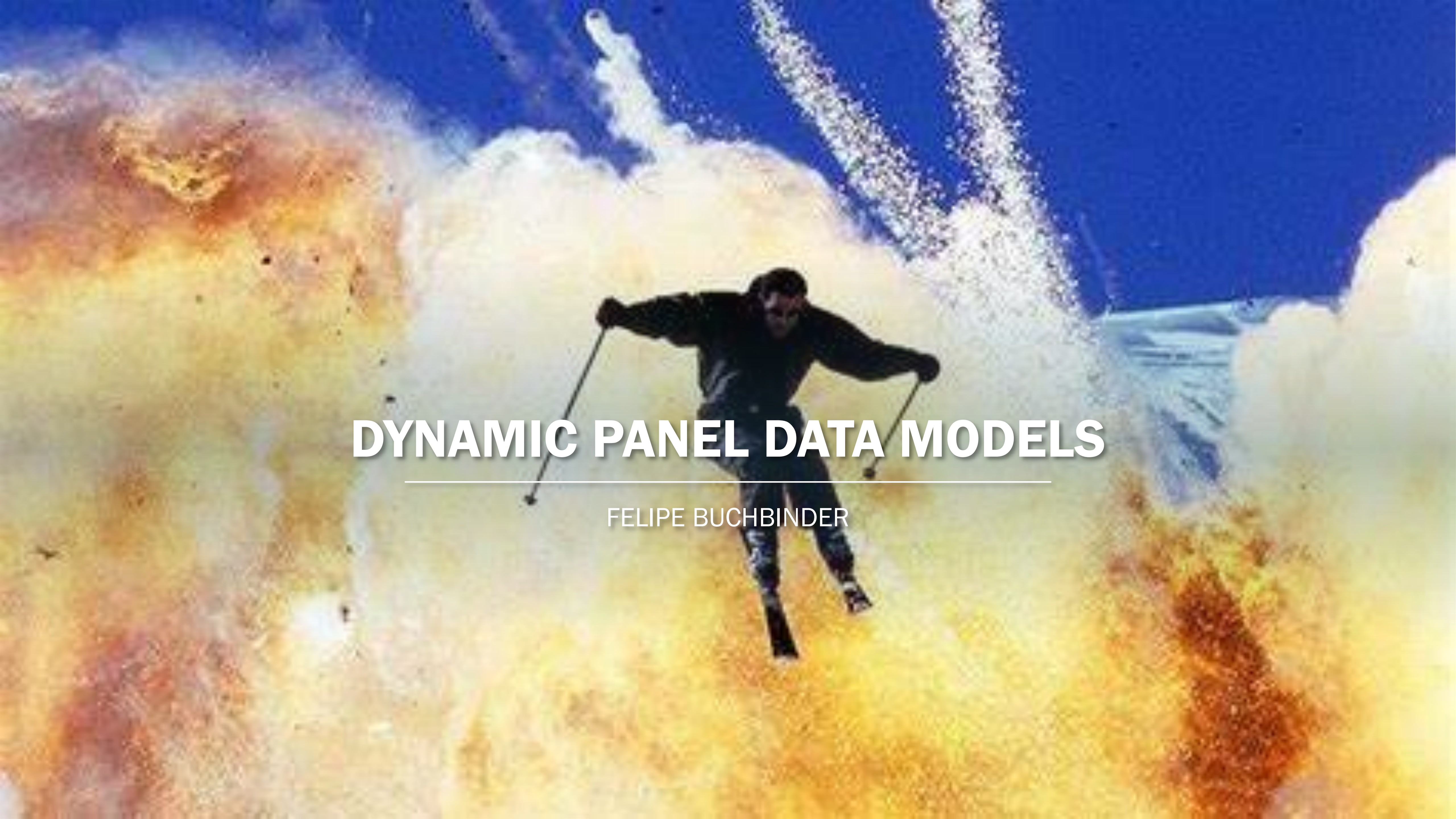


DYNAMIC PANEL DATA MODELS

FELIPE BUCHBINDER

THIS CLASS IN 2 SENTENCES

- A dynamic panel data model is one where y_{it} depends on passed values of $y_{i.}$, such as y_{it-1}, y_{it-2} etc.
- When this happens, $\boldsymbol{\beta}$ can be estimated using the Arellano-Bond (AB) GMM estimator.

DYNAMIC PANEL MODELS

(GENERAL EQUATION)

$$Y_{it} = \sum_{\tau=1}^{p} \alpha_\tau Y_{i,t-\tau} + \mathbf{X_{it}\beta} + \mathbf{U_i} + \boldsymbol{\epsilon_{it}}$$

DYNAMIC PANEL MODELS

(SIMPLEST CASE POSSIBLE: 1 LAG, NO COVARIATES)

$$Y_{it} = \sum_{\tau=1}^{p} \alpha_\tau Y_{i,t-\tau} + \mathbf{X_{it}\beta} + \mathbf{U_i} + \boldsymbol{\epsilon_{it}}$$

$$Y_{it} = \alpha Y_{i,t-1} + \mathbf{U_i} + \boldsymbol{\epsilon_{it}}$$

LET'S TRY FIXED EFFECTS...

- $\ddot{Y}_{it} = \alpha\ddot{Y}_{i,t-1} + \ddot{\epsilon}_{it}$

OLS ASSUMPTIONS ARE VIOLATED BECAUSE $\ddot{Y}_{i,t-1}$ AND $\ddot{\epsilon}_{i,t}$ ARE CORRELATED

- Y_{it-1} correlates with ϵ_{it-1} thus with $\overline{\epsilon_i}$ and finally with $\epsilon_{it} - \bar{\epsilon}_i$, This means $Y_{it-1} - \bar{Y}_i$ correlates with $\epsilon_{it} - \bar{\epsilon}_i$, and this violates the assumptions of linear regression.

PROOF

$$\operatorname{Cov}\begin{pmatrix}\ddot{Y}_{i,t-1}\\ \ddot{\epsilon}_{i,t}\end{pmatrix} = \operatorname{Cov}\begin{pmatrix}\ddot{Y}_{i,t-1}\\ \ddot{Y}_{i,t} - \alpha\ddot{Y}_{i,t-1}\end{pmatrix} = \underbrace{\operatorname{Cov}\begin{pmatrix}\ddot{Y}_{i,t-1}\\ \ddot{Y}_{i,t}\end{pmatrix}}_{\neq 0} + \alpha\,\underbrace{\operatorname{Cov}\begin{pmatrix}\ddot{Y}_{i,t-1}\\ \ddot{Y}_{i,t-1}\end{pmatrix}}_{\neq 0} \neq 0$$

$$\ddot{Y}_{it} = \alpha\ddot{Y}_{i,t-1} + \ddot{\epsilon}_{it}$$

PERHAPS YOU HAVE BETTER LUCK WITH FIRST DIFFERENCES, MR. BOND?

$$\Delta_t Y_{it} = \alpha \Delta_t Y_{it-1} + \Delta_t \epsilon_{it}$$

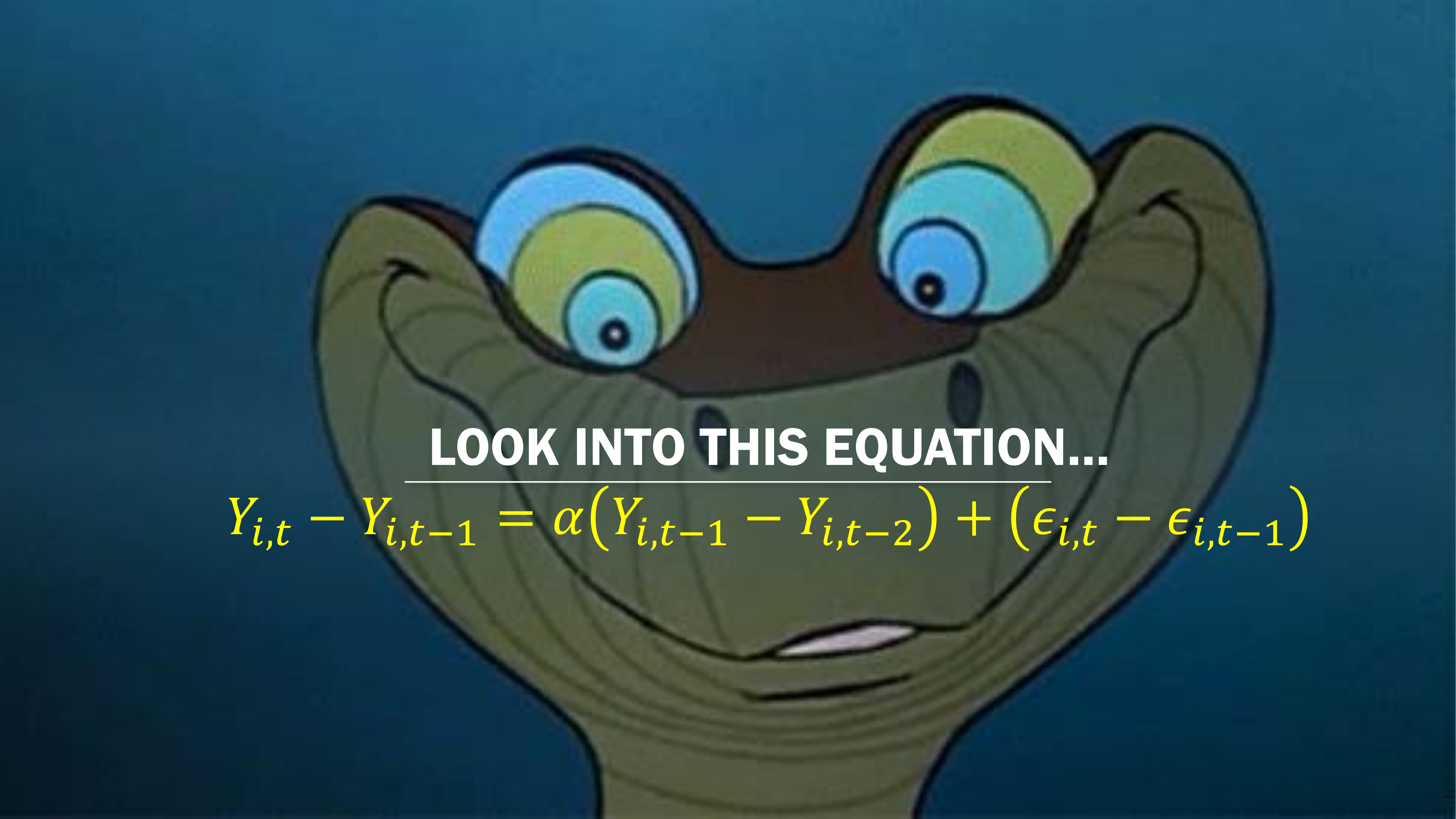
OLS ASSUMPTIONS ARE VIOLATED BECAUSE $\Delta_t Y_{i,t-1}$ AND $\Delta_t \epsilon_{i,t}$ ARE CORRELATED

- $\Delta_t Y_{it-1}$ correlates with Y_{it-1} thus with ϵ_{it-1} and finally with $\Delta_t \epsilon_{it}$.

PROOF

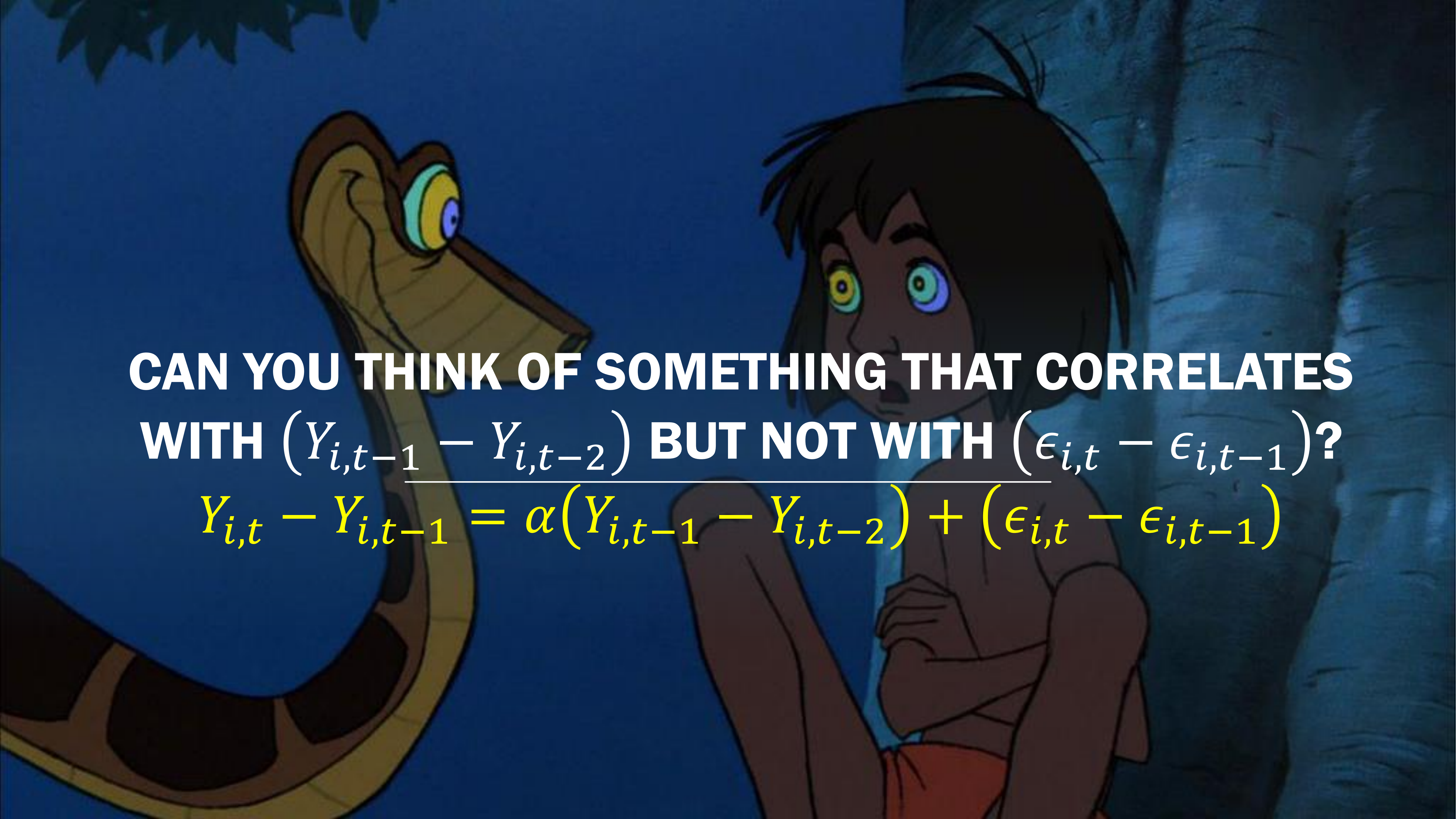
$$\text{Cov}\begin{pmatrix}\Delta Y_{i,t-1}\\ \Delta\epsilon_{i,t}\end{pmatrix} = \text{Cov}\begin{pmatrix}Y_{i,t-1} - Y_{i,t-2}\\ \epsilon_{i,t} - \epsilon_{i,t-1}\end{pmatrix} = \underbrace{\text{Cov}\begin{pmatrix}Y_{i,t-1}\\ \epsilon_{i,t-1}\end{pmatrix}}_{\neq 0} + \text{termos iguais a zero} \neq 0$$

HOW ABOUT WE FIND AN INSTRUMENTAL VARIABLE FOR THIS PROBLEM?



LOOK INTO THIS EQUATION...

$$Y_{i,t} - Y_{i,t-1} = \alpha\left(Y_{i,t-1} - Y_{i,t-2}\right) + \left(\epsilon_{i,t} - \epsilon_{i,t-1}\right)$$



CAN YOU THINK OF SOMETHING THAT CORRELATES WITH $\left(Y_{i,t-1} - Y_{i,t-2}\right)$ BUT NOT WITH $\left(\epsilon_{i,t} - \epsilon_{i,t-1}\right)$?

$$Y_{i,t} - Y_{i,t-1} = \alpha\left(Y_{i,t-1} - Y_{i,t-2}\right) + \left(\epsilon_{i,t} - \epsilon_{i,t-1}\right)$$

ANDERSON AND HSIAO (1981)
USE $Y_{i,t-2}$ AS AN INSTRUMENTAL VARIABLE FOR $\Delta_t Y_{i,t-1}$

1. Calculate a first-difference transform: $\Delta_t Y_{it} = \alpha \Delta_t Y_{it-1} + \Delta_t \epsilon_{it}$
2. Use $Y_{i,t-2}$ as an instrument for $\Delta_t Y_{i,t-1}$

ARELLANO AND BOND (1991)
USE MORE LAGS AS INSTRUMENTS

- In period $t = 3$, only $Y_{i,1}$ can be used as an instrument
- In period $t = 4$, we can use both $Y_{i,1}$ and $Y_{i,2}$ as instruments
- In period $t = 5$, we can use $Y_{i,1}$, $Y_{i,2}$ and $Y_{i,3}$ as instruments
- …

ARELLANO AND BOND (1991)
USE MORE LAGS AS INSTRUMENTS

- In period $t = 3$, only $Y_{i,1}$ can be used as an instrument
- In period $t = 4$, we can use both $Y_{i,1}$ and $Y_{i,2}$ as instruments
- In period $t = 5$, we can use $Y_{i,1}$, $Y_{i,2}$ and $Y_{i,3}$ as instruments
- …

In general, $Y_{i,\tau}$ can be used as an instrument for all $t \geq \tau - 2$, because it is correlated to $\Delta_t Y_{it}$ but uncorrelated to $\Delta_t \epsilon_{i,t}$

What would change if Y_{it} depended not only on $Y_{i,t-1}$ but also on $Y_{i,t-2}$?

MATHEMATICALLY,…

- The fact that $Y_{i,\tau}$ is correlated to $\Delta_t Y_{it}$ but uncorrelated to $\Delta_t \epsilon_{i,t}$ for all t $\geq$ 3 and $\tau \leq t-2$ can be expressed mathematically as …

$$\mathbb{E}\left(Y_{i,\tau}\Delta_t\epsilon_{i,t}\right) = 0 \; \forall \; t \geq 3 \wedge \tau \leq t-2$$

- This can be written in matrix notation as

$$\mathbb{E}\left(\mathbf{Z}^T\Delta_t\boldsymbol{\epsilon}_{\boldsymbol{i},\boldsymbol{t}}\right) = 0$$

Where $\mathbf{Z}$ is the matrix: $\begin{bmatrix} y_{i1} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & y_{i2} & y_{i1} & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$

- These equations can be solved using the Generalized Moments Method and yield unbiased and consistent estimates for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (in the general case where there are covariates present).

THIS CLASS IN 2 SENTENCES
(AGAIN ☺)

- A dynamic panel data model is one where y_{it} depends on passed values of $y_{i.}$, such as y_{it-1}, y_{it-2} etc.
- When this happens, $\boldsymbol{\beta}$ can be estimated using the Arellano-Bond (AB) GMM estimator.