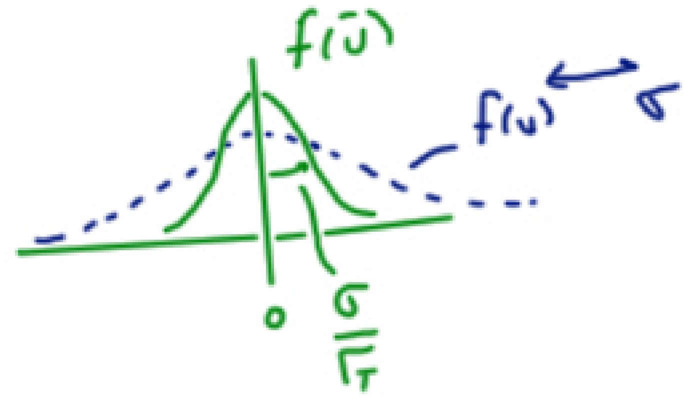
1 STANDARD ERROR OF THE MEAN

$$U_t : \begin{bmatrix} 5 \\ 4 \\ 8 \\ 8 \\ 2 \end{bmatrix} \quad E_T(U_t) = \frac{1}{T}\sum_{t=1}^{T} U_t = \bar{U} = 5.4$$

NEW SAMPLES

U_t DATA	5	3	5	8
	4	7	10	3
	8	7	4	6
	8	2	6	7
	2	2	3	9
$\bar{U} = E_T(U_t)$	5.4	4.2	5.6	6.6

$\bar{U}$ IS A RANDOM VARIABLE ACROSS SAMPLES (IMAGINARY)



$$E(\bar{U}) = E\left(\frac{1}{T}\sum_{t=1}^{T} U_t\right) = 0$$

$$\sigma^2(\bar{U}) = \sigma^2\left(\frac{1}{T}\sum_{t=1}^{T} U_t\right) = \frac{1}{T^2}\sigma^2\left(\sum_{t=1}^{T} U_t\right)$$

IF U_t STATIONARY + UNCORRELATED $\quad cov(U_t\, U_{t-j}) = 0$

$$\rightarrow \sigma^2(\bar{U}) = \frac{1}{T}\sigma^2(U) \quad ; \quad \sigma(\bar{U}) = \frac{\sigma(U)}{\sqrt{T}}$$

AMAZING: INFER $\sigma(\bar{U})$ OVER WORLDS NOT SEEN

① · IF $\{U_t\}$ STATIONARY

$$\sigma^2(\bar{U}) = \frac{1}{T}\sum_{j=-T}^{T}\frac{|T-j|}{T}\operatorname{cov}(U_t U_{t-j}) = \frac{1}{T}\left[\sigma^2(U) + 2\sum_{j=1}^{T}\frac{|T-j|}{T}\operatorname{cov}(U_t, U_{t-j})\right]$$

AS $T \to \infty$ $\quad \sigma^2(\bar{U}) \to \frac{1}{T}\sum_{j=-\infty}^{\infty}\operatorname{cov}(U_t, U_{t-j}) = \frac{1}{T}\left[\sigma^2(U) + 2\sum_{j=1}^{\infty}\operatorname{cov}(U_t, U_{t-j})\right] = \frac{S}{T}$

. $\sigma/\sqrt{T}$ FOR AUTOCORRELATED "FEWER DATA POINTS"

• IF $\quad U_t = \rho U_{t-1} + \varepsilon_t \qquad \sigma^2(\bar{U}) = \frac{1}{T}\sigma^2(U)\sum_{j=-\infty}^{\infty}\rho^{|j|} = \frac{1}{T}\sigma^2(U)\left(\frac{1+\rho}{1-\rho}\right)$

· "PARAMETRIC" CORRECTION

• GMM : MOST ECONOMETRICS BOILS DOWN TO THIS!

2. GMM – MOMENTS + ESTIMATE (NOTATION!)

EXPRESS

- MODEL → MEAN OF f(DATA, PARAMETER) = 0

VECTOR OF PARAMETERS

- ASSET PRICING $\quad 0 = E_t\left(M_{t+1}(b)\, R^e_{t+1}\right) \qquad 0 = E\left(M_{t+1}(b)\, R_{t+1} - 1\right)$

$$\rightarrow \quad 0 = E\left(M_{t+1}(b)\, R^e_{t+1} \cdot z_t\right)$$

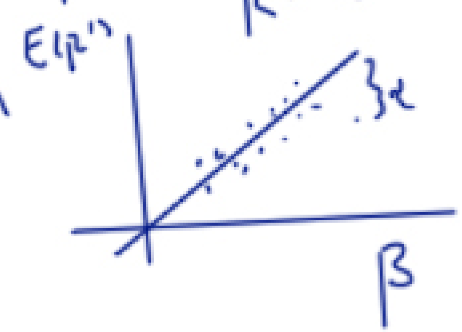
- CAPM $\quad 0 = E\left((a - b R^{em}_{t+1}) \cdot R^e_{t+1}\right) = 0 \qquad b = \{a, b\}$
- CONS. $\quad 0 = E\left(\beta \cdot \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \cdot R^e_{t+1}\right) = 0 \qquad b = \{\beta, \gamma\}$

TRUE VALUE (b_0 TO BE REALLY CLEAR)

- VECTOR $\quad R^e_{t+1} = \begin{bmatrix} R^{e1}_{t+1} \\ R^{e2}_{t+1} \\ \vdots \\ R^{e1}_{t+1} \cdot z_t \\ R^{e2}_{t+1} \cdot z_t \\ \vdots \end{bmatrix}$

- $E(MR^e) = E(M)E(R^e) + cov(M, R^e) = \frac{1}{R^f}\left[E(R^e) - \beta \cdot \lambda\right] = \frac{1}{R^f}\alpha$

"$E(MR^e)$ SHOULD BE 0" = "α SHOULD BE 0"



(2)

- "POPULATION". USE SAMPLE MOMENTS $E_T \equiv \frac{1}{T}\sum \ldots$
  - $E_T(M_{t+1}(b) R^e_{t+1}) = \frac{1}{T}\sum_{t=1}^{T} M_{t+1}(b) R^e_{t+1}$
- → "ESTIMATE" PICK $\hat{b} = b_T$ TO MAKE $E_T(MR^e) \approx 0$
  ESTIMATE $\neq b$
- → "STD ERROR" $\sigma(\hat{b})$
- → "TEST" IS $E(MR^e) \neq 0$ JUST LUCK? (AFTER ACCOUNTING FOR $\hat{b}$!)

"MODEL" "STRUC"

(VECTOR OF) MOMENTS $g(b) = E(f(\underset{\text{data}}{x_t}, b)) = E(u_t) = E(M_{t+1}(b) R^e_{t+1}) = 0$

SAMPLE MOMENTS $g_T(b) = E_T(f(x_t, b)) = E_T \ldots$

function of b

- ESTIMATION

2 PARAMETERS. N (25) MOMENTS? → SET 2 LIN. COMB. OF MOMENTS TO 0

$a_T \, g_T(\hat{b}) = 0$

YOU CHOOSE

EQUATION DEFINES $\hat{b}$

Ex.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{a_T} \underbrace{\begin{bmatrix} E_T(m(\hat{a},\hat{b}) R^{em}) \\ E_T(m(\hat{a},\hat{b}) R^{f}) - 1 \\ E_T(m(\hat{a},\hat{b}) R^{SMB}) \\ E_T(m(\hat{a},\hat{b}) R^{HML}) \end{bmatrix}}_{g_T(\hat{b})} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

≈ 0

② · ESTIMATION

· SOMETIMES YOU CAN SOLVE (LINEAR)

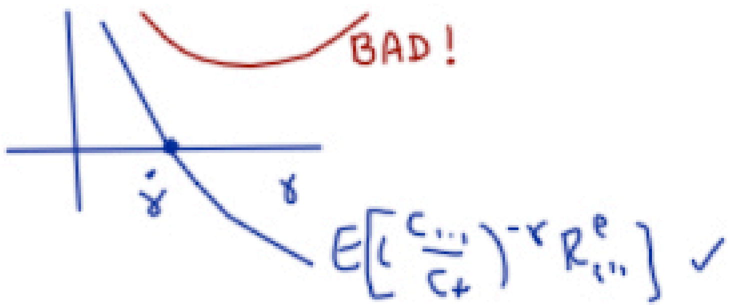
$$m = a - bR^{m}$$

$$\begin{bmatrix} a E_T(R^{em}) - b E_T(R^{em}R^{m}) \\ a E_T(R^f) - b E_T(R^m R^f) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

· SOMETIMES COMPUTER SEARCH — "NONLINEAR INSTRUMENTAL VARIABLES"

$$\begin{bmatrix} E_T\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^{em}_{t+1}\right) \\ E_T\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^{f}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

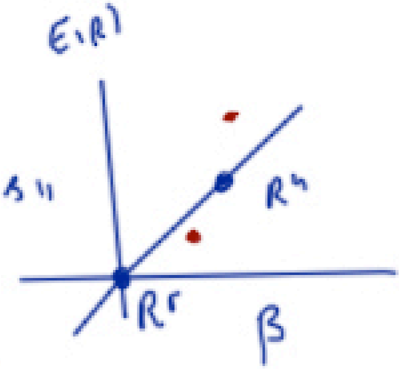
· NOTE:



· a_T IS FREE TO CHOOSE! [RANK 2]

HERE "ESTIMATE ON R^m, R^f, TEST ON R^{m-1}, R^{m5}"

OTHER STUFF! FLEXIBLE TOOL!



· EXACT ID (2b, 2 MOMENTS) IS OK. CAN'T "TEST."

3. GMM DISTRIBUTION

- PRODUCE $\hat{b}$ by $a_T g_T(\hat{b})=0$. DIST. OF $\hat{b}$? DIST OF $g_T(\hat{b})$ (=0?)

- IF MODEL IS TRUE $[g(b)=0]$, x_t STATIONARY $\Rightarrow$ $\hat{b}$ IS CONSISTENT, ASY. NORMAL

$$\sqrt{T}(\hat{b}-b) \to N\Big(0, \overbrace{(ad)^{-1} a S a' (ad)^{-1\prime}}^{VAR(\hat{b});\ STD\ ERRORS.}\Big)$$

$$d = \frac{\partial g(b)}{\partial b'} \qquad a = \text{plim}\ a_T \qquad S = \sum_{j=-\infty}^{\infty} E\left(f(x_t, b) f(x_{t-j}, b)'\right) = \sum_{j=-\infty}^{\infty} E(u_t u_{t-j}')$$

and sample moments $g_T(\hat{b})$ ← AFTER $\hat{b}$ SETS SOME TOO!

$$\sqrt{T}\, g_T(\hat{b}) \sim N\Big(0, \underbrace{[I - d(ad)^{-1}a]\, S\, [\cdot]'}_{VAR(g_T(\hat{b}))}\Big)$$

USE: "IS α OF SMALL FIRM SIGNIFICANT"?

TESTS! χ^2 TEST "are ALL α COLLECTIVELY BIGGER THAN 0?"

$$g_T(\hat{b})\ VAR(g_T(\hat{b}))^{+}\ g_T(\hat{b}) \sim \chi^2_{\#\,MOMS - \#\,PARAMS}$$

"TEST OF OVERIDENTIFYING RESTRICTIONS"

③ INTUITION

• WHY S? $$g_T = \frac{1}{T}\sum u_t = \frac{1}{T}\sum \overbrace{m_{t+1}(b)\, R^e_{t+1}}^{u_t,\ E(u_t)=0}$$

$$VAR(g_T(b)) = VAR\left(\frac{1}{T}\sum_{t=1}^{T} u_t\right) \rightarrow \frac{1}{T}\sum_{j=-\infty}^{\infty} E(u_t u'_{t-j}) = \frac{1}{T} S!$$

$\frac{S}{T}$ IS SAMPLING VARIANCE OF SAMPLE MOMENTS AT TRUE b

(IN ASSET PRICING $E_t(M_{t+1}R_{t+1})=0 \rightarrow E(u_{t-j}u_t)=0$ UNDER NULL.
A) NOT ALWAYS B) ALLOW ALTERNATIVE C) LONG HORIZONS, SHORT Δt)

- STD ERRORS "Δ METHOD"

$$a\, g_T(\hat{b}) = 0$$

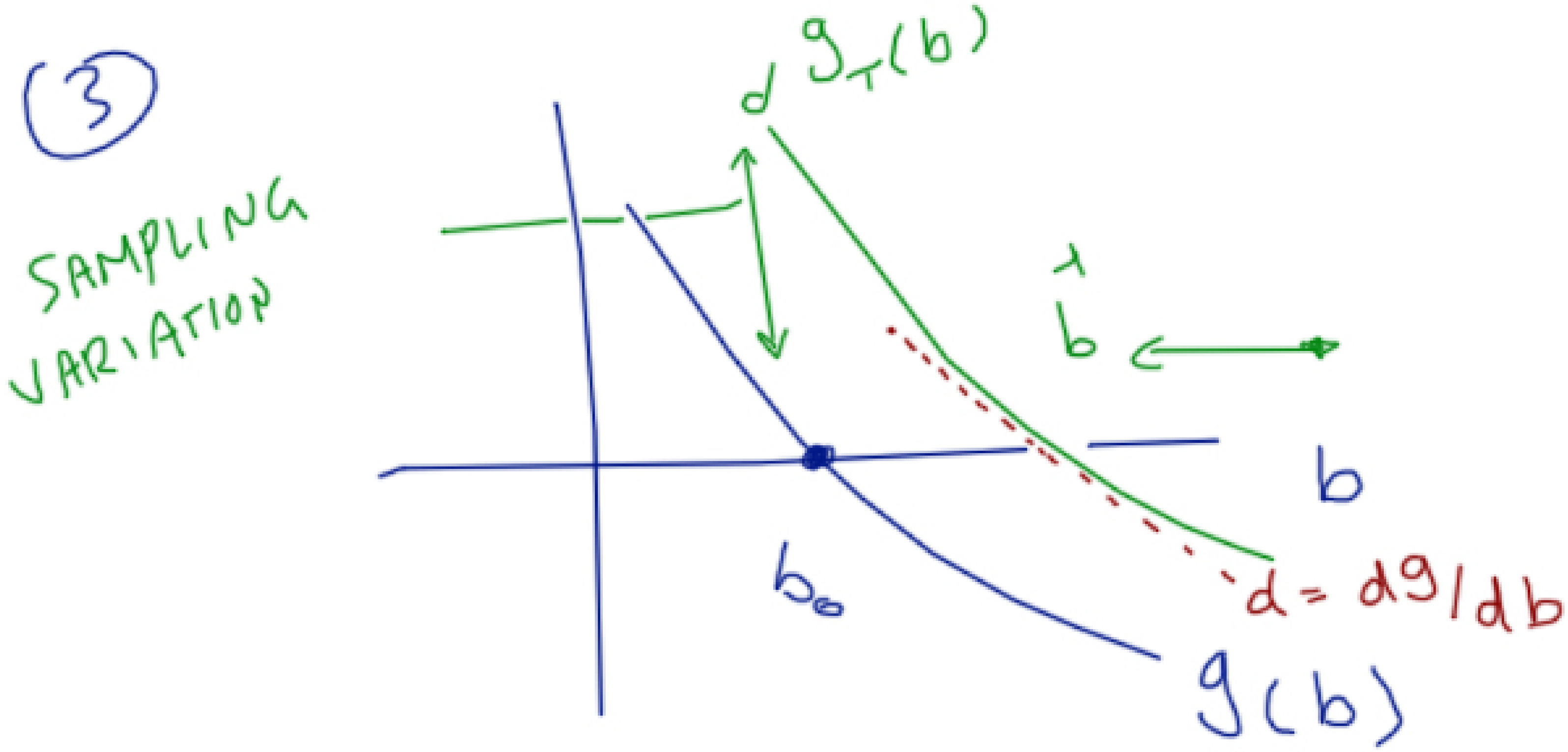
($g(b_0)=0$
$g_T(b_0) \neq 0$!) →

$$a\, g_T(b_0) + a \overbrace{\frac{\partial g_T}{\partial b'}}^{d} (\hat{b} - b_0) = 0$$

$$\hat{b} - b_0 = -(ad)^{-1} a\, g_T(b_0) \leftarrow S = VAR(g_T(b_0))$$

$$\sigma^2(\hat{b} - b_0) = \frac{1}{T}(ad)^{-1} a S a' (ad)^{-1'}$$

(3)



$\Rightarrow$ ESTIMATE ; STD ERROR ; TEST

$\hat{b}$; $\sigma(\hat{b})$; $g_T(\hat{b}) = 0$?

$g_T' \, VAR(g_T)^{-1} g_T \sim \chi^2$

$$(\hat{b} - b_0) = d^{-1}[g_T(b_0) - \underset{0}{\underbrace{g(b_0)}}]$$

- $S = cov(g_T(b_0))$ NEED $cov(g_T(\hat{b}))$

$$a\, g_T(\hat{b}) = 0$$

$$g_T(\hat{b}) \simeq g_T(b_0) + \overset{d}{\overbrace{\frac{\partial g_T}{\partial b'}}} \underbrace{(\hat{b} - b_0)}_{-(ad)^{-1} a\, g_T(b_0)}$$

$$g_T(\hat{b}) = [I - d(ad)^{-1}a]\, g_T(b_0)$$

$$VAR(g_T(\hat{b})) = [I - d(ad)^{-1}a]\, S\, [\cdot]'$$

- $a = I$? $[I - d d^{-1}] S [\cdot\cdot] = 0$ ... YES!

4. EFFICENT GMM. / WHAT a?

$$a\, g_T(\hat{b}) = 0$$

- WHAT a?
  - JUDGEMENT: "ECONOMICALLY IMPORTANT"
    "ROBUST" (MEASUREMENT, SPECIFICATION ERROR)
  - STATISTICS?
- STATISTICAL ANSWER "MOST EFFICIENT" = MIN $\sigma(\hat{b})$ GIVEN CHOICE OF g_T

  $a = d'S^{-1}$ IS MOST EFFICIENT

- FURMULAS SIMPLIFY

ONLY WHEN $a = d'S^{-1}$

$$\left\{\begin{array}{l} \mathrm{VAR}(\hat{b}) = \frac{1}{T}(d'S^{-1}d) \\ \mathrm{cov}(g_T(\hat{b})) = \frac{1}{T}[S - d(d'S^{-1}d)d'] \\ g_T'\,\mathrm{cov}^{-1}(g_T(\hat{b}))\,g_T = T g_T' S^{-1} g_T = T J_T \sim \chi^2_{\#M - \#PAR} \end{array}\right.$$

- USE ANY $a \rightarrow \hat{b}_1$ "FIRST STAGE" $\rightarrow S \rightarrow$ USE $a = d'S^{-1} \rightarrow \hat{b}_2$ "SECOND STAGE"
  - ITERATE OR SIMULTANEOUS ARE = ASYMPTOTICALLY
  - LIKE $\hat{b}_1 = OLS$, $\hat{b}_2 = GLS$

④

- MINIMIZATION APPROACH (WHAT a?)

$$\hat{b} = \min_{\{b\}} g_T(b)' W g_T(b)$$

W = "WEIGHING MATRIX"
AT LEAST RANK = # PARAM'S.

- COMPUTE MIN IS EASY

$$\left\{ \frac{\partial g_T(b)}{\partial b'} W \right\} g_T(\hat{b}) = 0$$

$$a_T \, g_T(\hat{b}) = 0 \leftarrow \text{A WAY TO PICK } a$$

FREE TO PICK W! $\quad$ MIN $\sum_{i=1} \alpha_i^2$; $\sum \frac{\alpha_i^2}{\sigma^2(\varepsilon_i)}$; $\alpha' \Sigma^{-1} \alpha$; $\alpha_1^2 + \alpha_2^2$

$$W = \begin{bmatrix} 1 & \\ & 100 \end{bmatrix}$$

- EFFICIENCY

$W = S^{-1}$ IS MOST EFFICIENT

- $$J_T = \min_{\{b\}} g_T(b)' S^{-1} g_T(b) \sim \chi^2_{\#M - \#PAR}$$

- WHY EFFICIENT ?

$$S = \begin{bmatrix} S_1^2 & 0 \\ 0 & S_2^2 \end{bmatrix} \rightarrow \frac{E(MR^1)^2}{S_1^2}, \frac{E(MR^2)^2}{S_2^2}$$

PAY ATTENTION TO BEST MEASURED MOMENT = MINIMUM VARIANCE [OF MR] PORTFOLIO. {DANGER!}

5. GMM DOES OLS

$$y_t = x_t' \beta + \varepsilon_t \qquad x_t = \begin{bmatrix} 1 \\ x_t \\ x_t^2 \end{bmatrix}$$

MAP
$E(\cdot)=0?$ $\rightarrow$

$$g_T(\hat\beta) = E_T[x_t(y_t - x_t'\hat\beta)] = 0$$

$$\hat\beta = E_T(x_t x_t')^{-1} E_T(x_t y_t) \qquad = (X'X)^{-1}X'Y$$

$$d = \frac{\partial g_T}{\partial b'} = -E(x_t x_t') \; ; \; a = I$$

$$f(x_t, b) = u_t(b) = x_t(y_t - x_t'\beta) = x_t \varepsilon_t$$

$$S = \sum_{j=-\infty}^{\infty} E(u_t u_{t-j}') = \sum_{j=-\infty}^{\infty} E(\varepsilon_t x_t x_{t-j}' \varepsilon_{t-j})$$

$$\boxed{\sigma^2(\hat\beta) = \frac{1}{T} d^{-1} S d^{-1\prime} = \frac{1}{T} E(x_t x_t')^{-1} \sum_{j=-\infty}^{\infty} E(\varepsilon_t x_t x_{t-j}' \varepsilon_{t-j}) E(x_t x_t')^{-1}}$$

(5) HELLO? $\sigma^2(\hat{\beta}) = (X'X)^{-1}s^2$??

A) ε_t UNCORRELATED OVER TIME $E(\varepsilon_t \mid x_t, x_{t-1} \ldots \varepsilon_{t-1}, \varepsilon_{t-2} \ldots) = 0$

$\rightarrow$ no LAGS

B ε_t NOT CONDITIONALLY HETEROSKEDASTIIC $E(\varepsilon_t^2 \mid x_t, x_{t-1}, \varepsilon_{t-1} \ldots) = \sigma_\varepsilon^2$

$\rightarrow E(\varepsilon_t^2 x_t x_t') = E(E(\varepsilon_t^2 \mid x_t x_t') \cdot x_t x_t') = \sigma_\varepsilon^2 E(x_t x_t')$

$\rightarrow \sigma^2(\hat{\beta}) = \frac{1}{T} E(x_t x_t')^{-1} \sigma^2(\varepsilon)$!

NO B? "WHITE ERRORS" $\sigma^2(\hat{\beta}) = \frac{1}{T} E(x_t x_t')^{-1} E(\varepsilon_t^2 x_t x_t') E(x_t x_t')^{-1}$

NO A, B? HANSEN-HODRICK / NEWEY WEST ERROR !

NOTE ESTIMATE $\sum_{j=-k}^{k} w_j E(\varepsilon_t x_t x_{t-j}' \varepsilon_{t-j})$ OR PARAMETRIC.

6. CHOOSING a / W, EFFICIENT VS ROBUST

- WHAT a / WHAT W

$$a\, g_T(\hat{b}) = 0 \qquad \min_{\{b\}} g_T(b)\, W\, g_T(b)$$

$$\underbrace{d' W}_{a}\, g_T(\hat{b}) = 0$$

- "ROBUST" "IMPORTANT". BIG CHANGE IN PHILOSOPHY (FAMA-FRENCH)

  MEASURE, EVALUATE, .. TEST?

- ECONOMICS

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} E_T(w_L) - \text{MODEL} \\ E_T(\Delta y) - \text{MODEL} \\ \sigma(\Delta)\sigma(\Delta i) \\ \text{CORR}(\Delta c, \Delta i) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

⟸ "CALIBRATE"

⟸ "EVALUATE"

- THOUGH MODEL EARLY "REJECTED"
- BUT DON'T GIVE UP ON $\sigma(\hat{\varepsilon})$!

(b)

- FINANCE. M IS SCALAR. $g_T(b) = E_T(m(b) \cdot R)$: $\quad a\, g_T(b) = E_T(m(b) \cdot a \underset{\rightarrow}{R})$

a = WHICH PORTFOLIOS PRICED PERFECTLY?

$$\min \; g_T' \underset{Q'Q}{W} g_T = \min \; E_T(m(b)\, [\underbrace{Q} R]')\, E_T[(m(b)\, [QR])$$

PORTFOLIOS

- 1^{st} STAGE ("INTERESTING" a) VS EFFICIENT?
  - 1^{st} IS CONSISTENT, CORRECT $\sigma(\hat{b})$, $\sigma(g_T(\hat{b}))$. NOT "WRONG"
  - 1^{st} "INEFFICIENT" GIVEN MOMENTS. MATTERS!

- EFFICIENT: DANGEROUS!
  - OLS VS GLS $\quad$ OLS $\quad \min\, (Y - X\beta)'(Y - X\beta) \rightarrow \beta = (X'X)^{-1} X'Y$
  - Σ AC, HETEROSKEDASTIC? - OLS CONSISTENT; UNBIASED; NOT "EFFICIENT"
    - $\sigma(\hat{\beta}) = s^2 (X'X)^{-1}$ IS WRONG
  - STATISTICIAN: $\beta_{OLS} \rightarrow E(\varepsilon\varepsilon') = \Omega$; GLS $\min\, (Y - X\beta)' \Omega^{-1} (Y - X\beta)$

$$\beta = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1} Y \qquad \sigma^2(\hat{\beta}) = (X'\Omega^{-1}X)^{-1} s^2$$

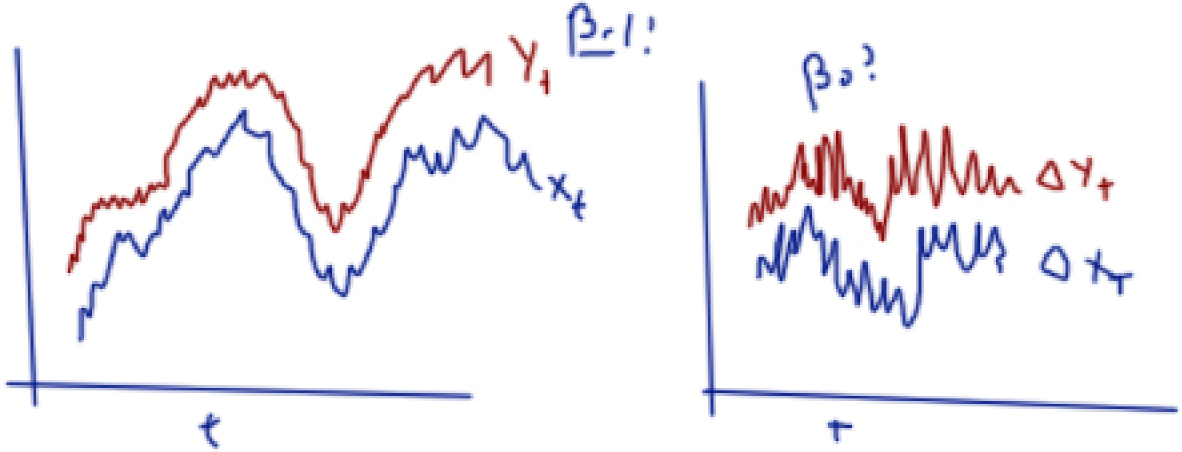
CONSISTENT (CORRECT $\sigma(\hat{\beta})$)
EFFICIENT

(5) · ECONOMIST DANGER

$$Y_t = X_t \beta + U_t$$

$$U_t = U_{t-1} + \varepsilon_t$$

$\rightarrow$ GLS $\quad Y_t - Y_{t-1} = (X_t - X_{t-1})\beta + \varepsilon$



• STATISTICIAN: YOU DON'T TELL ME THE MODEL WASN'T 100% RIGHT

ANSWER: RUN OLS, USE RIGHT SE

$$\sigma(\hat{\beta}_{OLS}) - VAR((X'X)^{-1}X'Y) = VAR((X'X)^{-1}X'\varepsilon) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

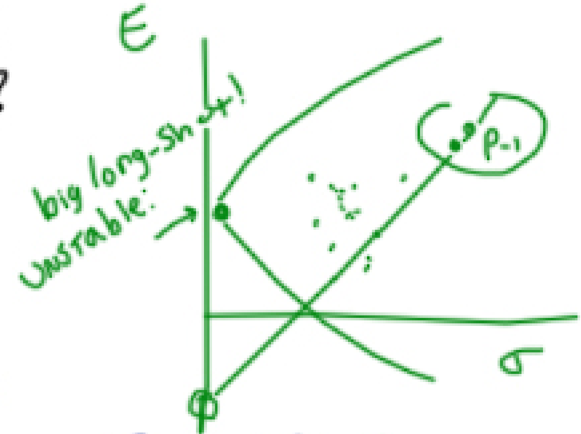
– THIS IS 1ST STAGE GMM.

7 S DANGER

MIN $g_T' S^{-1} g_T$ DANGER

? $a\, g_T(\hat{b}) = 0$

min $g_T' W g_T$?



$S = Q \Lambda Q'$; min $g_T' S^{-1} g_T$

$$\min \ (Q' g_T)' \underbrace{\Lambda^{-1}}_{\text{VAR}} \underbrace{(Q' g_T)}_{\text{PORTFOLIOS}}$$

$$S = \begin{bmatrix} s_1^2 & 0 \\ 0 & s_2^2 \end{bmatrix} \quad - \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \min \frac{g_1^2}{s_1^2} + \frac{g_1^2}{s_2^2} \ ✓$$

PAY ATTENTION TO WELL MEASURED $\sigma^2(g_T) = \sigma^2(E_T(MR)) \approx \sigma^2(R)/T$

$$S = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1+\rho & \\ & 1-\rho \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\rightarrow \min \frac{[E(m \cdot (R^{e1} + R^{e2}))]^2}{1+\rho} + \frac{[E(m(R^{e1} - R^{e2}))]^2}{1-\rho}$$

ASP →) PAY ALL ATTENTION TO SAMPLE MINIMUM VARIANCE

$R^{e2} - R^{e1}$ PORTFOLIO

⇒ LOOK AT S! BIG LONG-SHORT?

- BOTTOM LINE: "EFFICIENCY" USUALLY MORE ASSETS?
- GMM GREAT TOOL FOR MODERN STYLE
- BUT... BIG T SMALL N (CAREFULLY PICKED PORTFOLIOS)

ISSUE: BIG N SMALL T, MUCH CROSS CORRELATION.