GSERM - Oslo 2019

Hierarchical / Multilevel Models

January 8, 2019 (morning session)

“Robust” Variance-Covariance Estimators

Linear Model: $\text{Var}(\hat{\beta})$ with $\mathbf{uu}' = \sigma^2\Omega$:

omega should be I in standard OLS

$$\begin{aligned} \text{Var}(\beta_{\text{Het.}}) &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{Q}\,(\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

where $\mathbf{Q} = (\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})$ and $\mathbf{W} = \sigma^2\boldsymbol{\Omega}$.

Rewrite:

$$\begin{aligned} \mathbf{Q} &= \sigma^2(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}) \\ &= \sum_{i=1}^{N} \sigma_i^2 \mathbf{X}_i \mathbf{X}_i' \end{aligned}$$

“Robust” Variance-Covariance Estimators

squaring the empircal
errors gives a very rough
insight if the error variance
is small or big due its
exponential nature

White's Insight:

$$\widehat{\mathbf{Q}} = \sum_{i=1}^{N} \hat{u}_i^2 \mathbf{X}_i \mathbf{X}_i'$$

$$\begin{aligned}
\widehat{\mathrm{Var}(\beta)}_{\mathrm{Robust}} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\widehat{\mathbf{Q}}^{-1}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1} \\
&= (\mathbf{X}'\mathbf{X})^{-1}\left[\mathbf{X}'\left(\sum_{i=1}^{N} \hat{u}_i^2 \mathbf{X}_i \mathbf{X}_i'\right)^{-1}\mathbf{X}\right](\mathbf{X}'\mathbf{X})^{-1}
\end{aligned}$$

errors are adjusted for the empirical error,
big errors more weighted, small errors less weighted

it is consistent estimator even in heterosecadistic situation

What about MLE?

Recall:

$$\begin{aligned}\mathrm{Var}(\hat{\theta}) &= \mathrm{E}[(\hat{\theta}-\theta)(\hat{\theta}-\theta)'] \\ &= \mathrm{E}\left[\left(-\frac{\partial^2 \ln L}{\partial \theta^2}\right)^{-1} \frac{\partial \ln L}{\partial \theta}\frac{\partial \ln L}{\partial \theta}' \left(-\frac{\partial^2 \ln L}{\partial \theta^2}\right)^{-1}\right]\end{aligned}$$

We assumed:

$$\mathrm{E}\left[\frac{\partial \ln L}{\partial \theta}\frac{\partial \ln L}{\partial \theta}'\right] = \mathrm{E}\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right]$$

So,

$$\begin{aligned}\mathrm{Var}(\hat{\theta}) &= \left[-\mathrm{E}\left(\frac{\partial^2 \ln L}{\partial \theta^2}\right)\right]^{-1} \\ &= [\mathbf{I}(\theta)]^{-1}\end{aligned}$$

Alternatively:

$$\mathrm{Var}(\hat{\theta})_{\mathrm{Robust}} = [\mathbf{I}(\theta)]^{-1}\left(\frac{\partial \ln L}{\partial \hat{\theta}}\frac{\partial \ln L}{\partial \hat{\theta}}'\right)[\mathbf{I}(\theta)]^{-1}$$

"Clustering"

when you know certain observations are common to each other
with similar error variability, but other cluster have different error variability

should be derived from substantive knowledge, that within cluster variability is similar, but not between. but we do not know the exact error variance.

Suppose N "clusters" $i = \{1, 2, ...N\}$, each with n_i observations $j = \{1, 2, ...n_i\}$.

Model:

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

Then:

$$\widehat{\text{Var}(\boldsymbol{\beta})}_{\text{Clustered}} = (\mathbf{X}'\mathbf{X})^{-1}\left\{\mathbf{X}'\left[\sum_{i=1}^{N}\left(\sum_{j=1}^{n_j}\hat{u}_{ij}^2\mathbf{X}_{ij}\mathbf{X}'_{ij}\right)\right]^{-1}\mathbf{X}\right\}(\mathbf{X}'\mathbf{X})^{-1}$$

within cluster

sum across cluster----------

then use in the formula

An Illustration

"Regular" OLS:

```
> id<-seq(1,100,1) # 100 observations
> set.seed(7222009)
> x<-rnorm(100) # N(0,1) noise
> y<-1+1*x+rnorm(100)*abs(x)
> library(rms)
> fit<-ols(y~x,x=TRUE,y=TRUE)
> fit

Linear Regression Model

ols(formula = y ~ x, x = TRUE, y = TRUE)
                Model Likelihood     Discrimination
                   Ratio Test           Indexes
Obs      100    LR chi2     61.54    R2       0.460
sigma 0.9538    d.f.            1    R2 adj   0.454
d.f.      98    Pr(> chi2) 0.0000    g        1.002

Residuals

     Min       1Q   Median       3Q      Max
-3.27767 -0.54898  0.09069  0.35771  2.95014

          Coef   S.E.   t    Pr(>|t|)
Intercept 0.8867 0.0954 9.30 <0.0001
x         0.8822 0.0966 9.13 <0.0001
```

Further Illustration: “Robust” $\hat{V}$

```
> RVCV<-robcov(fit)
> RVCV

Linear Regression Model

ols(formula = y ~ x, x = TRUE, y = TRUE)
                Model Likelihood     Discrimination
                   Ratio Test           Indexes
Obs       100    LR chi2     61.54    R2       0.460
sigma 0.9538    d.f.            1    R2 adj   0.454
d.f.       98    Pr(> chi2) 0.0000    g        1.002

Residuals

     Min        1Q   Median       3Q      Max
-3.27767 -0.54898  0.09069  0.35771  2.95014

          Coef   S.E.   t    Pr(>|t|)
Intercept 0.8867 0.0943 9.41 <0.0001
x         0.8822 0.1352 6.52 <0.0001
```

robust s.e. are bigger

Attack of the Clones

replicate 16x times

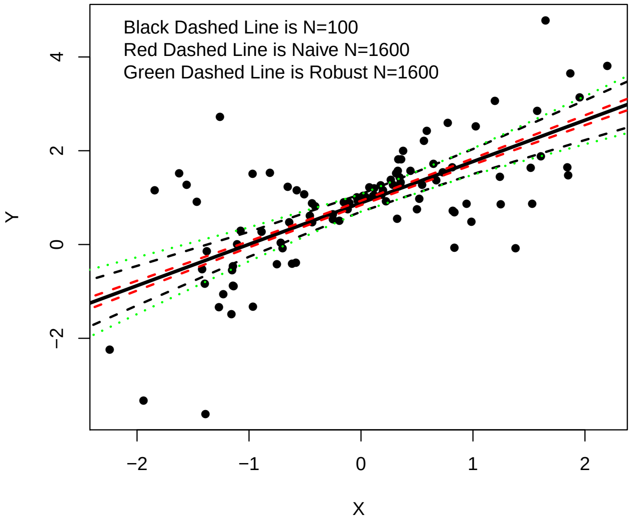
```
> bigID<-rep(id,16)
> bigX<-rep(x,16)
> bigY<-rep(y,16)
> bigdata<-as.data.frame(cbind(bigID,bigY,bigX))
> bigOLS<-ols(bigY~bigX,data=bigdata,x=TRUE,y=TRUE)
> bigOLS

Linear Regression Model

ols(formula = bigY ~ bigX, data = bigdata, x = TRUE, y = TRUE)
                Model Likelihood     Discrimination
                   Ratio Test           Indexes
Obs      1600   LR chi2     984.69    R2         0.460
sigma 0.9448    d.f.             1    R2 adj     0.459
d.f.     1598   Pr(> chi2) 0.0000     g          0.993

Residuals

     Min       1Q   Median       3Q      Max
-3.27767 -0.54898  0.09069  0.35771  2.95014

          Coef   S.E.   t     Pr(>|t|)
Intercept 0.8867 0.0236 37.54 <0.0001
bigX      0.8822 0.0239 36.86 <0.0001
```

```
> BigRVCV<-robcov(bigOLS,bigdata$bigID)
> BigRVCV

Linear Regression Model

ols(formula = bigY ~ bigX, data = bigdata, x = TRUE, y = TRUE)
                            Model Likelihood     Discrimination
                               Ratio Test           Indexes
Obs                  1600    LR chi2    984.69    R2        0.460
sigma              0.9448    d.f.            1    R2 adj    0.459
d.f.                 1598    Pr(> chi2) 0.0000    g         0.993
Cluster on bigdata$bigID
Clusters              100

Residuals

     Min        1Q    Median        3Q       Max
-3.27767 -0.54898  0.09069  0.35771  2.95014

          Coef   S.E.   t    Pr(>|t|)
Intercept 0.8867 0.0943 9.41 <0.0001
bigX      0.8822 0.1352 6.52 <0.0001
```

cluster by bigID

same s.e. as before with 100 observations

Illustrated...



‘Robust” Variance Estimators: Cautions

- Are *only* consistent (Chesher and Jewitt 1987)
- Efficiency loss if homoscedastic (Kauermann and Carroll 2001)
- “Even if the key assumption holds, bias should be of greater interest than variance, especially when the sample is large and causal inferences are based on a model that is incorrectly specifed. Variances will be small, and bias may be large.” (Freedman 2006)

robust estimators are only asymptotically unbiased, so bad (may get wrong inferences in small samples)

Things you should read...

see also slide before about freedman

Freedman, D. A. 2006. "On the So-Called 'Huber Sandwich Estimator' and 'Robust' Standard Errors." *The American Statistician* 60:299-302.

Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions." *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* I:221-33.

White, H. 1994. *Estimation, Inference, and Specifcation Analysis*. New York: Cambridge University Press.

Big parts of this course are based on White 1994

Hierarchical Linear Models

HLM Starting Points

Begin by considering a two-level "nested" data structure, with:

both variables are invariant to the other level

$$\begin{aligned} i &\in \{1, 2, ...N\} \text{ indexing first-level units, and} \\ j &\in \{1, 2, ...J\} \text{ indexing second-level groups.} \end{aligned}$$

A general two-level HLM is an equation of the form:

$$Y_{ij} = \beta_{0j} + \mathbf{X}_{ij}\boldsymbol{\beta}_j + u_{ij} \tag{1}$$

where β_{0j} is a "constant" term, $\mathbf{X}_{ij}$ is a $NJ \times K$ matrix of K covariates, $\boldsymbol{\beta}_j$ is a $K \times 1$ vector of parameters, and $u_{ij} \sim$ i.i.d. $N(0, \sigma_u^2)$ is the usual random-disturbance assumption.

HLMs, continued

Each of these $K+1$ "level-one" parameters is then allowed to vary across Q "level-two" variables $\mathbf{Z}_j$, so that:

$$\beta_{0j} = \gamma_{00} + \mathbf{Z}_j\gamma_0 + \varepsilon_{0j} \tag{2}$$

for the "intercept" and

$$\beta_{kj} = \gamma_{k0} + \mathbf{Z}_j\gamma_k + \varepsilon_{kj} \tag{3}$$

for the "slopes" of $\mathbf{X}$. The εs are typically assumed to be distributed multivariate Normal, with parameters for the variances and covariances selected by the analyst. Substitution of (3) and (2) into (1) yields:

first level / 2nd level / interaction /errors

$$Y_{ij} = \gamma_{00} + \mathbf{Z}_j\gamma_0 + \mathbf{X}_{ij}\gamma_{k0} + \mathbf{X}_{ij}\mathbf{Z}_j\gamma_k + \mathbf{X}_{ij}\varepsilon_{kj} + \varepsilon_{0j} + u_{ij} \tag{4}$$

The form is essentially one of a model with "saturated" interaction effects across the various levels, as well as "errors" which are multivariate Normal.

- Linearity / Additivity
- Normality of us
- Homoscedasticity
- Residual Independence:
  - $\text{Cov}(\varepsilon_{\cdot j}, u_{ij}) = 0$
  - $\text{Cov}(u_{ij}, u_{i\ell}) = 0$

independent errors between hierarchical levels
and independent errors between units

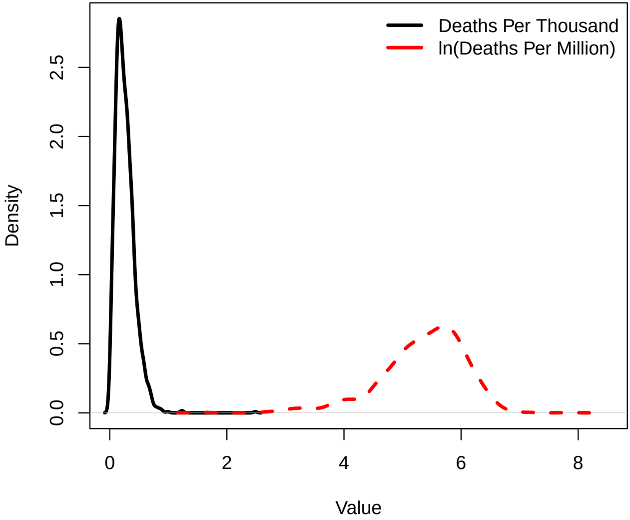
Two main alternatives:

- MLE
- "Restricted" MLE ("RMLE")
- Choosing:
  - MLE is biased in small samples, especially for estimating variance effects. variance in second level group can be small (many students, but only 12 classes
  - RMLE is not, but prevents use of LR tests when the models do not have identical fixed effects.
  - In general: RMLE is better with small sample sizes, but MLE is fine in larger ones.

An Example: HIV Death Rates, 1990-2007

```
> temp<-getURL("https://raw.githubusercontent.com/PrisonRodeo/GSERM-Oslo-2019-git/master/Data/HIVDeaths.csv")
> HIV<-read.csv(text=temp, header=TRUE)
> HIV<-HIV[ which(is.na(HIV$HIVDeathRate)==FALSE), ]
> HIV$LnDeathPM <- log(HIV$HIVDeathRate*1000)
> summary(HIV)
      country           ISO3           year       HIVDeathRate
 Angola   :  18   AGO    :  18   Min.   :1990   Min.   :0.00478
 Argentina:  18   ARG    :  18   1st Qu.:1995   1st Qu.:0.14429
 Australia:  18   AUS    :  18   Median :2000   Median :0.23303
 Benin    :  18   BDI    :  18   Mean   :1999   Mean   :0.26126
 Botswana :  18   BEN    :  18   3rd Qu.:2004   3rd Qu.:0.34889
 Brazil   :  18   BFA    :  18   Max.   :2007   Max.   :2.48542
 (Other)  :1540   (Other):1540
 CivilWarDummy      OPENLag         GDPGrowthLag       POLITYLag
 Min.   :0.000   Min.   :  1.09   Min.   :-62.368   Min.   :-10.00
 1st Qu.:0.000   1st Qu.: 44.31   1st Qu.: -0.458   1st Qu.: -4.00
 Median :0.000   Median : 61.21   Median :  1.961   Median :  6.00
 Mean   :0.181   Mean   : 74.29   Mean   :  1.899   Mean   :  2.97
 3rd Qu.:0.000   3rd Qu.: 97.37   3rd Qu.:  4.428   3rd Qu.:  9.00
 Max.   :1.000   Max.   :456.56   Max.   : 88.748   Max.   : 10.00
                 NA's   :30       NA's   :32        NA's   :63
  POLITYSQLag    InterstateWarLag   PolityLag    BatDeaths1000Lag
 Min.   :  0.0   Min.   :0.00000   Min.   : 0   Min.   : 0.000
 1st Qu.: 25.0   1st Qu.:0.00000   1st Qu.: 6   1st Qu.: 0.000
 Median : 49.0   Median :0.00000   Median :16   Median : 0.000
 Mean   : 49.5   Mean   :0.00364   Mean   :13   Mean   : 0.264
 3rd Qu.: 81.0   3rd Qu.:0.00000   3rd Qu.:19   3rd Qu.: 0.000
 Max.   :100.0   Max.   :1.00000   Max.   :20   Max.   :30.239
 NA's   :63                        NA's   :63
    GDPLagK          LnDeathPM
 Min.   : 0.153   Min.   :1.57
 1st Qu.: 1.576   1st Qu.:4.97
 Median : 5.011   Median :5.45
 Mean   : 8.582   Mean   :5.35
 3rd Qu.:10.265   3rd Qu.:5.85
 Max.   :42.683   Max.   :7.82
 NA's   :30
```

Log? Si.


Deaths Per Thousand
ln(Deaths Per Million)
Density
Value
0.0
0.5
1.0
1.5
2.0
2.5
0
2
4
6
8

OLS Regression

```
> OLSfit<-with(HIV, lm(LnDeathPM~GDPLagK+GDPGrowthLag+
+                      OPENLag+POLITYLag+POLITYSQLag+CivilWarDummy+
+                      InterstateWarLag+BatDeaths1000Lag))
> summary(OLSfit)

Call:
lm(formula = LnDeathPM ~ GDPLagK + GDPGrowthLag + OPENLag + POLITYLag +
    POLITYSQLag + CivilWarDummy + InterstateWarLag + BatDeaths1000Lag)

Residuals:
   Min     1Q Median     3Q    Max
-3.940 -0.388  0.095  0.447  1.953

Coefficients:
                  Estimate Std. Error t value    Pr(>|t|)
(Intercept)       5.493740   0.044516  123.41     < 2e-16 ***
GDPLagK          -0.027965   0.002509  -11.15     < 2e-16 ***
GDPGrowthLag     -0.002261   0.002430   -0.93      0.3524
OPENLag           0.001972   0.000368    5.35 0.000000099 ***
POLITYLag         0.010009   0.003356    2.98      0.0029 **
POLITYSQLag      -0.002182   0.000734   -2.97      0.0030 **
CivilWarDummy     0.051862   0.047026    1.10      0.2703
InterstateWarLag  0.129922   0.283361    0.46      0.6467
BatDeaths1000Lag -0.024675   0.011732   -2.10      0.0356 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.651 on 1548 degrees of freedom
  (91 observations deleted due to missingness)
Multiple R-squared:  0.177,Adjusted R-squared:  0.173
F-statistic: 41.7 on 8 and 1548 DF,  p-value: <2e-16
```

Fixed Effects

```
> FEfit<-plm(LnDeathPM~GDPLagK+GDPGrowthLag+OPENLag+POLITYLag+POLITYSQLag+CivilWarDummy+
+                       InterstateWarLag+BatDeaths1000Lag,data=HIV,effect="individual", model="within",
+                       index=c("ISO3","year"))
> summary(FEfit)
Oneway (individual) effect Within Model

Call:
plm(formula = LnDeathPM ~ GDPLagK + GDPGrowthLag + OPENLag +
    POLITYLag + POLITYSQLag + CivilWarDummy + InterstateWarLag +
    BatDeaths1000Lag, data = HIV, effect = "individual", model = "within",
    index = c("ISO3", "year"))

Unbalanced Panel: n=117, T=1-18, N=1557

Coefficients :
                   Estimate Std. Error t-value Pr(>|t|)
GDPLagK          -0.0987550  0.0094605 -10.439  < 2e-16 ***
GDPGrowthLag      0.0045675  0.0020894   2.186    0.029 *
OPENLag           0.0077044  0.0009468   8.138 8.67e-16 ***
POLITYLag         0.0505600  0.0051147   9.885  < 2e-16 ***
POLITYSQLag      -0.0006743  0.0009589  -0.703    0.482
CivilWarDummy     0.0751139  0.0534712   1.405    0.160
InterstateWarLag -0.3030380  0.2396271  -1.265    0.206
BatDeaths1000Lag  0.0004229  0.0103239   0.041    0.967
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Total Sum of Squares:    445.6
Residual Sum of Squares: 378.6
R-Squared:      0.1505
Adj. R-Squared: 0.1384
F-statistic: 31.7023 on 8 and 1432 DF, p-value: < 2.2e-16
```

Random Effects (using `lmer`)

```
> REfit<-lmer(LnDeathPM~GDPLagK+GDPGrowthLag+OPENLag+POLITYLag+POLITYSQLag+CivilWarDummy+
+               InterstateWarLag+BatDeaths1000Lag+(1|ISO3),data=HIV,REML=FALSE)
> summary(REfit)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula:
LnDeathPM ~ GDPLagK + GDPGrowthLag + OPENLag + POLITYLag + POLITYSQLag +
    CivilWarDummy + InterstateWarLag + BatDeaths1000Lag + (1 | ISO3)
   Data: HIV

     AIC      BIC   logLik deviance df.resid
  2698.9   2757.7  -1338.4   2676.9     1546

Random effects:
 Groups   Name        Variance Std.Dev.
 ISO3     (Intercept) 0.265    0.515
 Residual             0.270    0.520
Number of obs: 1557, groups:  ISO3, 117

Fixed effects:
                  Estimate Std. Error t value
(Intercept)       5.272156   0.086694    60.8
GDPLagK          -0.050509   0.005092    -9.9
GDPGrowthLag      0.002749   0.002077     1.3
OPENLag           0.004776   0.000706     6.8
POLITYLag         0.044502   0.004565     9.7
POLITYSQLag      -0.000964   0.000888    -1.1
CivilWarDummy     0.060362   0.052101     1.2
InterstateWarLag -0.251942   0.240937    -1.0
BatDeaths1000Lag -0.003502   0.010331    -0.3

Correlation of Fixed Effects:
            (Intr) GDPLgK GDPGrL OPENLg POLITYL POLITYS CvlWrD IntrWL
GDPLagK     -0.172
GDPGrowthLg -0.032 -0.051
OPENLag     -0.554 -0.222 -0.015
POLITYLag   -0.047 -0.222  0.002  0.017
POLITYSQLag -0.373 -0.341  0.000  0.054 -0.051
CivilWrDmmy -0.194 -0.002  0.076  0.074  0.126   0.060
IntrsttWrLg -0.005  0.014 -0.025 -0.009 -0.028   0.013   0.023
BtDths1000L -0.045 -0.013  0.129  0.044  0.056  -0.019  -0.105 -0.329
```

HLM with Random β for GDP

```
> HLMfit1<-lmer(LnDeathPM~GDPLagK+(GDPLagK|ISO3)+GDPGrowthLag+OPENLag+POLITYLag+POLITYSQLag+CivilWarDummy+
+              InterstateWarLag+BatDeaths1000Lag,data=HIV,REML=FALSE)
> summary(HLMfit1)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula:
LnDeathPM ~ GDPLagK + (GDPLagK | ISO3) + GDPGrowthLag + OPENLag +  POLITYLag + POLITYSQLag + CivilWarDummy +
InterstateWarLag + BatDeaths1000Lag
   Data: HIV

     AIC      BIC   logLik deviance df.resid
  2298.8   2368.4  -1136.4   2272.8     1544

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ISO3     (Intercept) 9.168    3.028
          GDPLagK     0.200    0.447    -0.74
 Residual             0.136    0.369
Number of obs: 1557, groups:  ISO3, 117

Fixed effects:
                  Estimate Std. Error t value
(Intercept)       4.791024   0.302393   15.84
GDPLagK           0.155304   0.048233    3.22
GDPGrowthLag      0.000872   0.001555    0.56
OPENLag           0.005995   0.000834    7.19
POLITYLag         0.039930   0.003959   10.09
POLITYSQLag      -0.003896   0.000770   -5.06
CivilWarDummy     0.009747   0.040489    0.24
InterstateWarLag -0.261331   0.178583   -1.46
BatDeaths1000Lag  0.013020   0.007920    1.64

Correlation of Fixed Effects:
            (Intr) GDPLgK GDPGrL OPENLg POLITYL POLITYS CvlWrD IntrWL
GDPLagK     -0.686
GDPGrowthLg  0.018 -0.067
OPENLag     -0.120 -0.085  0.002
POLITYLag   -0.018 -0.033 -0.007 -0.074
POLITYSQLag -0.084 -0.055  0.002 -0.019  0.039
CivilWrDmmy -0.041 -0.004  0.080  0.025  0.101   0.052
IntrsttWrLg -0.009  0.005 -0.020  0.018 -0.039   0.017   0.019
BtDths1000L -0.009 -0.008  0.101  0.065  0.063  -0.052  -0.095 -0.353
```

Testing

```
> anova(REfit,HLMfit1)

Data: HIV

Models:
REfit: LnDeathPM ~ GDPLagK + GDPGrowthLag + OPENLag + POLITYLag + POLITYSQLag +
REfit:     CivilWarDummy + InterstateWarLag + BatDeaths1000Lag + (1 |
REfit:     ISO3)
HLMfit1: LnDeathPM ~ GDPLagK + (GDPLagK | ISO3) + GDPGrowthLag + OPENLag +
HLMfit1:     POLITYLag + POLITYSQLag + CivilWarDummy + InterstateWarLag +
HLMfit1:     BatDeaths1000Lag

        Df  AIC  BIC logLik deviance Chisq Chi Df Pr(>Chisq)
REfit   11 2699 2758  -1338     2677
HLMfit1 13 2299 2368  -1136     2273 404.1      2     <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

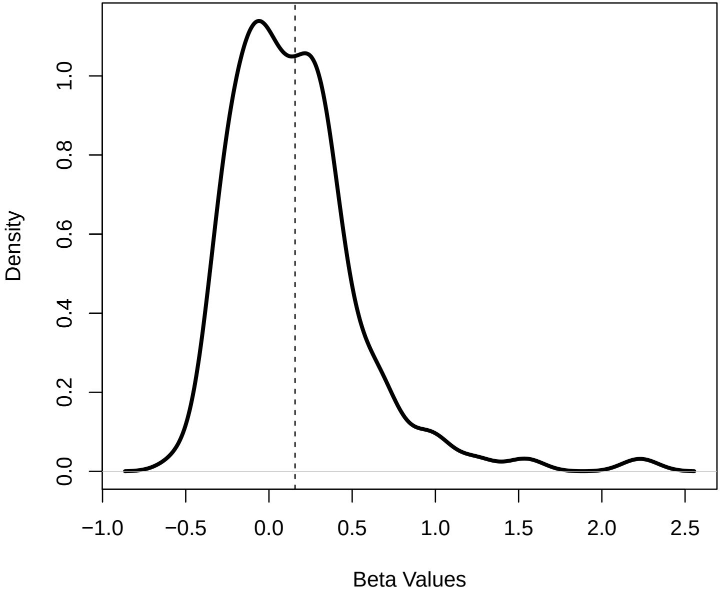
Random Coefficients

some random slopes, some fixed.

```
> Bs<-data.frame(coef(HLMfit1)[1])
>
> head(Bs)
    ISO3..Intercept. ISO3.GDPLagK ISO3.GDPGrowthLag ISO3.OPENLag
AGO          3.96339    0.3234238       0.000869237   0.00598492
ARG          3.57905    0.1164726       0.000869237   0.00598492
ARM          5.07487    0.1142131       0.000869237   0.00598492
AUS          9.97544   -0.1999752       0.000869237   0.00598492
AUT          7.08153   -0.0845660       0.000869237   0.00598492
AZE          3.80985    0.0133378       0.000869237   0.00598492...
>
>
> mean(Bs$ISO3.GDPLagK)
[1] 0.156798
```

Random Coefficients (Plotted)

for gdp



Wrap-Up & Extensions

- Can expand to 3- and 4- and higher-level models (e.g., students in classrooms in schools in districts)
- Cross-Level Interactions...
- Widely used in education, psychology, ecology, etc. (less so in economics, political science)
- There are many, many excellent books, websites, etc. that address HLMs