In our discussion today, I think we missed to consider that the integral over one of the variables of the joint distribution gives us the marginal.

$$\int f(Y, X)dX = f(Y)$$

which is the same as

$$\int f(Y|X)f(X)dX = f(Y)$$

With a conditioning variable,

$$\int f(Y, X|Z)dX = f(Y|Z)$$

$$\int f(Y|X, Z)f(X|Z)dX = f(Y|Z)$$

And if X is independet of Z but not of Y

$$\int f(Y|X, Z)f(X)dX = f(Y|Z)$$

In our case then,

$$\int f(W_{t-1}, \varepsilon_{t-1}|k_{t-1})d\varepsilon_{t-1} = f(W_{t-1}|k_{t-1})$$

I don't know the joint but I know the marginal of ε_{t-1} and althoug it is not independent of W_{t-1} it is independent of k_{t-1}

$$\int f(W_{t-1}, \varepsilon_{t-1}|k_{t-1})d\varepsilon_{t-1} =$$

$$\int f(W_{t-1}|\varepsilon_{t-1}k_{t-1})f(\varepsilon_{t-1}|k_{t-1})d\varepsilon_{t-1} =$$

$$\int f(W_{t-1}|\varepsilon_{t-1}k_{t-1})f(\varepsilon_{t-1})d\varepsilon_{t-1} = f(W_{t-1}|k_{t-1})$$

Lastly, by the **interchangeability of integrals** I don't see why I cannot integrate ε_{t-1} first

$$
\begin{aligned}
&E[h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})|k_{t-1}] = \\
&\qquad \int\int h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1}) f(W_{t-1}, \varepsilon_{t-1}|k_{t-1}) dW_{t-1} d\varepsilon_{t-1} \\
&\qquad \int\int h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1}) f(W_{t-1}, \varepsilon_{t-1}|k_{t-1}) d\varepsilon_{t-1} dW_{t-1} \\
&\qquad \int\int h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1}) f(W_{t-1}|k_{t-1}, \varepsilon_{t-1}) f(\varepsilon_{it-1}|k_{t-1}) d\varepsilon_{t-1} dW_{t-1} \\
&\qquad \int\int h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1}) f(\varepsilon_{it-1}) f(W_{t-1}|k_{t-1}, \varepsilon_{t-1}) d\varepsilon_{t-1} dW_{t-1} \\
&\qquad E\left[\int h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1}) f(\varepsilon_{it-1}) d\varepsilon_{t-1}|k_{t-1}, \varepsilon_{t-1}\right] \\
&\qquad E[E_{\varepsilon_{t-1}}[h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})]|k_{t-1}, \varepsilon_{t-1}]
\end{aligned}
$$

This is showing that I can integrate $h(\cdot)$ over ε_{t-1} first and then use k_{t-1} as instrument and, in principle, I could use ε_{t-1} as instrument too.

For example, when the markov process is a second degree polynomial,$h(\xi) = \delta_0 + \delta_1\xi + \delta_2\xi^2$

$$
\begin{aligned}
&E[h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})|k_{t-1}] = \\
&\quad = E[E_{\varepsilon_{t-1}}[h(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})]|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = E[E_{\varepsilon_{t-1}}[\delta_0 + \delta_1(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1}) + \delta_2(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})^2]|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = E[\delta_0 + \delta_1 E_{\varepsilon_{t-1}}[(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})] + \delta_2 E_{\varepsilon_{t-1}}[(W_{t-1} - (1-\beta)\varepsilon_{t-1} - \alpha k_{t-1})^2]|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = E[\delta_0 + \delta_1\omega_{t-1} + \delta_2 E_{\varepsilon_{t-1}}[((\omega_{t-1} + (1-\beta)\varepsilon_{t-1}) - (1-\beta)\varepsilon_{t-1})^2]|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = E[\delta_0 + \delta_1\omega_{t-1} \\
&\qquad + \delta_2 E_{\varepsilon_{t-1}}[(\omega_{t-1} + (1-\beta)\varepsilon_{t-1})^2 + (1-\beta)^2\varepsilon_{t-1}^2 - 2(1-\beta)\varepsilon_{t-1}(\omega_{t-1} + (1-\beta)\varepsilon_{t-1})]|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = E[\delta_0 + \delta_1\omega_{t-1} + \delta_2\omega_{t-1}^2 \\
&\qquad + \delta_2 E_{\varepsilon_{t-1}}[\Big((1-\beta)^2\varepsilon_{t-1}^2 + 2(1-\beta)\varepsilon_{t-1}\omega_{t-1}\Big) \\
&\qquad + (1-\beta)^2\varepsilon_{t-1}^2 - 2(1-\beta)\varepsilon_{t-1}(\omega_{t-1} + (1-\beta)\varepsilon_{t-1})]|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = \delta_0 + \delta_1 E[\omega_{t-1}|k_{t-1}, \varepsilon_{t-1}] + \delta_2 E[\omega_{t-1}^2|k_{t-1}, \varepsilon_{t-1}] \\
&\quad = \delta_0 + \delta_1 E[\omega_{t-1}|k_{t-1}] + \delta_2 E[\omega_{t-1}^2|k_{t-1}]
\end{aligned}
$$

Hence, all the parameters of $h(\cdot)$ are identified, if α is identified.

And for a third-degree polynomial $h(\cdot) = \delta_0 + \delta_1\xi + \delta_2\xi^2 + \delta_3\xi^3$,

$$
\begin{aligned}
&E[h(W_{t-1}-(1-\beta)\varepsilon_{t-1}-\alpha k_{t-1})|k_{t-1}] = \\
&\qquad E[E_{\varepsilon_{t-1}}[h(W_{t-1}-(1-\beta)\varepsilon_{t-1}-\alpha k_{t-1})]|k_{t-1},\varepsilon_{t-1}] \\
&\quad =\delta_0+\delta_1 E[\omega_{t-1}|k_{t-1},\varepsilon_{t-1}]+\delta_2 E[\omega_{t-1}^2|k_{t-1},\varepsilon_{t-1}] \\
&\qquad +\delta_3 E[E_{\varepsilon_{t-1}}[(W_{t-1}-(1-\beta)\varepsilon_{t-1}-\alpha k_{t-1})^3]|k_{t-1},\varepsilon_{t-1}] \\
&\quad =\delta_0+\delta_1 E[\omega_{t-1}|k_{t-1},\varepsilon_{t-1}]+\delta_2 E[\omega_{t-1}^2|k_{t-1},\varepsilon_{t-1}] \\
&\qquad +\delta_3 E[E_{\varepsilon_{t-1}}[((\omega_{t-1}+(1-\beta)\varepsilon_{t-1})-(1-\beta)\varepsilon_{t-1})^3]|k_{t-1},\varepsilon_{t-1}] \\
&\quad =\delta_0+\delta_1 E[\omega_{t-1}|k_{t-1},\varepsilon_{t-1}]+\delta_2 E[\omega_{t-1}^2|k_{t-1},\varepsilon_{t-1}] \\
&\qquad +\delta_3 E[E_{\varepsilon_{t-1}}[((\omega_{t-1}+(1-\beta)\varepsilon_{t-1})-(1-\beta)\varepsilon_{t-1})^2\,((\omega_{t-1}+(1-\beta)\varepsilon_{t-1})-(1-\beta)\varepsilon_{t-1})]|k_{t-1},\varepsilon_{t-1}] \\
&\quad =\delta_0+\delta_1 E[\omega_{t-1}|k_{t-1},\varepsilon_{t-1}]+\delta_2 E[\omega_{t-1}^2|k_{t-1},\varepsilon_{t-1}]+\delta_3 E[\omega_{t-1}^3|k_{t-1},\varepsilon_{t-1}] \\
&\qquad +\delta_3 E[E_{\varepsilon_{t-1}}[\left(\omega_{t-1}^2(1-\beta)\varepsilon_{t-1}\right)-\omega_{t-1}^2(1-\beta)\varepsilon_{t-1}]|k_{t-1},\varepsilon_{t-1}] \\
&\quad =\delta_0+\delta_1 E[\omega_{t-1}|k_{t-1},\varepsilon_{t-1}]+\delta_2 E[\omega_{t-1}^2|k_{t-1},\varepsilon_{t-1}]+\delta_3 E[\omega_{t-1}^3|k_{t-1},\varepsilon_{t-1}] \\
&\quad =\delta_0+\delta_1 E[\omega_{t-1}|k_{t-1}]+\delta_2 E[\omega_{t-1}^2|k_{t-1}]+\delta_3 E[\omega_{t-1}^3|k_{t-1}]
\end{aligned}
$$

Again, all the parameters of $h(\cdot)$ are identified, if α is identified.