$$\begin{aligned}&\text{ATE}(\mathbf{x};t)=\text{E}(y_1-y_0\,|\,\mathbf{x},t)\\&\text{ATET}(\mathbf{x};t>0)=\text{E}(y_1-y_0\,|\,\mathbf{x},t>0)\\&\text{ATENT}(\mathbf{x};t=0)=\text{E}(y_1-y_0\,|\,\mathbf{x},t=0)\end{aligned} \qquad (2)$$

where ATE indicates the overall average treatment effect, ATET the average treatment effect on treated, and ATENT the one on untreated units. By the law of iterated expectation (LIE), we know that the population unconditional ATEs are obtained as:

$$\begin{aligned}&\text{ATE}=\text{E}_{(\mathbf{x};t)}\{\text{ATE}(\mathbf{x};t)\}\\&\text{ATET}=\text{E}_{(\mathbf{x};t>0)}\{\text{ATE}(\mathbf{x};t>0)\}\\&\text{ATENT}=\text{E}_{(\mathbf{x};t=0)}\{\text{ATE}(\mathbf{x};t=0)\}\end{aligned} \qquad (3)$$

where $\text{E}_{\mathbf{z}}(\cdot)$ identifies the mean operator taken over the support of a generic vector of variables **z**. By assuming a linear-in-parameters parametric form for $g_0(\mathbf{x})=\mathbf{x}\boldsymbol{\delta}_\mathbf{0}$ and $g_1(\mathbf{x})=\mathbf{x}\boldsymbol{\delta}_1$ the Average Treatment Effect (ATE) conditional on **x** and *t* becomes:

$$\text{ATE}(\mathbf{x},t,w)=w\cdot[\mu+\mathbf{x}\boldsymbol{\delta}+h(t)]+(1-w)\cdot[\mu+\mathbf{x}\boldsymbol{\delta}] \qquad (4)$$

where $\mu=(\mu_1-\mu_0)$ and $\boldsymbol{\delta}=(\boldsymbol{\delta}_1-\boldsymbol{\delta}_0)$ and the unconditional Average Treatment Effect (ATE) related to model (1) is equal to:

$$\text{ATE}=p(w=1)\cdot(\mu+\overline{\mathbf{x}}_{t>0}\boldsymbol{\delta}+\overline{h}_{t>0})+p(w=0)\cdot(\mu+\overline{\mathbf{x}}_{t=0}\boldsymbol{\delta})$$

where $p(\cdot)$ is a probability, and $\overline{h}_{t>0}$ is the average of the response function taken over $t > 0$. Since, by LIE, we have that ATE = $p(w=1)\cdot$ATET + $p(w=0)\cdot$ATENT, we obtain from the previous formula that:

$$\begin{cases}\text{ATE}=p(w=1)(\mu+\overline{\mathbf{x}}_{t>0}\boldsymbol{\delta}+\overline{h}_{t>0})+p(w=0)(\mu+\overline{\mathbf{x}}_{t=0}\boldsymbol{\delta})\\ \text{ATET}=\mu+\overline{\mathbf{x}}_{t>0}\boldsymbol{\delta}+\overline{h}_{t>0}\\ \text{ATENT}=\mu+\overline{\mathbf{x}}_{t=0}\boldsymbol{\delta}\end{cases}$$

where the dose-*response function* is given by averaging ATE(**x**, *t*) over **x**:

$$\text{ATE}(t)=\begin{cases}\text{ATET}+(h(t)-\overline{h}_{t>0}) & \text{if} \quad t>0\\ \text{ATENT} & \text{if} \quad t=0\end{cases} \qquad (6)$$

that is a function of the treatment intensity t. The estimation of equation (6) under different identification hypothesises is the main purpose of next sections.

3. THE REGRESSION APPROACH

In this section we consider the conditions for a consistent estimation of the causal parameters defined in (2) and (3) and thus of the dose-response function in (6).

What it is firstly needed, however, is a consistent estimation of the parameters of the potential outcomes in (1) – we call here "basic" parameters – as both ATEs and the dose-response function are functions of these parameters.

Under previous definitions and assumptions, and in particular the form of the potential outcomes in model (1), to be substituted into Rubin's potential outcome equation $y_i=y_{0i}+w(y_{1i}-y_{0i})$, the following *Baseline random-coefficient regression* can be obtained (Wooldridge, 1997; 2003):