```
* CVER presentation demo : analyses restricted to first round of scores
use "...", clear /* dataset not public */
drop if thirdreviewer==1
keep paperid reviewerid overallscore0to100
rename overallscore0to100 overall

codebook paperid reviewerid overall, c
bysort paperid: gen wpaper=_n
bysort reviewerid: gen wrev=_n
by reviewerid: gen npaper=_N
tab npaper if wrev==1
bysort paperid: egen rawmean=mean(overall)

* fixed effects model
reg overall i.paperid i.reviewerid /* no collinearity */

* cross-classified random effects model
mixed overall || _all:R.paperid || reviewerid:, reml
* same results with paperid and reviewerid reversed
* same results as with lme4 in R using model specification:
*   overall ~ (1|reviewerid)+(1|paperid)
di 63.95901/(63.95901+126.869+160.2216) /* abstract variance proportion */
di 126.869/(63.95901+126.869+160.2216) /* reviewer variance proportion */

predict stdres, rstandard
qnorm stdres
sum stdres, d
predict reff*, reffects
qnorm reff1 if wpaper==1
qnorm reff2 if wrev==1
predict fit
gen fullfit=fit+reff1+reff2
```

```
scatter stdres fullfit
* transformation of outcome seems needed
* work instead with overall2=overall^2 (determined by Box-Cox analyses, see added code below)

gen sqoverall=overall^2/100
capture drop stdres reff* fit fullfit
mixed sqoverall || _all:R.paperid || reviewerid:, reml
predict stdres, rstandard
qnorm stdres
sum stdres, d
predict reff*, reffects
qnorm reff1 if wpaper==1
qnorm reff2 if wrev==1
predict fit
gen fullfit=fit+reff1+reff2
scatter stdres fullfit
* much better
predict modelscore, reffects relevel(_all)

* abstract random effects only
mixed sqoverall i.reviewerid || paperid:, reml
predict modelfixscore, reffects

format modelscore modelfixscore %9.2f
preserve
keep if wpaper==1 /* one record per paper */
egen rawrank=rank(rawmean)
egen modelrank=rank(modelscore)
egen modelfixrank=rank(modelfixscore)
gen diffrank_crudemod=abs(rawrank-modelrank)
gen diffrank_crudemfix=abs(rawrank-modelfixrank)
gen diffrank_modmodfix=abs(modelrank-modelfixrank)
```

```
sum diffrank*, d


sort rawrank

list paperid rawmean modelscore modelfixscore rawrank modelrank modelfixrank if abs(diffrank_crudemod)>30,
sep(20)


gen highrank_crude=(rawrank>=119-68+1) /* 68 highest scores ~ all abstracts ranked >=52 */

gen highrank_model=(modelrank>=119-68+1)

tab highrank_crude highrank_model

gen highrank_modelfix=(modelfixrank>=119-68+1)

tab highrank_crude highrank_modelfix

tab highrank_model highrank_modelfix


list paperid rawmean modelscore modelfixscore rawrank modelrank highrank_crude if
highrank_crude!=highrank_model, sep(20)

list paperid rawmean modelscore modelfixscore rawrank modelfixrank highrank_crude if
highrank_crude!=highrank_modelfix, sep(20)

list paperid rawmean modelscore modelfixscore modelrank modelfixrank highrank_model if
highrank_model!=highrank_modelfix, sep(20)

restore


** added code for Box-Cox analyses

gen overallp1=overall+1

xi: boxcox overallp1 i.paperid i.reviewerid /* estimated power=2.06 */


preserve

tempname memhold

postfile memhold lambda lnL pl using temp.dta, replace

gen y=.


* loop over lambda values

gen lnoverall=ln(overall)

egen nobs=count(paperid)
```

```
egen meanln=mean(lnoverall)

foreach lambda of numlist 3 2.75 2.5 2.25 2 1.5 1 0.5 {
  di `lambda'
  replace y=(overall^(`lambda')-1)/(`lambda')
  replace y=lnoverall if (`lambda')==0
  mixed y || _all:R.paperid || reviewerid:
  scalar lnL=e(ll)
  scalar pl=lnL+nobs*(`lambda'-1)*meanln
  post memhold (`lambda') (lnL) (pl)
  }
* stop the process of posting results to a file
postclose memhold
* open the file that captured the results
use c:\data.avc\conf\ssc22\abstract\temp.dta, clear
list lambda lnL pl
scatter pl lambda, ytitle(profile log-likelihood) xtitle(lambda)
* optimal lambda somewhere 2-2.25
restore

* try also with inverted outcome
gen cover=100-overalls
xi: boxcox cover i.paperid i.reviewerid

preserve
tempname memhold
postfile memhold lambda lnL pl temp.dta, replace
gen y=.

* loop over lambda values
gen lncover=ln(cover)
egen nobs=count(paperid)
```

```
egen meanln=mean(lncover)


foreach lambda of numlist 1 0.5 0.3333 0.25 0.2 0.1 0 -0.1 -0.25 -0.3333 -0.5 {
  di `lambda'
  replace y=(cover^(`lambda')-1)/(`lambda')
  replace y=lncover if (`lambda')==0
  mixed y || _all:R.paperid || reviewerid:
  scalar lnL=e(ll)
  scalar pl=lnL+nobs*(`lambda'-1)*meanln
  post memhold (`lambda') (lnL) (pl)
  }
* stop the process of posting results to a file
postclose memhold
* open the file that captured the results
use c:\data.avc\conf\ssc22\abstract\temp.dta, clear
list lambda lnL pl
scatter pl lambda, ytitle(profile log-likelihood) xtitle(lambda)
* optimal lambda around 0.2
restore
```