(t|\epsilon \geq t, W)\mathbf{1}\{\epsilon > t\}$$

where ρ_1 and ρ_2 are the hazard rates of τ given W and respectively ϵ or $\epsilon \geq t$. Then α and β are characterized as the solution of:

$$\int d\epsilon \int_0^{\Phi_t} \rho(s) g(\epsilon|\tau \geq s, W) ds = t \tag{15}$$

where $g(\epsilon|\tau \geq s, W)$ is the conditional density of ϵ given $\tau \geq s$ and W. In this equation ρ and g are identified and α and β follows from the resolution of Equation (15).

3.4.3 Counting process with endogenous cofactor

Let us assume that Y_t is a counting process, i.e. a process valued in $\mathbb{N}$ such that $Y_0 = 0$ and with càdlàg trajectories which are step functions having jumps of size 1 i.e. there exists a sequence of (τ_j) such that:

$$Y_t = \sum_{j \geq 1} \mathbf{1}\{t \geq \tau_j\}.$$

If Z is assumed first to be fixed or assigned at a value $Z = z$ the process Y_t is modelled by its stochastic intensity $\lambda_t(z)$ or by its compensator $\Lambda_t(z) = \int_0^t \lambda_s(z) ds$. It is clear that if $\Lambda_t(z)$ is invertible and if we define:

$$\Phi_t(z) = \Lambda_t^{-1}(z)$$

the process $Y_{\Phi_t(z)} = U_t$ is an homogenous Poisson process. Indeed we have the decomposition:

$$Y_t = \Lambda_t(z) + M_t$$

[2] Indeed let's consider for a given t the event $E = \{\Phi_t(Z) \leq s\}$. If $s \geq \epsilon$, $\mathcal{Z}_s = \sigma\{\epsilon\}$ and then $E \in \mathcal{Z}_s$ for any s. If $s < \epsilon$, $\mathcal{Z}_s = \sigma\{\mathbf{1}(\epsilon > s)\}$. In that case if $t < \alpha\epsilon$, E is always true and if $t \geq \alpha\epsilon$, E is always false.

and

$$Y_{\Phi_t(z)} = t + M_{\Phi_t(z)}.$$

Therefore the compensator is equal to t, which fully characterizes the Poisson process.

If Z is now randomly generated but not necessarily independent of U_t but if U_t is independent of W, we face the situation described in Definition 3.1 . We limit ourself in the following to the case where Z and W are time-independent for sake of simplicity.

We first rewrite in that case the integral equation characterizing $\Phi_t(z)$. Note that the intensity k_t verifies $k_t = \sum_{j\geq 1} k_t^{(j)} \mathbf{1}\{\tau_{j-1} \leq t < \tau_j\}$ with:

$$k_t^{(j)} = \frac{f_j(s+\tau_{j-1}|W,z,\tau_1,\dots,\tau_{j-1})}{S_j(s+\tau_{j-1}|W,z,\tau_1,\dots,\tau_{j-1})}$$

where f_j and S_j are the density and the survivor function of the (difference in) durations $\tau_j - \tau_{j-1}$ conditional to W, Z, and the past of the durations. The equation (12) becomes the following sequence of integral equations:

$$\begin{aligned}\sum_{l=1}^{j} \int dz \int_0^{\tau_l-\tau_{l-1}} \frac{f_l(s+\tau_{l-1}|W,z,\tau_1,\dots,\tau_{l-1})}{S_l(s+\tau_{l-1}|W,z,\tau_1,\dots,\tau_{l-1})} g(z|W,\tau_1,\dots,\tau_{l-1},\tau_l > s+\tau_{l-1})ds \quad &+ \\ \int dz \int_0^{\Phi_t(z)-\tau_j} \frac{f_j(s+\tau_{j-1}|W,z,\tau_1,\dots,\tau_j)}{S_j(s+\tau_{j-1}|W,z,\tau_1,\dots,\tau_j)} g(z|W,\tau_1,\dots,\tau_{j-1},\tau_j > s+\tau_j)ds \quad &= \quad t\end{aligned} \tag{16}$$

One may add that

$$\frac{f_l(s+\tau_{l-1}|W,z,\tau_1,\dots,\tau_{l-1})}{S_l(s+\tau_{l-1}|W,z,\tau_1,\dots,\tau_{l-1})} g(z|W,\tau_1,\dots,\tau_{l-1},\tau_l > s+\tau_{l-1}) = \frac{f_l(s+\tau_{l-1},z|W,\tau_1,\dots,\tau_{l-1})}{S_l(s+\tau_{l-1}|W,\tau_1,\dots,\tau_{l-1})}.$$

All the elements inside the integral may be estimated and this sequence of integral equations characterizes $\Phi_t(z)$ by intervals. Let us now analyze in more details the nature of the function $\Phi_t(z)$ and come back to the structural model where Z is fixed or assigned. In this structural model the $\lambda_t(z)$ function takes the form $\lambda_t(z) = \lambda_{t-\tau_{j-1}^{(j)}(z}$ for $t \in]\tau_{j-1};\tau_j]$ where $\lambda_t^{(j)}(z)$ is the hazard rate of $\tau_j - \tau_{j-1}$ conditional on the past $(\tau_1,\dots,\tau_{j-1})$ and given z. Then $\Lambda_t(z) = \int_0^t \lambda_s(z)ds$ which implies that $\Lambda_t(z) = \Lambda_{\tau_{j-1}}(z) + \int_0^{t-\tau_{j-1}} \lambda_{s-\tau_{j-1}}^{(j)}(z)ds$ if $\tau_{j-1} < t \leq \tau_j$. From this follows:

$$\text{if } \Lambda_{\tau_{j-1}}(z) < t \leq \Lambda_{\tau_j}(z) \text{ , then } \Phi_t(z) = \tau_{j-1} + (\Lambda_{t-\Lambda_{\tau_{j-1}}(z)}^{(j)})^{-1}(z)$$

where $\Lambda_t^{(j)}(z)$ is the integral of $\lambda^{(j)}(z)$.

In practice, $\Phi_t(Z)$ will be selected such that some properties are satisfied in the model when Z is fixed. For example, Y_t may be in that case an accelerated life non homogenous Poisson process i.e.:

$$Y_t = F(\psi(Z)t) + M_t$$

where $\psi(Z)$ is a function depending on the variables Z and F is a baseline, cumulative function on $\mathbb{R}^+$. In that case we have obviously:

$$\Phi_t(Z) = \frac{F^{-1}(t)}{\psi(Z)}$$

depending on the functional parameters ψ and F. Note however that this assumption **does not imply** that Y_t given Z and W is a Poisson process.

An other example of structural modelling is given by the Hawkes process. Let us assume that for Z fixed Y_t is an Hawkes process whose intensity is:

$$\lambda_t(z) = \mu + \int_0^t g_z(t,s)dY_s$$

where the parameters are μ and g function of Z, t and s. For example g may take the semi-parametric form:

$$g_z(t,s) = e^{-\beta(Z)(t-s)}$$

where β is an unknown positive function of Z. More generally, Z may be a stochastic process and g may be equal to $g_z(t,s) = e^{-\beta(Z_t)(t-s)}$ or $e^{-\beta(Z_s)(t-s)}$. For simplicity, we concentrate our presentation to the case where Z is constant w.r.t. the time index. The compensator of Y_t for any fixed value of $Z = z$ is equal to:

$$\begin{aligned}\Lambda_t(z) &= \mu t + \int_0^t du \int_0^u g_z(u,s)dN_s \\ &= \mu t + \int_0^t dN_s \int_s^t g_z(u,s)du \\ &= \mu t + \sum_{j=1}^{N_t} \int_{\tau_j}^t g_z(u,\tau_j)du\mathbf{1}\{t \geq \tau_j\}\end{aligned}$$

The inverse function $\Phi_t(z) = \Lambda_t^{-1}(z)$ has not an explicit form but may be easily numerically computed if g is given and Theorem 3.1 gives the way to estimate $\Phi_t(z)$ and then $g_z(t,u)$. As in the Poisson case let us note that Y_t given Z and W is not in general an Hawkes process.

3.4.4 Markovian transition models

An other application could concern Markov processes with multiple states. We begin by considering a Markov process Y_t with two states $\{1,2\}$. We write I^Y the generator of Y and suppose that I^Y has the form $q_t(Z)I$ where I is the following matrix:

$$I = \begin{bmatrix} -1 & 1 \\ a & -a \end{bmatrix}$$

where $a \in \mathbb{R}_+^*$. We note $Q_t(Z) = \int_0^t q_s(Z)ds$. Z is assumed here to be static, endogenous. We assume that there exists a change of time $\Phi(Z) = \Lambda^{-1}(Z)$ such that $Y_t = U_{\Lambda_t(Z)}$ where U_t is a homogenous Markov process with two states and with a generator I. We make the assumption that U_t is independent from given instruments W. In the following we will skip the indexation in Z for simplicity (Z will be assumed to be fixed or assigned). It is possible to show[3] that:

$$\Phi(t) = \Lambda^{-1}(t) = Q^{-1}\Big(\frac{1-e^{(1+a)t}}{1+a}\Big).$$

[3] See Appendix .

We verify easily that this function is increasing in t.

We now consider the counting processes $N^{12}(t)$ and $N^{21}(t)$ that jump when respectively the process Y_t jumps from state 1 to 2 conditional on the fact that Y_t is in 1, and when Y_t jumps from state 2 to 1 conditional on the fact that Y_t is in state 2. We remark that in the general case, Y_t conditional on Z, W has no reason to remain Markovian. We note k_s^{12} (respectively k_s^{21}) the intensity of N^{12} (resp. N_s^{21}) conditional on Z, W, $\mathcal{H}_t$. Applying Theorem 3.1 we get:

$$\int_z \int_0^{\Phi_t(z)} k_s^{12} g(z|W, \mathcal{Y}_t) ds = \int_0^t \mathbf{1}\{U_s = 1\} ds,$$
$$\int_z \int_0^{\Phi_t(z)} k_s^{21} g(z|W, \mathcal{Y}_t) ds = \int_0^t a \mathbf{1}\{U_s = 2\} ds.$$

These equations are not useful because U_t is not observed and this right-hand side cannot be computed. But dividing the second equation by a, summing both and remarking that for each s, $\mathbf{1}\{U_s = 1\} + \mathbf{1}\{U_s = 2\} = 1$, then we get:

$$\int_z \int_0^{\Phi_t(z)} (k_s^{12} + \frac{1}{a} k_s^{21}) g(z|W, \mathcal{Y}_t) ds = t.$$

3.4.5 The diffusion model

Let us first consider a structural model which generates a zero-mean diffusion process for Z (assumed to be time-independent) fixed:

$$dY_t = \sigma(Y_t, Z) dB_t \tag{17}$$

We simplify our presentation by assuming σ independent from t. Let us consider the quadratic variation of Y_t:

$$\Lambda_t(Z) = < Y_t > = \int_0^t \sigma^2(Y_s, Z) ds.$$

We define $\Phi_t(Z)$ the inverse function of $\Lambda(Z)$ (which is invertible because σ is assumed not null for any Y_t). This function characterizes an increasing sequence of stopping times (the event $\Phi_t(Z) \leq s$ is equivalent to $t \leq \Lambda_s(Z)$, and only depends on the past of Y until s). The process:

$$Y_{\Phi_t(Z)} = U_t$$

is then a Brownian motion (see [Protter, 2003]). We now consider that Z is randomly generated and that W is an instrument. The model still assumes Equation (17) and that the process U is independent of the filtration $\mathcal{W}_t$ generated by W and the past of W. In order to characterize σ or Φ we applied Theorem 3.1 to the relation:

$$Y_{\Phi_t(Z)}^2 = U_t^2. \tag{18}$$

The compensator of U_t^2 is equal to t. Let k the stochastic intensity of Y_t^2 w.r.t. $\mathcal{Z}_t \vee \mathcal{W}_t$. We have:

$$\int dz \int_0^{\Phi_t(z)} k_s g_s(z|\mathcal{W}_s) ds = t.$$

In this expression k and g are identifiable for the DGP and Φ is obtained by solving this nonlinear, integral equation.

This approach may be generalized by considering a process Z_t instead of a fixed value Z if we assume σ depending only on the past up to t of Z (e.g. $\sigma(Y_t, Z_t)$) and $\mathcal{W}_t$ may be a filtration generated by a process W_t and Y_t. Let us underline that even if the structural model 17 is a zero-mean diffusion, this is in general not the case for the process Y_t given the filtration $\mathcal{Z}_t \vee \mathcal{W}_t$ and even if k is identifiable, its estimate may be complex.

An other extension is to consider a structural model with drift; if Z is fixed or assigned we assume the model:

$$dY_t = m_Z(Y_t, Z)dt + \sigma(Y_t, Z)dB_t.$$

We consider the same stopping time as before and the sequence of equations:

$$Y_{\Phi_t(Z)} = \int_0^{\Phi_t(Z)} m(Y_s, Z)ds + U_t.$$

The parameters of the model are $\Phi_t(Z)$ and $m(Y_s, Z)$ and in the case where Z is random we assume that U_t is a Brownian motion independent of W_t. We then apply twice Theorem 3.1: we compute the stochastic intensities $k_s^{(1)}$ of $Y_t - \int_0^t m(Y_s, Z)ds$ w.r.t. $\mathcal{Z}_t \vee \mathcal{W}_t$ and $k_s^{(2)}$ of $\left[Y_t - \int_0^t m(Y_s, Z)ds\right]^2$ w.r.t. $\mathcal{Z}_t \vee \mathcal{W}_t$ also and we derive from Theorem 3.1 that:

$$\int dz \int_0^{\Phi_t(z)} k_s^{(1)} g_s(z|\mathcal{W}_s)ds = 0$$
$$\int dz \int_0^{\Phi_t(z)} k_s^{(2)} g_s(z|\mathcal{W}_s)ds = t.$$

The functional parameters σ and m are solution of this system of nonlinear equations.

4 CONCLUSION

We have presented two classes of models for stochastic processes with endogenous variables treated with the instrumental variables method. Dynamic extension of separable models gives a generalization of the standard Doob-Meyer decomposition of semi-martingales and some probabilistic aspects of this model should be developed (extension for example to the case when martingales are only local). In the two kinds of approaches the functional parameters of interest are characterized as solutions of integral equations and their identification (unicity of the solution) is discussed. We have illustrated these concepts to many kinds if stochastic processes used in many fields of applied econometrics. All these examples need to be developed in connection with the infinitesimal generator.

This paper only treats modelling and not the practical aspects and the theoretical properties of the inference. In practice, many objects we have introduces depend on infinite past and cannot be estimated under this form. We have introduced models where the specification is made on the structural form and reduced forms are implicitly left unconstrained for the estimation. Tractable approximations for the reduced form should be selected in order to implement the presented methods. In the considered cases, the parameters are solutions of ill-posed inverse problems and their statistical properties have to be analyzed.

A APPENDIX

A.1 Proof of Theorem 2.1

Let us define U_t by $Y_t - \int_0^t \lambda_s ds$. As $\Lambda_t = \int_0^t \lambda_s ds$ is predictable by construction, we just have to proove that U_s satisfies condition **A2** of Definition 2.1. We have:

$$\begin{aligned} \mathbb{E}[U_t - U_s|\mathcal{Y}_s \vee \mathcal{W}_s] &= \mathbb{E}[Y_t - Y_s - \int_s^t \lambda_u du|\mathcal{Y}_s \vee \mathcal{W}_s] \\ &= \mathbb{E}[\int_s^t h_u du - \int_s^t \lambda_u du|\mathcal{Y}_s \vee \mathcal{W}_s] \end{aligned}$$

because $\mathbb{E}[M_t - M_s|\mathcal{Y}_s \vee \mathcal{W}_s] = 0$. We can commute the integration and the conditional expectation terms, and we get:

$$\int_s^t \mathbb{E}[h_u - \lambda_u|\mathcal{Y}_s \vee \mathcal{W}_s]du = \int_s^t \mathbb{E}[h_u - \mathbb{E}(\lambda_u|\mathcal{Y}_u \vee \mathcal{W}_u)|\mathcal{Y}_s \vee \mathcal{W}_s]du$$

because $\mathcal{Y}_s \vee \mathcal{W}_s \subset \mathcal{Y}_u \vee \mathcal{W}_u$ for each $s \leq u$. The second assumption allows then to conclude and to obtain the desired result $\mathbb{E}[U_t - U_s|\mathcal{Y}_s \vee \mathcal{W}_s] = 0$ and Y_t has a DMIV decomposition.

□

A.2 Proof of Theorem 3.1

Let us start with the decomposition of Y_t w.r.t. $\mathcal{H}_t$:

$$Y_t = K_t + E_t. \tag{19}$$

We consider $(\mathcal{H}_{\Phi_t})_t$ the filtration where for any t, $\mathcal{H}_{\Phi_t}$ is the stopping-time sub σ-field of $\mathcal{H}_\infty$ associated to Φ_t, i.e. :

$$\mathcal{H}_{\Phi_t} = \sigma\{A \in \mathcal{H}_\infty|\{\Phi_t < s\} \cap A \in \mathcal{H}_s\}. \tag{20}$$

Note that $(\mathcal{H}_{\Phi_t})_t$ is a filtration because Φ_t is increasing. Equivalently (see [Protter, 2003] - Chap. I - Theorem 6):

$$\mathcal{H}_{\Phi_t} = \sigma\{\mathcal{Y}_{\Phi_t}, \mathcal{Z}_{\Phi_t}, \mathcal{W}_{\Phi_t}\}. \tag{21}$$

Then:

$$Y_{\Phi_t} = K_{\Phi_t} + E_{\Phi_t}$$

is the semi-martingale decomposition of $(Y_{\Phi_t})_t$ w.r.t. the filtration $(\mathcal{H}_{\Phi_t})_t$. This result follows from Proposition 1 of [Kazamaki, 1972] which implies that E_{Φ_t} remains a martingale w.r.t. $(\mathcal{H}_{\Phi_t})_t$ and K_{Φ_t} is predictable under our assumption $K_t = \int_0^t k_s ds$. The continuity condition of Φ_t is obviously satisfied under our assumptions. Under the model specification $Y_{\Phi_t} = U_t$ and :

$$\mathcal{H}_{\Phi_t} = \mathcal{U}_t \vee \mathcal{Z}_{\Phi_t} \vee \mathcal{W}_{\Phi_t} \tag{22}$$

where $\mathcal{U}_t$ is the σ-field generated by $(U_s)_{0 \leq s \leq t}$. Then the decomposition (21) rewrites :

$$U_t = K_{\Phi_t} + E_{\Phi_t} \tag{23}$$

and is also the semi-martingale decomposition of U_t w.r.t. $(\mathcal{U}_t \vee \mathcal{Z}_{\Phi_t} \vee \mathcal{W}_{\Phi_t})$. Equivalently Equation (23) becomes:

$$U_t = \int_0^t \Phi_s' k_{\Phi_s} ds + E_{\Phi_t} \tag{24}$$

where Φ_t' is the derivative w.r.t. t of Φ_t.

If Φ_t is not differentiable at some point, we partition the integral before and after the point. For simplicity we assume here Φ to be differentiable.

The next step is to derive from Equation (24) the decomposition of U_t w.r.t. the sub-filtration $(\mathcal{U}_t \vee \mathcal{W}_{\Phi_t})_t$. We have (see [Karr, 1991]):

$$U_t = \int_0^t \mathbb{E}[\Phi_s' k_{\Phi_s} | \mathcal{U}_s \vee \mathcal{W}_{\Phi_s}] ds + \tilde{E}_t$$

where $\tilde{E}_t$ is a martingale adapted to $(\mathcal{U}_t \vee \mathcal{W}_{\Phi_t})_t$. The computation of the conditional expectation inside the integral may be conventionally written as an integral w.r.t. a conditional density of Z_t process given $\mathcal{U}_s \vee W_{\Phi_s}$, noted $g(z|\mathcal{U}_s \vee \mathcal{W}_{\Phi_s})$:

$$U_t = \int_0^t ds \int \Phi_s' k_{\Phi_s} g(z|\mathcal{U}_s \vee W_{\Phi_s}) dz + \tilde{E}_t$$

We commute the integrals and, after a change of variable $v = \Phi_s$, we get:

$$U_t = \int_0^t Q(dz) \int_0^{\Phi_t} k_v g(z|\mathcal{Y}_v \vee W_v) dv + \tilde{E}_t.$$

Finally let us consider the decomposition of U_t w.r.t. its own filtration:

$$U_t = H_t^U + E_t^U.$$

As $(U_t)_t$ and $(W_t)_t$ are independent, (U_t) and (W_{Φ_t}), are also independent and this last decomposition is also the decomposition w.r.t. $(\mathcal{U}_t \vee \mathcal{W}_{\Phi_t})$. By unicity of the decomposition we get:

$$\int dz \int_0^{\Phi_t} k_v g(z|\mathcal{Y}_v \vee \mathcal{W}_v) dv = H_t^U.$$

□

A.3 Expression of Φ_t for Markov models with two states

By definition, we have that $\mathbb{P}[U_{t+s}|U_t] = e^{Is}$. We remark that I has for eigenvalues 0 and $-(1+a)$ and U writes $U = DLD^{-1}$ with D the matrix of eigenvectors which are respectively $(1;1)'$ and $(-1;1)'$. It follows that for $t, s \geq 0$:

$$\mathbb{P}[Y_{t+s}|Y_t] = \mathbb{P}[U_{\Lambda(t+s)}|U_{\Lambda(t)}] = e^{I(\Lambda(t+s)-\Lambda(s))}.$$

Matrix $e^{I(\Lambda(t+s)-\Lambda(s))}$ rewrites:

$$D \begin{bmatrix} 1 & 0 \\ 0 & e^{-(1+a)(\Lambda(t+s)-\Lambda(s))} \end{bmatrix} D^{-1}.$$

The generator I^Y of Y is then given by the derivative in s, taken in $s = 0$ of the former expression. That is:

$$I^Y = \begin{bmatrix} 0 & 0 \\ 0 & -(1+a)\Lambda'(t)e^{-(1+a)\Lambda(t)} \end{bmatrix} D^{-1} = \Lambda'(t)e^{-(1+a)\Lambda(t)} DLD^{-1} = \Lambda'(t)e^{-(1+a)\Lambda(t)} I.$$

Consequently, the generator matrix I^Y is of the form $q(t)I$ with $q(t) = \Lambda'(t)e^{-(1+a)\Lambda(t)}$. Then we have that

$$Q(t) = \frac{1}{1+a}\Big(1 - e^{-(1+a)\Lambda(t)}\Big).$$

The expression of Φ follows.

□

References

[Abbring and VanDenBerg, 2003] Abbring, J. and VanDenBerg, G. (2003). The Nonparametric Identification of Treatment Effects in Durations Models.

[Aït-Sahalia, 2002] Aït-Sahalia, Y. (2002). Maximum-Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach. *Econometrica*, 70(1):223–262.

[Carrasco, 2008] Carrasco, M. (2008). A Regularization Approach to the Many Instruments Problem. *Mimeo Montreal*.

[Carrasco et al., 2003] Carrasco, M., Florens, J.-P., and Renault, E. (2003). Estimation of the Solution of an Integral Equation of the Second Kind.

[Chernozhukov et al., 2007a] Chernozhukov, V., Gaglliardini, P., and Scaillet, O. (2007a). Nonparametric Instrumental Variable Estimators of Quantile Structural Effects. *HEC, Swiss Finance Institue, Technical Report*.

[Chernozhukov et al., 2007b] Chernozhukov, V., Imbens, G., and Newey, W. (2007b). Instrumental Variable Identification and Estimation of Nonseparable Models via Quantile Conditions. *Journal of Econometrics*, 139:4–14.

[Darolles et al., 2010] Darolles, S., Fan, Y., Florens, J.-P., and Renault, E. (2010). Nonparametric Instrumental Regression. *GREMAQ Discussion Paper*.

[Dellacherie and Meyer, 1971] Dellacherie, C. and Meyer, P.-A. (1971). Probabilités et Potentiels. *Hermann*.

[d'Haultfoeuille, 2008] d'Haultfoeuille, X. (2008). On the Completeness Condition in Nonparametric Instrumental Problems. *ENSAE, CREST-INSEE, University Paris I Panthéon-Sorbonne. Forthcoming in Econometric Theory*.

[El-Karoui and Meyer, 1977] El-Karoui, N. and Meyer, P.-A. (1977). Les Changements de Temps en Théorie Générale des Processus. *Séminaire de probabilités (Strasbourg), Springer-Verlag New-York*, pages 65–78.

[El-Karoui and Weidenfeld, 1977] El-Karoui, N. and Weidenfeld, G. (1977). Théorie Générale et Changements de Temps. *Séminaire de probabilités (Strasbourg), Springer-Verlag New-York*, 11:79–108.

[Engle et al., 1983] Engle, R., Hendry, D., and Richard, J.-F. (1983). Exogeneity. *Econometrica*, 51(2):277–304.

[Florens, 2003] Florens, J.-P. (2003). Inverse Problems in Structural Econometrics. The Example of Instrumental Variables. *In Dewatripont M., Hansen L.P., Turnovsky, S.J. (Eds), Advances in Economics and Econometrics, Volume 2, Cambridge University Press, Cambridge*.

[Florens and Fougère, 1996] Florens, J.-P. and Fougère, D. (1996). Noncausality in Continuous Time. *Econometrica*, 64(5):1195–1212.

[Florens and Mouchart, 1982] Florens, J.-P. and Mouchart, M. (1982). A Note On Noncausality. *Econometrica*, 50(3):583–592.

[Florens et al., 1990] Florens, J.-P., Mouchart, M., and Rolin, J.-M. (1990). Elements of Bayesian Statistics. *Marcel Dekker, New-York*.

[Hall and Horowitz, 2005] Hall, P. and Horowitz, J. (2005). Non Parametric Methods for Inference in the Presence of Instrumental Variables. *Annals of Statistics*, 33(6):2904–2929.

[Horowitz and Lee, 2007] Horowitz, J. and Lee, S. (2007). Nonparametric Instrumental Variables Estimation of a Quantile Regression Model. 75:1191–1208.

[Karr, 1991] Karr, A. (1991). Point Processes And Their Statistical Inference. *Marcel Dekker Inc., New York and Basel*.

[Kazamaki, 1972] Kazamaki, N. (1972). Changes of Time, Stochastic Integrals and Weak Martingales. *Probability Theory and Related Fields, Springer, Berlin*, 22(1):25–32.

[LeJan, 1979] LeJan, Y. (1979). Martingales et Changements de Temps. *Séminaire de Probabilités, XIII, (Univ. Strasbourg, 1977), Lecture Notes in Mathematics, 721, Springer, Berlin*, pages 385–399.

[Newey and Powell, 2003] Newey, W. and Powell, J. (2003). Instrumental Variable Estimation of Nonparametric Models. *Econometrica*, 71:1565–1578.

[Protter, 2003] Protter, P. (2003). Stochastic Integration and Differential Equations (2nd Edition). *Springer Verlag, New-York*.

[VanDenBerg, 2008] VanDenBerg, G. (2008). Duration Models: Specification, Identification and Multiple Durations. *Handbook of Econometrics, Chapter 55*, 6B.

[Wang and Phillips, 2009] Wang, Q. and Phillips, P. (2009). Structural Nonparametric Cointegrating Regression. *Econometrica*, 77(6):1901–1948.