An example of a linear model with instrumental variables

Consider the process $(y_t)_{t\in\mathbb{Z}}$ defined by $y = b_0 x_t + \epsilon_t$, $b_0 \in \mathbb{R}$, where $(\epsilon_t)_{t\in\mathbb{Z}}$ is a strictly stationary and ergodic white noise process $(WN(\sigma^2))$ and $(x_t)_{t\in\mathbb{Z}}$ is a linear process based on the white noise process $(\epsilon_t)_{t\in\mathbb{Z}}$ thus for a real sequence $(a_j)_{j\in\mathbb{N}}$, $\sum_{j=0}^{\infty}|a_j| < +\infty$, $a_j, a_{j+1} \neq 0$ for some $j \in \mathbb{N}$, $x_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j}$.
Given a sample $(y_t, x_t)_{t=1,\dots,T}$, suppose that our interest centers on estimating the unknown b_0. Propose an estimator for b_0 and derive its asymptotic properties.

Let's consider the OLSE for b_0:

$$b_T = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2} = \frac{\sum_{t=1}^{T} x_t(b_0 x_t + \epsilon_t)}{\sum_{t=1}^{T} x_t^2} = b_0 + \frac{\frac{1}{T}\sum_{t=1}^{T} x_t \epsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}$$

Given the results in the paragraph Transformations and Heridarity and the ones about Causal Linear Processes, since $(\epsilon_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic, we have that $(x_t)_{t\in\mathbb{Z}}$, $(x_t^2)_{t\in\mathbb{Z}}$, $(x_t\epsilon_t)_{t\in\mathbb{Z}}$ are strictly stationary and ergodic processes.

Also,

$$E(x_t\epsilon_t) = E\left[\left(\sum_{j=0}^{\infty} a_j \epsilon_{t-j}\right)\epsilon_t\right] = E\left(\sum_{j=0}^{\infty} a_j \epsilon_{t-j}\epsilon_t\right) = \sum_{j=0}^{\infty} a_j E(\epsilon_{t-j}\epsilon_t)$$

$$= a_0 E(\epsilon_t^2) + \sum_{j=1}^{\infty} a_j \underbrace{E(\epsilon_{t-j}\epsilon_t)}_{=0} = a_0\sigma^2 < +\infty \qquad \text{since } \epsilon \text{ is white noise}$$

Obviously, we have an endogeneity issue here, as the error term ϵ_t is not orthogonal to the regressor x_t.

Moreover,

$$E(x_t^2) = E\left[\left(\sum_{j=0}^{\infty} a_j \epsilon_{t-j}\right)^2\right] = E\left[\sum_{j=0}^{\infty}\sum_{j^*=0}^{\infty} a_j a_{j^*} \epsilon_{t-j}\epsilon_{t-j^*}\right] = \sum_{j,j^*=0}^{\infty} a_j a_{j^*} E(\epsilon_{t-j}\epsilon_{t-j^*})$$

since ϵ is white noise $E(\epsilon_{t-j}\epsilon_{t-j^*}) = \begin{cases} 0, & j \neq j^* \\ \sigma^2, & j = j^* \end{cases}$ hence,

$$E(x_t^2) = \sum_{j=j^*=0}^{\infty} a_j^2 E(\epsilon_{t-j}^2) = \sigma^2 \sum_{j=0}^{\infty} a_j^2 < +\infty. \quad \text{since } \sum_{j=0}^{\infty}|a_j| < +\infty.$$

Recall, Birkhoff's Law of Large Numbers:
Suppose that $(x_t)_{t\in\mathbb{Z}}$ is stationary ergodic, and that $E(|x_0|) < +\infty$ exists as a real number. Then, $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{\infty} x_t \to E(x_0)$ $\mathbb{P}$ a.s.

Thus, by Birkhoff's LLN:

- $\frac{1}{T}\sum_{t=0}^{\infty} x_t \epsilon_t \to E(x_0\epsilon_0) = a_0 \sigma^2$ $\mathbb{P}$ a.s.

- $\frac{1}{T}\sum_{t=0}^{\infty} x_t^2 \to E(x_0^2) = \sigma^2 \sum_{j=0}^{\infty} a_j^2 > 0$ $\mathbb{P}$ a.s.

and thereby due to the Continuous Mapping Theorem:

$$b_T = b_0 + \frac{\frac{1}{T}\sum_{t=0}^{\infty} x_t\epsilon_t}{\frac{1}{T}\sum_{t=0}^{\infty} x_t^2} \to b_0 + \underbrace{\frac{a_0\sigma^2}{\sigma^2\sum_{j=0}^{\infty} a_j^2}}_{\neq 0} \quad \mathbb{P} \text{ a.s.}$$

Therefore, the OLSE for b_0 is inconsistent.
To overcome the endogeneity problem we use the so-called instumental variables estimator (IV). This approach assumes the existence of a variable z which we call instrument, that is correlated with the endogenous variable $Cov(z_t, x_t) \neq 0$ but exogenous to the error term i.e $Cov(z_t \epsilon_t) = 0$.
Given the structure of x_t, a valid instrument in our case would be x_{t-1} defined by $x_{t-1} = \sum_{j=0}^{\infty} a_j \epsilon_{t-1-j}$, since:

- $Cov(x_t, x_{t-1}) = E(x_t x_{t-1}) = \sigma^2 \sum_{j=0}^{\infty} a_j a_{j+1} < +\infty$

given that we have already proved that for a linear process $(x_t)_{t\in\mathbb{Z}}$, $Cov(x_t, x_{t-k}) = E(\epsilon_0^2)\sum_{j=0}^{\infty} a_j a_{j+k}$ and $\sum_{j=0}^{\infty} |a_j a_{j+k}| < +\infty$

- $Cov(x_{t-1}, \epsilon_t) = E(x_{t-1}\epsilon_t) = E\left[\left(\sum_{j=0}^{\infty} a_j \epsilon_{t-1-j}\right)\epsilon_t\right] = \sum_{j=0}^{\infty} a_j E(\epsilon_{t-1-j}\epsilon_t) = 0$

since $(\epsilon_t)_{t\in\mathbb{Z}}$ is a WN.

Hence, x_{t-1} is a valid instrument which enables the definition of a moment condition which along with an identification condition and the analogy principle, give us enough structure so as a criterior function can be constructed, upon the optimization of which, we can design the IV estimation (see Tutorial - Linear model with instrumental variables).
Remember that the IV estimator is denoted by:

$$b_{IV} = (Z'X)^{-1} Z'Y = (Z'X)^{-1} Z'(Xb_0 + \epsilon) = b_0 + (Z'X)^{-1} Z'\epsilon$$
$$= b_0 + \left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t\right)^{-1} \frac{1}{T}\sum_{t=1}^{T} z_t \epsilon_t$$

where z_t in our case is x_{t-1}, therefore given that we also know x_0,

$$b_{IV} = b_0 + \left(\frac{1}{T}\sum_{t=1}^{T} x_{t-1} x_t\right)^{-1} \frac{1}{T}\sum_{t=1}^{T} x_{t-1} \epsilon_t$$

Given the results in the paragraph Transformations and Heridarity, since $(\epsilon_t)_{t\in \mathbb{Z}}$ is stationary ergodic, we have that $(x_{t-1}x_t)_{t\in \mathbb{Z}}$ and $(x_{t-1}\epsilon_t)_{t\in \mathbb{Z}}$ are stationary ergodic processes. Also, $E(x_t x_{t-1}) < +\infty$ and $E(x_{t-1}\epsilon_t) < +\infty$.
Hence, by Birkhoff's LLN:

- $\frac{1}{T}\sum_{t=1}^{T} x_{t-1} x_t \longrightarrow E(x_{-1}x_0) = \sigma^2 \sum_{j=0}^{\infty} a_j a_{j+1}$ $\mathbb{P}$ a.s.

- $\frac{1}{T}\sum_{t=1}^{T} x_{t-1} \epsilon_t \longrightarrow E(x_{-1}\epsilon_0) = 0$ $\mathbb{P}$ a.s

Since $\sum_{j=0}^{\infty} a_j a_{j+1} \neq 0$ since $a_j, a_{j+1} \neq 0$ for some $j \in \mathbb{N}$, due to the CMT:

$$b_{IV} = b_0 + \frac{\frac{1}{T}\sum_{t=1}^{T} x_{t-1}\epsilon_t}{\frac{1}{T}\sum_{t=1}^{T} x_{t-1} x_t} \longrightarrow b_0 + \frac{0}{\sigma^2 \sum_{j=0}^{\infty} a_j a_{j+1}} = b_0 \quad \mathbb{P} \text{ a.s.}$$

Therefore, the IV estimator is strongly consistent

Consider now that we want to find the asymptotic distribution of the proposed IV estimator.

$$\sqrt{T}(b_{IV}-b_0) = \left(\frac{1}{T}\sum_{t=1}^{T} x_{t-1}x_t\right)^{-1} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{t-1}\epsilon_t$$

We have already specified the limit of $\frac{1}{T}\sum_{t=1}^{T} x_{t-1}x_t$, therefore the limit distribution is going to be determined by $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{t-1}\epsilon_t$.
We are examing the properties of $(x_{t-1}\epsilon_t)_{t\in\mathbb{Z}}$ process and we have already said that is stationary ergodic, thus we need to investigate whether it is a square martingale difference sequence (s.m.d.) or not in order to use the appropriate Central Limit Theorem.

Remember that:
The process $(x_t)_{t\in\mathbb{Z}}$ is called a square martingale difference sequence wrt. the filtration $(f_t)_{t\in\mathbb{Z}}$, iff
1. $(x_t)_{t\in\mathbb{Z}}$ is adopted to $(f_t)_{t\in\mathbb{Z}}$
2. $E(x_t^2) < +\infty \quad \forall\, t\in\mathbb{Z}$
3. $E(x_t / f_{t-1}) = 0 \quad \forall\, t\in\mathbb{Z}$

Strict stationarity, ergodicity and the s.m.d. property imply the following CLT:
Suppose that $(x_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic, and that for some $(f_t)_{t\in\mathbb{Z}}$, $\left((x_t)_{t\in\mathbb{Z}}, (f_t)_{t\in\mathbb{Z}}\right)$ is s.m.d.
Then, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_t \xrightarrow{d} z \sim N\left(0, E(x_0^2)\right)$, as $T\to\infty$.

In our case consider the filtration $f_t := 6(\epsilon_{t-i}, i\geq 0)$ and further assume that $(\epsilon_t)_{t\in\mathbb{Z}}$ is s.m.d wrt. some filtration
- $(x_{t-1}\epsilon_t)_{t\in\mathbb{Z}}$ is adopted to f_t by construction
- $E(x_{t-1}\epsilon_t / f_{t-1}) = x_{t-1}\, E(\epsilon_t / f_{t-1}) = x_{t-1}\cdot 0 = 0 \quad \forall\, t\in\mathbb{Z}$
  since $(\epsilon_t)_{t\in\mathbb{Z}}$ is s.m.d.
- $E\left[(x_{t-1}\epsilon_t)^2\right] = E(x_{t-1}^2\cdot\epsilon_t^2) = E\left[\left(\sum_{j=0}^{\infty} a_j\epsilon_{t-1-j}\right)^2\epsilon_t^2\right]$

$$= E\left[\left(\sum_{j=0}^{\infty} a_j^2\epsilon_{t-1-j}^2 + \sum_{\substack{i,j=0\\ i\neq j}}^{\infty} a_i a_j \epsilon_{t-1-i}\epsilon_{t-1-j}\right)\epsilon_t^2\right]$$

$$= \sum_{j=0}^{\infty} a_j^2 E(\epsilon_t^2 \epsilon_{t-1-j}^2) + \sum_{\substack{i,j=0 \\ i \neq j}}^{\infty} a_i a_j E(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i})$$

$$\leq \sum_{j=0}^{\infty} a_j^2 E(\epsilon_t^2 \epsilon_{t-1-j}^2) + \left| \sum_{\substack{i,j=0 \\ i \neq j}}^{\infty} a_i a_j E(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i}) \right|$$

However, no already stated assumptions specify that the above sums converge. Thereby, we need to assume that $\sum_{j=0}^{\infty} a_j^2 E(\epsilon_t^2 \epsilon_{t-1-j}^2) < +\infty$ and $\left| \sum_{\substack{i,j=0 \\ i \neq j}}^{\infty} a_i a_j E(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i}) \right| < +\infty$.

Hence, the process $(x_{t-1} \epsilon_t)_{t \in \mathbb{Z}}$ is smd and we can use the relevant CLT.

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_{t-1} \epsilon_t \xrightarrow{d} z \sim N\left(0, E(x_{-1}^2 \epsilon_0^2)\right) \quad \text{as } T \to \infty.$$

and due to the CMT

$$\sqrt{T}(b_{1V} - b_0) = \left(\frac{1}{T} \sum_{t=1}^{T} x_{t-1} x_t\right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_{t-1} \epsilon_t \xrightarrow{d} \left[E(x_{-1} x_0)\right]^{-1} z \sim N\left(0, \frac{E(x_{-1}^2 \epsilon_0^2)}{\left[E(x_{-1} x_0)\right]^2}\right)$$