

GSERM - St. Gallen 2019

GLS-ARMA and Dynamics

June 17, 2019 (afternoon session)

For:

$$Y_{it} = \mathbf{X}_{it}\beta + u_{it}$$

i.i.d. u_{it} s require:

$$\begin{aligned} \mathbf{u}\mathbf{u}' \equiv \mathbf{\Omega} &= \sigma^2 \mathbf{I} \\ &= \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} \end{aligned}$$

That is, within units:

- $\text{Var}(u_{it}) = \text{Var}(u_{is}) \forall t \neq s$ (temporal homoscedasticity)
- $\text{Cov}(u_{it}, u_{is}) = 0 \forall t \neq s$ (no within-unit autocorrelation)

and between units:

- $\text{Var}(u_{it}) = \text{Var}(u_{jt}) \forall i \neq j$ (cross-unit homoscedasticity)
- $\text{Cov}(u_{it}, u_{jt}) = 0 \forall i \neq j$ (no between-unit / spatial correlation)

Estimator:

$$\hat{\beta}_{GLS} = (\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1}\mathbf{X}'\Omega^{-1}\mathbf{Y}$$

with:

$$\widehat{V(\beta_{GLS})} = (\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1}$$

Two approaches:

- Use OLS \hat{u}_{it} s to get $\hat{\Omega}$ (“feasible GLS”)
- Use substantive knowledge about the data to structure Ω

Assume:

- $E(u_{it}^2) = E(u_{is}^2) \forall t \neq s$
- $E(u_{it}, u_{jt}) = \sigma_{ij} \forall i \neq j,$
- $E(u_{it}, u_{js}) = 0 \forall i \neq j, t \neq s$
- $E(u_{it}, u_{is}) = \rho$ or ρ_i

(B&K: “panel error assumptions”).

Then

1. Use OLS to generate $\hat{u}s \rightarrow \hat{\rho} (\rightarrow \hat{\Omega}),$
2. Use $\hat{\rho}$ for Prais-Winsten.

This method was widely used prior to B&K (1995)

$$\Omega = \begin{pmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \Sigma & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma \end{pmatrix} = \Sigma \otimes \mathbf{I}_N$$

where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1N} & \sigma_{2N} & \cdots & \sigma_N^2 \end{pmatrix}$$

Means:

- $\frac{N(N-1)}{2}$ distinct contemporaneous correlations,
- NT observations,
- $\rightarrow 2T/(N+1)$ observations per $\hat{\sigma}$

Panel-Corrected Standard Errors

Key to PCSEs:

$$\hat{\sigma}_{ij} = \frac{\sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt}}{T}$$

Define:

$$\mathbf{U}_{T \times N} = \begin{pmatrix} \hat{u}_{11} & \hat{u}_{21} & \cdots & \hat{u}_{N1} \\ \hat{u}_{12} & \hat{u}_{22} & \cdots & \hat{u}_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{u}_{1T} & \hat{u}_{2T} & \cdots & \hat{u}_{NT} \end{pmatrix}$$

$$\hat{\Sigma} = \frac{(\mathbf{U}'\mathbf{U})}{T}$$

$$\hat{\Omega}_{PCSE} = \frac{(\mathbf{U}'\mathbf{U})}{T} \otimes \mathbf{I}_T$$

Panel-Corrected Standard Errors

Correct formula:

$$\text{Cov}(\hat{\beta}_{PCSE}) = (\mathbf{X}'\mathbf{X})^{-1}[\mathbf{X}'\mathbf{\Omega}\mathbf{X}](\mathbf{X}'\mathbf{X})^{-1}$$

PCSEs:

- Do not fix unit-level heterogeneity (a la “fixed” / “random” effects)
- Do not deal with dynamics
- Depend critically on the “panel data assumptions” of Park / B&K

Panel Assumptions and Numbers of Parameters to be Estimated

Panel Assumptions	No AR(1)	Common $\hat{\rho}$	Separate $\hat{\rho}_i$ s
$\sigma_i^2 = \sigma^2, \text{Cov}(\sigma_{it}, \sigma_{jt}) = 0$	$k + 1$	$k + 2$	$k + N + 1$
$\sigma_i^2 \neq \sigma^2, \text{Cov}(\sigma_{it}, \sigma_{jt}) = 0$	$k + N$	$k + N + 1$	$k + 2N$
$\sigma_i^2 \neq \sigma^2, \text{Cov}(\sigma_{it}, \sigma_{jt}) \neq 0$	$\frac{N(N-1)}{2} + k + N$	$\frac{N(N-1)}{2} + k + N + 1$	$\frac{N(N-1)}{2} + k + 2N$

Example: Central Banks, Unions, Unemployment

- Hall and Franzese (1998 *IO*)
- 18 OECD countries, 1955-1990 ($N = 18$, $T = 36$, $NT = 648$)
- Y = unemployment
- Covariates: GDP, openness, union density, left cabinets, central bank independence, coordinated wage bargaining, interaction

Example: Data

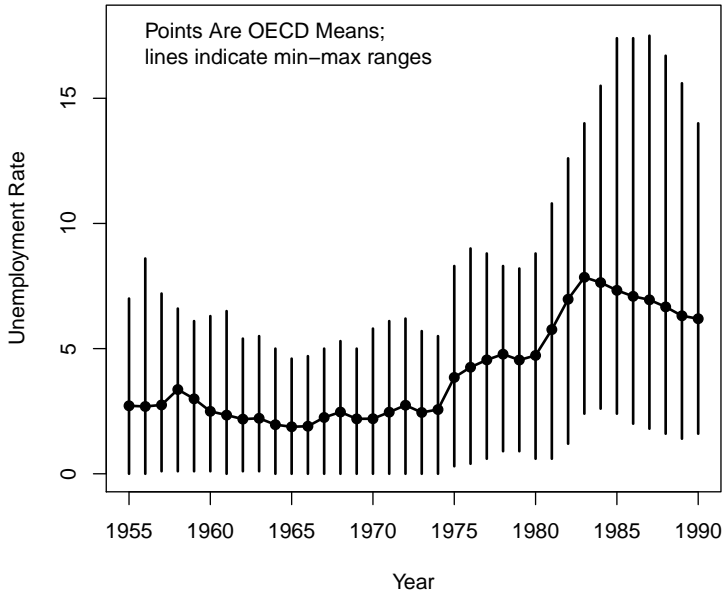
```
> summary(HF)
```

country	year	ue	inf	cbi
Min. : 1.0	Min. :1955	Min. : 0.0	Min. : -1.7	Min. :0.12
1st Qu.: 5.0	1st Qu.:1964	1st Qu.: 1.6	1st Qu.: 3.2	1st Qu.:0.41
Median : 9.5	Median :1972	Median : 3.0	Median : 4.9	Median :0.47
Mean :10.3	Mean :1972	Mean : 4.0	Mean : 6.0	Mean :0.50
3rd Qu.:15.0	3rd Qu.:1981	3rd Qu.: 5.7	3rd Qu.: 7.7	3rd Qu.:0.61
Max. :21.0	Max. :1990	Max. :17.5	Max. :27.2	Max. :0.93

cwagebrg	GDP_PC	open	uden	lcab
Min. :0.00	Min. :7.6	Min. :0.07	Min. :0.10	Min. :0.00
1st Qu.:0.25	1st Qu.:8.9	1st Qu.:0.31	1st Qu.:0.32	1st Qu.:0.00
Median :0.50	Median :9.2	Median :0.43	Median :0.41	Median :0.07
Mean :0.49	Mean :9.1	Mean :0.46	Mean :0.44	Mean :0.31
3rd Qu.:0.75	3rd Qu.:9.4	3rd Qu.:0.54	3rd Qu.:0.56	3rd Qu.:0.58
Max. :1.00	Max. :9.8	Max. :1.40	Max. :0.85	Max. :1.00

wagexcbi	HasLCAB
Min. :0.00	Min. :1
1st Qu.:0.04	1st Qu.:1
Median :0.21	Median :1
Mean :0.25	Mean :1
3rd Qu.:0.37	3rd Qu.:1
Max. :0.70	Max. :1

Unemployment in 18 Nations, 1955-1990



Example: OLS

```
> summary(HF.OLS)
Oneway (individual) effect Pooling Model

Balanced Panel: n=18, T=36, N=648

Residuals :
    Min. 1st Qu.  Median 3rd Qu.    Max.
-5.120  -1.500   -0.241    1.230    9.290

Coefficients :
              Estimate Std. Error t-value    Pr(>|t|)
(Intercept) -13.579      2.328   -5.83 0.0000000086 ***
GDP_PC       1.603      0.263    6.09 0.0000000020 ***
open         5.119      0.418   12.24 < 2e-16 ***
uden         0.709      0.808    0.88    0.38
lcab         0.236      0.293    0.81    0.42
cbi          5.169      1.097    4.71 0.0000030150 ***
cwagebrg     -1.292      0.792   -1.63    0.10
wagexcbi     -7.030      1.505   -4.67 0.0000036327 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Total Sum of Squares:    6500
Residual Sum of Squares: 3730
R-Squared:               0.426
Adj. R-Squared: 0.421
F-statistic: 67.9634 on 7 and 640 DF, p-value: <2e-16
```

Example: Prais-Winsten

```
> HF.prais <- prais.winsten(ue~GDP_PC+open+uden+lcab+cbi+cwagebrg+wagexcbi,  
                           data=HF,iter=100)
```

```
> HF.prais
```

Residuals:

	Min	1Q	Median	3Q	Max
	-8.456	-0.431	-0.144	0.314	4.615

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
Intercept	-15.2783	2.2128	-6.90	1.2e-11	***
GDP_PC	2.3415	0.2515	9.31	< 2e-16	***
open	-0.3491	0.7950	-0.44	0.6608	
uden	5.5466	1.1492	4.83	1.7e-06	***
lcab	-0.0593	0.1861	-0.32	0.7501	
cbi	-3.4801	2.4753	-1.41	0.1602	
cwagebrg	-10.5954	2.0019	-5.29	1.7e-07	***
wagexcbi	10.6805	3.4942	3.06	0.0023	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.95 on 640 degrees of freedom

Multiple R-squared: 0.279, Adjusted R-squared: 0.27

F-statistic: 31 on 8 and 640 DF, p-value: <2e-16

Rho	Rho.t.statistic	Iterations
0.94	73	10

Example: GLS with Homoscedastic AR(1) Errors

```
> HF.GLS <- gls(ue~GDPPC+open+uden+lcab+cbi+cwagebrg+wagexcbi,
               HF,correlation=corAR1(form=~1|country))
> summary(HF.GLS)
Generalized least squares fit by REML
Model: ue ~ GDPPC + open + uden + lcab + cbi + cwagebrg + wagexcbi
Data: HF
    AIC   BIC logLik
1484 1529   -732

Correlation Structure: AR(1)
Formula: ~1 | country
Parameter estimate(s):
Phi
0.99
```

Coefficients:

	Value	Std.Error	t-value	p-value
(Intercept)	43	7.3	5.8	0.000
GDPPC	-4	0.7	-5.5	0.000
open	-1	0.8	-1.8	0.072
uden	1	2.2	0.3	0.792
lcab	0	0.1	-0.7	0.473
cbi	-1	7.5	-0.2	0.848
cwagebrg	-5	6.4	-0.8	0.402
wagexcbi	3	11.7	0.3	0.770

Correlation:

	(Intr)	GDPPC	open	uden	lcab	cbi	cwgbrg
GDPPC	-0.827						
open	0.124	-0.207					
uden	-0.145	0.017	-0.048				
lcab	0.041	-0.054	0.005	-0.003			
cbi	-0.421	-0.100	0.033	0.069	0.009		
cwagebrg	-0.371	-0.065	0.018	-0.084	-0.014	0.721	
wagexcbi	0.334	0.082	-0.028	0.017	0.004	-0.813	-0.905

Standardized residuals:

	Min	Q1	Med	Q3	Max
	-1.49	-0.64	-0.20	0.46	2.55

More GLS: Unit-Wise Heteroscedasticity

```
> HF.GLS2 <- gls(ue~GDPPC+open+uden+lcab+cbi+cwagebrg+wagexcbi,  
+ HF,correlation=corAR1(form=~1|country),  
+ weights = varIdent(form = ~1|country))  
> summary(HF.GLS2)  
Generalized least squares fit by REML  
   AIC   BIC logLik  
1326 1446   -636
```

Correlation Structure: AR(1)

Formula: ~1 | country
Parameter estimate(s):
Phi
0.98

Variance function:

Structure: Different standard deviations per stratum
Formula: ~1 | country
Parameter estimates:

	1	2	3	4	5	6	7	8	9	10	11	13	14	15	18	19	20	21
	1.00	0.19	0.75	0.54	0.60	1.00	0.95	0.32	0.88	0.89	0.78	1.09	0.92	0.57	0.33	0.38	0.86	0.60

Coefficients:

	Value	Std.Error	t-value	p-value
(Intercept)	21.1	4.7	4.5	0.0000
GDPPC	-1.6	0.4	-4.3	0.0000
open	-2.2	0.6	-3.4	0.0008
uden	0.9	1.5	0.6	0.5415
lcab	-0.1	0.1	-1.2	0.2206
cbi	-1.9	7.3	-0.3	0.7984
cwagebrg	-6.3	4.7	-1.3	0.1794
wagexcbi	5.0	9.5	0.5	0.5996
.				
.				
.				

Example: PCSEs

```
> library(lmtest)
> coeftest(HF.OLS,vcov=vcovBK)
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-13.579	7.320	-1.86	0.064	.
GDP_PC	1.603	0.821	1.95	0.051	.
open	5.119	1.304	3.93	0.000096	***
uden	0.709	2.518	0.28	0.778	
lcab	0.236	0.668	0.35	0.724	
cbi	5.169	3.439	1.50	0.133	
cwagebrg	-1.292	2.478	-0.52	0.602	
wagexcbi	-7.030	4.726	-1.49	0.137	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Alternative Approach: pcse

```
> HF.lm<-lm(ue~GDP_PC+open+uden+lcab+cbi+cwagebrg+wagexcbi,data=HF)
> HF.pcse<-pcse(HF.lm,groupN = HF$country, groupT = HF$year)
> summary(HF.pcse)
```

Results:

	Estimate	PCSE	t value	Pr(> t)
(Intercept)	-13.58	4.74	-2.87	4.3e-03
GDP_PC	1.60	0.53	3.01	2.7e-03
open	5.12	0.53	9.71	7.0e-21
uden	0.71	0.53	1.35	1.8e-01
lcab	0.24	0.27	0.88	3.8e-01
cbi	5.17	0.85	6.10	1.9e-09
cwagebrg	-1.29	0.77	-1.67	9.6e-02
wagexcbi	-7.03	1.05	-6.68	5.3e-11

Valid Obs = 648; # Missing Obs = 0; Degrees of Freedom = 640.

General advice...

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\beta_{LDV} + \epsilon_{it}$$

If ϵ_{it} is perfect...

- $\hat{\beta}_{LDV}$ is biased (but consistent),
- $O(\text{bias}) = \frac{-1+3\beta_{LDV}}{T}$

If ϵ_{it} is autocorrelated...

- $\hat{\beta}_{LDV}$ is biased and inconsistent
- IV is one (bad) option...

Lagged Y s and GLS-ARMA

Can rewrite:

$$\begin{aligned}Y_{it} &= \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + u_{it} \\ u_{it} &= \phi u_{it-1} + \eta_{it}\end{aligned}$$

as

$$\begin{aligned}Y_{it} &= \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + \phi u_{it-1} + \eta_{it} \\ &= \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + \phi(Y_{it-1} - \mathbf{X}_{it-1}\boldsymbol{\beta}_{AR}) + \eta_{it} \\ &= \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + \mathbf{X}_{it-1}\psi + \eta_{it}\end{aligned}$$

where $\psi = \phi\boldsymbol{\beta}_{AR}$ and $\psi = 0$ (by assumption).

Lagged Y s and World Domination

In:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\beta_{LDV} + \epsilon_{it}$$

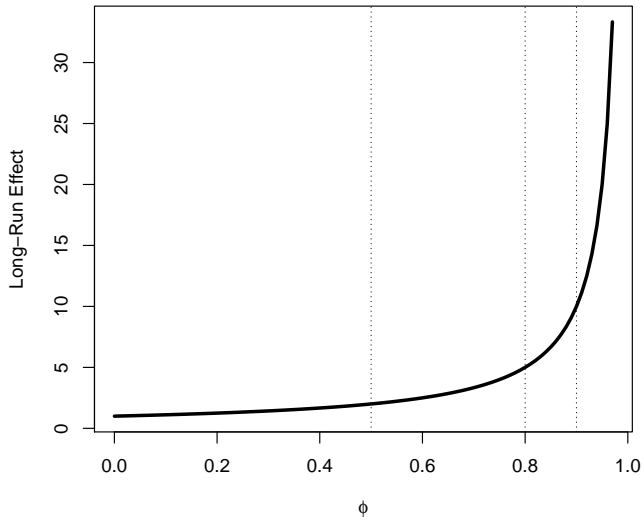
Achen: Bias “deflates” $\hat{\beta}$ relative to $\hat{\phi}$, “suppress” the effects of \mathbf{X} ...

Keele & Kelly (2006):

- Contingent on ϵ s having autocorrelation
- Key: In LDV, *long-run impact of a unit change in X is:*

$$\hat{\beta}_{LR} = \frac{\hat{\beta}_{LDV}}{1 - \hat{\phi}}$$

Long-Run Impact for $\hat{\beta} = 1$



Lagged Y s and Unit Effects

Consider:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + u_{it}.$$

If we omit the unit effects, we have:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta} + u_{it}^*$$

with

$$u_{it}^* = \alpha_i + u_{it}$$

Lagging yields:

$$Y_{it-1} = \phi Y_{it-2} + \mathbf{X}_{it-1}\boldsymbol{\beta} + \alpha_i + u_{it-1}$$

which means

$$\text{Cov}(Y_{it-1}, u_{it}^*) \neq 0.$$

Bias in $\hat{\phi}$ is

- toward zero when $\phi > 0$,
- increasing in ϕ .

Including unit effects still yields bias in $\hat{\phi}$ of $O(\frac{1}{T})$, and bias in $\hat{\beta}$.

Solutions:

- Difference/GMM estimation
- Bias correction approaches

First Difference Estimation

$$\begin{aligned}Y_{it} - Y_{it-1} &= \phi(Y_{it-1} - Y_{it-2}) + (\mathbf{X}_{it} - \mathbf{X}_{it-1})\beta + (\alpha_i - \alpha_i) + (u_{it} - u_{it-1}) \\ \Delta Y_{it} &= \phi \Delta Y_{it-1} + \Delta \mathbf{X}_{it} \beta + \Delta u_{it}\end{aligned}$$

Anderson/Hsiao: If \nexists autocorrelation, then use ΔY_{it-2} or Y_{it-2} as instruments for ΔY_{it-1} ...

- Consistent in theory,
- in practice, the former is preferred, and both have issues if ϕ is high;
- both are inefficient.

Arellano & Bond (also Wawro): Use *all* lags of Y_{it} and \mathbf{X}_{it} from $t - 2$ and before.

- “Good” estimates, better as $T \rightarrow \infty$,
- Easy to handle higher-order lags of Y ,
- Easy software (plm in R , xtabond in Stata).
- Model *is* fixed effects...
- \mathbf{Z}_i has $T - p - 1$ rows, $\sum_{i=p}^{T-2} i$ columns \rightarrow difficulty of estimation declines in p , grows in T .

Kiviet (1995, 1999; Bun and Kiviet 2003; Bruno 2005a,b): Derive the bias in $\hat{\phi}$ and $\hat{\beta}$, then correct it...

- More accurate than the instrumental-variables/GMM estimators of A&H/A&B...
- ...especially when T is small; but not as T gets reasonably large ($T \approx 20$)

Stationarity: Quick Intro

Mean stationarity:

$$E(Y_t) = \mu \forall t$$

Variance stationarity:

$$\text{Var}(Y_t) = E[(Y_t - \mu)^2] \equiv \sigma_Y^2 \forall t$$

Covariance stationarity:

$$\text{Cov}(Y_t, Y_{t-s}) = E[(Y_t - \mu)(Y_{t-s} - \mu)] = \gamma_s \forall s$$

I(1) Series and Unit Roots

I(1) (“integrated”) series:

$$Y_t = Y_{t-1} + u_t$$

vs. AR(1) series:

$$Y_t = \rho Y_{t-1} + u_t$$

or *trending* series:

$$Y_t = \beta t + u_t$$

Differencing:

$$\begin{aligned}\Delta Y_t \equiv Y_t - Y_{t-1} &= Y_t + u_t - Y_{t-1} \\ &= u_t\end{aligned}$$

and

$$\begin{aligned}\Delta Y_t \equiv Y_t - Y_{t-1} &= \beta t + u_t - (\beta(t-1) + u_{t-1}) \\ &= \beta t + u_t - \beta t + \beta - u_{t-1} \\ &= u_t - u_{t-1} + \beta\end{aligned}$$

I(1) series (continued)

More generally:

- $|\rho| > 1$
 - Series is nonstationary / *explosive*
 - Past shocks have a greater impact than current ones
 - Uncommon
- $|\rho| < 1$
 - *Stationary* series
 - Effects of shocks die out exponentially according to ρ
 - Is mean-reverting
- $|\rho| = 1$
 - Nonstationary series
 - Shocks persist at full force
 - Not mean-reverting; variance increases with t

Unit Root Tests: Dickey-Fuller

Two steps:

- Estimate $Y_t = \rho Y_{t-1} + u_t$,
- test the hypothesis that $\hat{\rho} = 0$, *but*
- this requires that the u s are uncorrelated.

But suppose:

$$\Delta Y_t = \sum_{i=1}^p d_i \Delta Y_{t-i} + u_t$$

which yields

$$Y_t = Y_{t-1} + \sum_{i=1}^p d_i \Delta Y_{t-i} + u_t.$$

D.F. tests will be incorrect.

Augmented Dickey-Fuller Tests:

- Estimate

$$\Delta Y_t = Y_{t-1} + \sum_{i=1}^p d_i \Delta Y_{t-i} + u_t$$

- Test $\hat{\rho} = 0$

Phillips-Perron Tests:

- Estimate:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + u_t$$

- Calculate modified test statistics (Z_ρ and Z_t)
- Test $\hat{\rho} = 0$

Issues with Unit Roots in Panel Data

- Short series + Asymptotic tests \rightarrow “borrow strength”
- Requires uniform unit roots across i s
- Various alternatives:
 - Maddala and Wu (1999)
 - Hadri (2000)
 - Levin, Lin and Chu (2002)
- What to do?
 - Difference the data...
 - Error-correction models

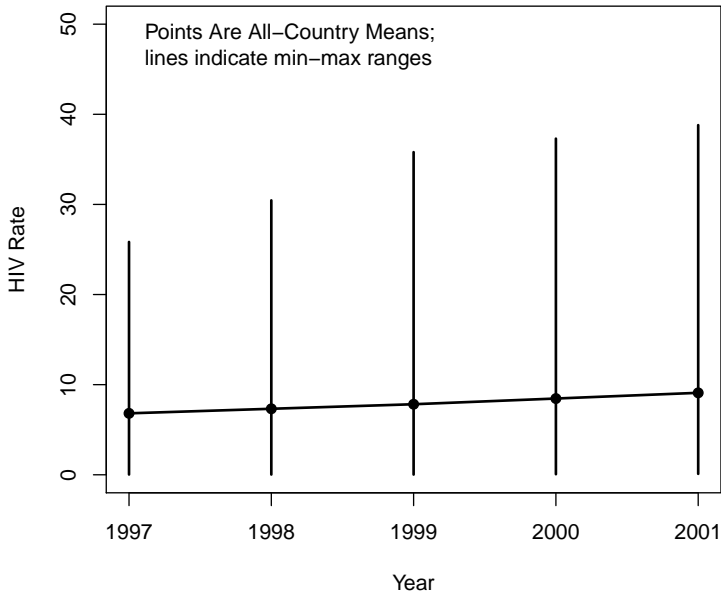
Example: HIV/AIDS in Africa, 1997-2001

```
> summary(AIDS)
```

cocode	year	lnAIDS	lnAIDSlag	warlag	popden
Min. :404	Min. :1997	Min. : -3.8	Min. : -4	Min. :0.00	Min. :0.00
1st Qu.:451	1st Qu.:1998	1st Qu.: 0.6	1st Qu.: 1	1st Qu.:0.00	1st Qu.:0.01
Median :506	Median :1999	Median : 1.6	Median : 2	Median :0.00	Median :0.03
Mean :510	Mean :1999	Mean : 1.2	Mean : 1	Mean :0.14	Mean :0.06
3rd Qu.:560	3rd Qu.:2000	3rd Qu.: 2.4	3rd Qu.: 2	3rd Qu.:0.00	3rd Qu.:0.07
Max. :651	Max. :2001	Max. : 3.7	Max. : 4	Max. :1.00	Max. :0.57

refsin
Min. : 0
1st Qu.: 1
Median : 10
Mean : 56
3rd Qu.: 46
Max. :543

HIV/AIDS in Africa, 1997-2001



Panel Unit Root Tests: R

```
> lnAIDS<-cbind(AIDS$ccode,AIDS$year,AIDS$lnAIDS)
> purtest(lnAIDS,exo="trend",test=c("levinlin"))
```

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

```
data: lnAIDS
z.x1 = 3e+12, p-value <2e-16
alternative hypothesis: stationarity
```

```
> purtest(lnAIDS,exo="trend",test=c("hadri"))
```

Hadri Test (ex. var.: Individual Intercepts and Trend)

```
data: lnAIDS
z = 60, p-value <2e-16
alternative hypothesis: at least one series has a unit root
```

```
> purtest(lnAIDS,exo="trend",test=c("ips"))
```

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts and Trend)

```
data: lnAIDS
z = 4, p-value = 0.0002
alternative hypothesis: stationarity
```

Final Thoughts: Dynamic Panel Models

- N vs. T ...
- Are dynamics nuisance or substance?
- What problem(s) do you *really* care about?