Advanced Quantitative Techniques (Class 13)

Gregory M. Eirich
QMSS

Agenda

1. Multilevel models
2. Two-stage multilevel models

1. Multilevel models

Multilevel modeling

- We have already seen multilevel models with fixed effects, random effects and growth curves
- There, we mostly looked at “person-years” nested within individuals; aka, the same person over time

Basic multilevel model terminology

We have smaller things nested in bigger things:

- Students (Level 1) within Schools (Level 2)
- Person-years (Level 1) within Individuals (Level 2)
- Voters (Level 1) within States (Level 2)
- Etc., etc.

Motivation for multilevel models

The technical reason:

- We have within-unit correlations that invalidate our “i.i.d.” errors assumption
- Observations from the same unit may share unmeasured similarities that lead them to be more “clustered” together than they are to a random observation from some other unit
- If our “i.i.d.” assumption is violated, then the typical estimation of our standard errors is incorrect (and hence, underestimated)

Motivation for multilevel models

The substantive reason(s):

- We want to exploit and explore this within-unit correlation, or clustering
- Why?
- Mainly because **context matters**: members within one unit may have an effect on other members within that unit
- We are interested in how much variation there is within units – and why it is high or low

Motivation for multilevel models

The substantive reason(s), continued:

- We want to explore **interactions** across units of analysis
- Things that happen at a higher level may have differential impacts on entities at the lower level
- E.g., School characteristics may affect some students differently than others

Heterogeneity

Causal processes in psychology are heterogeneous.

EXPORT   Add To My List   © Request Permissions   Database: APA PsycArticles   Journal Article

Bolger, Niall   Zee, Katherine S.   Rossignac-Milon, Maya   Hassin, Ran R.

Citation

Bolger, N., Zee, K. S., Rossignac-Milon, M., & Hassin, R. R. (2019). Causal processes in psychology are heterogeneous. *Journal of Experimental Psychology: General, 148*(4), 601–618. https://doi.org/10.1037/xge0000558

Abstract

All experimenters know that human and animal subjects do not respond uniformly to experimental treatments. Yet theories and findings in experimental psychology either ignore this causal effect heterogeneity or treat it as uninteresting error. This is the case even when data are available to examine effect heterogeneity directly, in within-subjects designs where experimental effects can be examined subject by subject. Using data from four repeated-measures experiments, we show that effect heterogeneity can be modeled readily, that its discovery presents exciting opportunities for theory and methods, and that allowing for it in study designs is good research practice. This evidence suggests that experimenters should work from the assumption that causal effects are heterogeneous. Such a working assumption will be of particular benefit, given the increasing diversity of subject populations in psychology.

Multilevel modeling

- The GSS “community” variable, *sampcode*
- *sampcode* = the large area or “community” where the respondent lives
- Smaller than a state, but bigger than a city … an MSA or so
- These sampcodes change over time, as new Censuses are conducted, but we can make a complete list of them based on their id and each year

To create our community IDs

We create an id for each person’s community in that given year of the survey

```
> vars <- c("sampcode", "year", "closeblk", "race", "age")
> sub <- GSS[,vars]
>
> # create community IDs
> sub$sampyear <- with(sub, sampcode*10^4 + year)
> head(Tab(sub$sampyear), 15)
      Count  Pct Cum.Pct
11993    52 0.09    0.09
11994   189 0.34    0.43
11996   189 0.34    0.78
11998   182 0.33    1.10
12000   195 0.35    1.46
12002   195 0.35    1.81
21993    36 0.06    1.87
21994   135 0.24    2.12
21996   133 0.24    2.36
21998   108 0.19    2.55
22000   123 0.22    2.77
22002   109 0.20    2.97
```

Our research question:

Does one's race influence how (emotionally) close they feel to African-Americans? Does it matter how many African-Americans live around the person? [Other questions too …]

- N.B., For ease of explanation, we will limit ourselves to just whites and African-Americans

Our dependent variable:

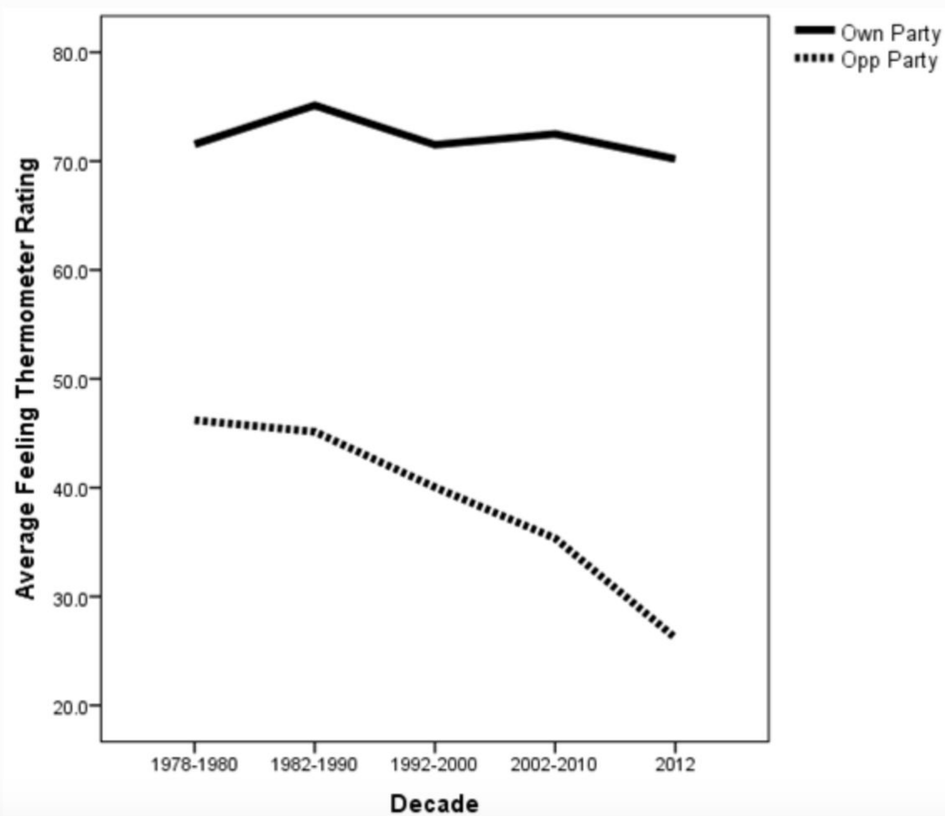
Closeblk = "In general, how close do you feel to blacks," ranging from *not at all close (1)* to *very close (9)*

```
> Tab(sub$closeblk)
  Count   Pct Cum.Pct
1   686  5.58    5.58
2   333  2.71    8.29
3   490  3.98   12.27
4   502  4.08   16.35
5  5235 42.57   58.92
6  1020  8.29   67.22
7  1479 12.03   79.25
8   698  5.68   84.92
9  1854 15.08  100.00
```

These questions akin to feeling thermometers

Especially prominent in politics (like here)

Figure 2: Average feeling thermometer ratings of own party and opposing party by decade



*

Thermometers along other dimensions too

Like [here]()

Americans Express Increasingly Warm Feelings Toward Religious Groups

Jews, Catholics continue to receive warmest ratings, atheists and Muslims move from cool to neutral

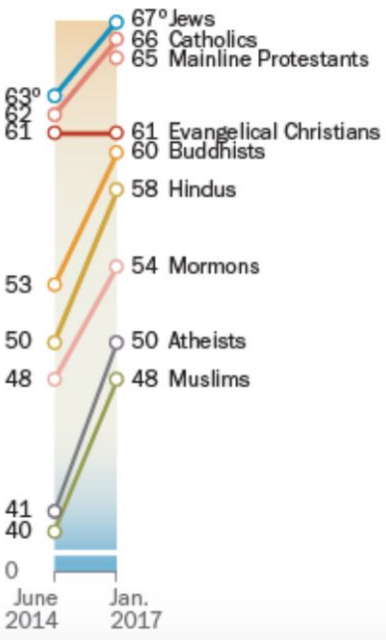
On the heels of a contentious election year in which [partisan politics increasingly divided Americans](), a new Pew Research Center survey finds that when it comes to religion, Americans generally express more positive feelings toward various religious groups today than they did just a few years ago. Asked to rate a variety of groups on a "feeling thermometer" ranging from 0 to 100, U.S. adults give nearly all groups warmer ratings than they did in a June 2014 Pew Research Center survey.

While Americans still feel coolest toward Muslims and atheists, mean ratings for these two groups increased from a somewhat chilly 40 and 41 degrees, respectively, to more neutral ratings of 48 and 50. Jews and Catholics continue to be among the groups that receive the warmest ratings – even warmer than in 2014.

Evangelical Christians, rated relatively warmly at 61 degrees, are the only group for which the mean rating did not change since the question was last asked in 2014. Americans' feelings toward Mormons and Hindus have shifted from relatively neutral places on the thermometer to somewhat warmer ratings of 54 and 58,

Americans feeling warmer toward variety of religious groups

Mean thermometer ratings



I can make community-level variables

I can construct the percent African-American in each community. Here are the distributions of this variable by race:

```
> sub <- ddply(sub, "sampyear", mutate, pct.black = 100*mean(race==2)) # percent
black

> with(sub, by(pct.black, race, summary)) # summary of pct.black by race
race: 1

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00    0.00    5.00   10.40   16.18   89.47
------------------------------------------------------------
race: 2
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.818  18.750  31.250  34.840  45.710 100.000
------------------------------------------------------------
race: 3
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.000   2.326   7.692  12.340  19.350  68.750
```

I will also make a variable that counts how many observations there are for each community:

- *countn* = gives each observation a number from 1 through N

-*countN* = gives each observation the total number of observations from the same cluster

```
> sub <- ddply(sub, "sampyear", mutate,
+              N = length(sampyear),  # number of obs per community
+              n = seq_along(N))      # index each obs within communities
```

(8 people have a community of 1 person only; 36 people have a community of 4 people; 1613 people come from the community fo 1613 people)

Our naïve OLS regression

We add clustered standard errors to account for the fact that people are coming from similar regions

```
> sub <- sub[-which(sub$race == 3), ] # only want to compare blacks vs whites
> sub$white <- sub$race == 1
> sub <- na.omit(sub)

> lm.closeblk <- plm(closeblk ~ white + pct.black + age + year,
+                    data = sub, index = "sampyear", model = "pooling")
> clusterSE(fit = lm.closeblk, cluster.var = "sampyear")

t test of coefficients:

               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept) -66.9535457   8.4013952  -7.9693 1.749e-15 ***
whiteTRUE    -2.0131133   0.0652328 -30.8604 < 2.2e-16 ***
pct.black     0.0076689   0.0016293   4.7069 2.545e-06 ***
age          -0.0099051   0.0010730  -9.2309 < 2.2e-16 ***
year          0.0372860   0.0041926   8.8933 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Our naïve OLS regression

Net of other factors, whites are over 2 categories less close to African-Americans that African-Americans are (to other African-Americans)

```
> clusterSE(fit = lm.closeblk, cluster.var = "sampyear")

t test of coefficients:

               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept) -66.9535457   8.4013952  -7.9693 1.749e-15 ***
whiteTRUE    -2.0131133   0.0652328 -30.8604 < 2.2e-16 ***
pct.black     0.0076689   0.0016293   4.7069 2.545e-06 ***
age          -0.0099051   0.0010730  -9.2309 < 2.2e-16 ***
year          0.0372860   0.0041926   8.8933 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Our naïve OLS regression

Net of other factors, for every percent more people in one's community are African-American, they are 0.0076 points closer to African-Americans

```
> clusterSE(fit = lm.closeblk, cluster.var = "sampyear")

t test of coefficients:

               Estimate  Std. Error  t value  Pr(>|t|)    
(Intercept) -66.9535457   8.4013952  -7.9693 1.749e-15 ***
whiteTRUE    -2.0131133   0.0652328 -30.8604 < 2.2e-16 ***
pct.black     0.0076689   0.0016293   4.7069 2.545e-06 ***
age          -0.0099051   0.0010730  -9.2309 < 2.2e-16 ***
year          0.0372860   0.0041926   8.8933 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Our naïve OLS regression

Looking at the Betas: 1 standard deviation increase in the %age black of one's community, translates into a 0.056 increase in the st. dev. of closeness to Blacks

```
> stdCoef(lm.closeblk) # check standardized coefficients
Standardized Coefficients for lm.closeblk
  whiteTRUE   pct.black         age        year
-0.34921885  0.05651031 -0.08268438  0.09131848
```

Now, the random intercept model

The random intercept model is this:

$$y_{ij} = \alpha + \beta x_1 + u_j + e_i$$

which provides a unique error (u) for each unit *j*, which is just how far a particular unit's intercept is from the overall intercept (α) … and then e_i is all the other sources of error that are not related to the unit-specific average difference from the overall intercept

Start with the “empty” random intercept model

- There are no Xs (hence, no βs) in this model yet
- Also called a “variance components” model

```
> library(lme4)

> nullmodel <- lmer(closeblk ~ (1 | sampyear), data = sub, REML = FALSE)
> summary(nullmodel)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ (1 | sampyear)
   Data: sub

     AIC      BIC   logLik deviance df.resid
 48210.2  48232.2 -24102.1  48204.2    11257

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.3018   0.5494
 Residual             4.0365   2.0091
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.67123    0.02834   200.1
```

Start with the “empty” random intercept model

Rho means that 6.96% of the variance in closeness to African-Americans is between different communities

```
> summary(nullmodel)

     AIC       BIC   logLik deviance df.resid
 48210.2  48232.2 -24102.1  48204.2    11257

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.3018   0.5494
 Residual             4.0365   2.0091
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.67123    0.02834   200.1


> rho(nullmodel)
[1] 0.06957642
```

Calculating rho ...

Here is a function that will calculate rho

```
> # Note that in the output from summary(nullmodel) the Std. Dev. column in the
> # Random Effects section contains what STATA refers to as sigma_u and sigma_e.
> # STATA also reports rho (fraction of variance due to u_i) but this is not given
> # in the R output. We'll want to compute it for the models below, so we can
> # write a function to avoid retyping the same commands multiple times.

> rho <- function(fit){
+   varcor <- VarCorr(fit) # extract the variance components using VarCorr()
+   varcor <- as.data.frame(varcor)[, "sdcor"] # get just the std devs we want
+   sigma_u <- varcor[1] # get sigma_u
+   sigma_e <- varcor[2] # get sigma_e
+   rho <- sigma_u^2 / (sigma_u^2 + sigma_e^2) # compute rho (fraction of variance
due to u_i)
+   rho
+ }

> # For future use the rho function is included in QMSS package
> ?rho
```

Now, the random intercept model

```
> lmer.closeblk1 <- lmer(closeblk ~ white + pct.black + age + year + (1 |
sampyear), data = sub, REML = FALSE)
> summary(lmer.closeblk1)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ white + pct.black + age + year + (1 | sampyear)
   Data: sub

     AIC      BIC   logLik deviance df.resid
 46410.1  46461.4 -23198.1  46396.1    11253

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.06747  0.2598
 Residual             3.54788  1.8836
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept) -66.104237   8.092759   -8.17
whiteTRUE    -2.014115   0.055273  -36.44
pct.black     0.007947   0.001444    5.50
age          -0.009698   0.001041   -9.32
year          0.036854   0.004040    9.12

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.016
pct.black  0.020  0.397
age        0.062 -0.082  0.009
```

Random intercept model

Virtually the same coefficients as OLS:

- White was -2.033 (z=-30), now is -2.035 (z=-36.44)
- %_black was 0.0083 (z=4.7), now is 0.0085 (z=5.50)

```
> summary(lmer.closeblk1)

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.06747  0.2598
 Residual             3.54788  1.8836

Fixed effects:
              Estimate Std. Error t value
(Intercept) -66.104237   8.092759   -8.17
whiteTRUE    -2.014115   0.055273  -36.44
pct.black     0.007947   0.001444    5.50
age          -0.009698   0.001041   -9.32
year          0.036854   0.004040    9.12

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.016
pct.black  0.020  0.397
age        0.062 -0.082  0.009
year      -1.000  0.010 -0.025 -0.067
```

Random intercept model

Now, rho implies that only 1.86% of the variance in closeness to African-Americans is still due to (unobserved) differences between communities

Why is that?

```
> rho(lmer.closeblk1)
[1] 0.01866329
```

Why did rho go down so much?

- Sigma_u drops from 0.575 to 0.259
- Sigma_e drops from 2.005 to 1.88
- We explained a lot with our Xs (a little bit with individual Xs like race and a lot with community Xs like %_black)

```
Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.06747  0.2598
 Residual             3.54788  1.8836
```

What is all this “Random-effects” at bottom?

- *sd(_sampyear)* is how much variation there is in the constant, once every community gets their own intercept
- I.e., the constant has a mean of -63.75 and now a sd of 0.254

```
Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.06747  0.2598
 Residual             3.54788  1.8836
```

What is up with that extreme constant?

- The constant now has a mean of 6.57 and now a sd of 0.259

```
> sub$n.year <- sub$year-1972
> lmer.closeblk1 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 | sampyear),
+                        data = sub, REML = FALSE)
> summary(lmer.closeblk1)

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.06747  0.2598
 Residual             3.54788  1.8836
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.571124   0.146906   44.73
whiteTRUE   -2.014115   0.055273  -36.44
pct.black    0.007947   0.001444    5.50
age         -0.009698   0.001041   -9.32
n.year       0.036854   0.004040    9.12

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.357
pct.black -0.248  0.397
age       -0.252 -0.082  0.009
n.year    -0.853  0.010 -0.025 -0.067
```

Do you need separate slopes (on white) for each community as well?

Now, the random intercept model

The random intercept, random slope model is this:

$$y_{ij} = \alpha + \beta x_1 + u_{j(\alpha)} + u_{j(\beta)} + e_i$$

where all we have done is allowed each community to have its own slope for how x_1 affects y_{ij}, by providing a unique error ($u_{i(\beta)}$) for each unit *j*, which is just how far a particular unit's slope differs from the average slope (β)

Everybody gets their own regression!

For community 1 in 1996, white has a coefficient of -2.2, with a constant of 7.1 (in STATA)

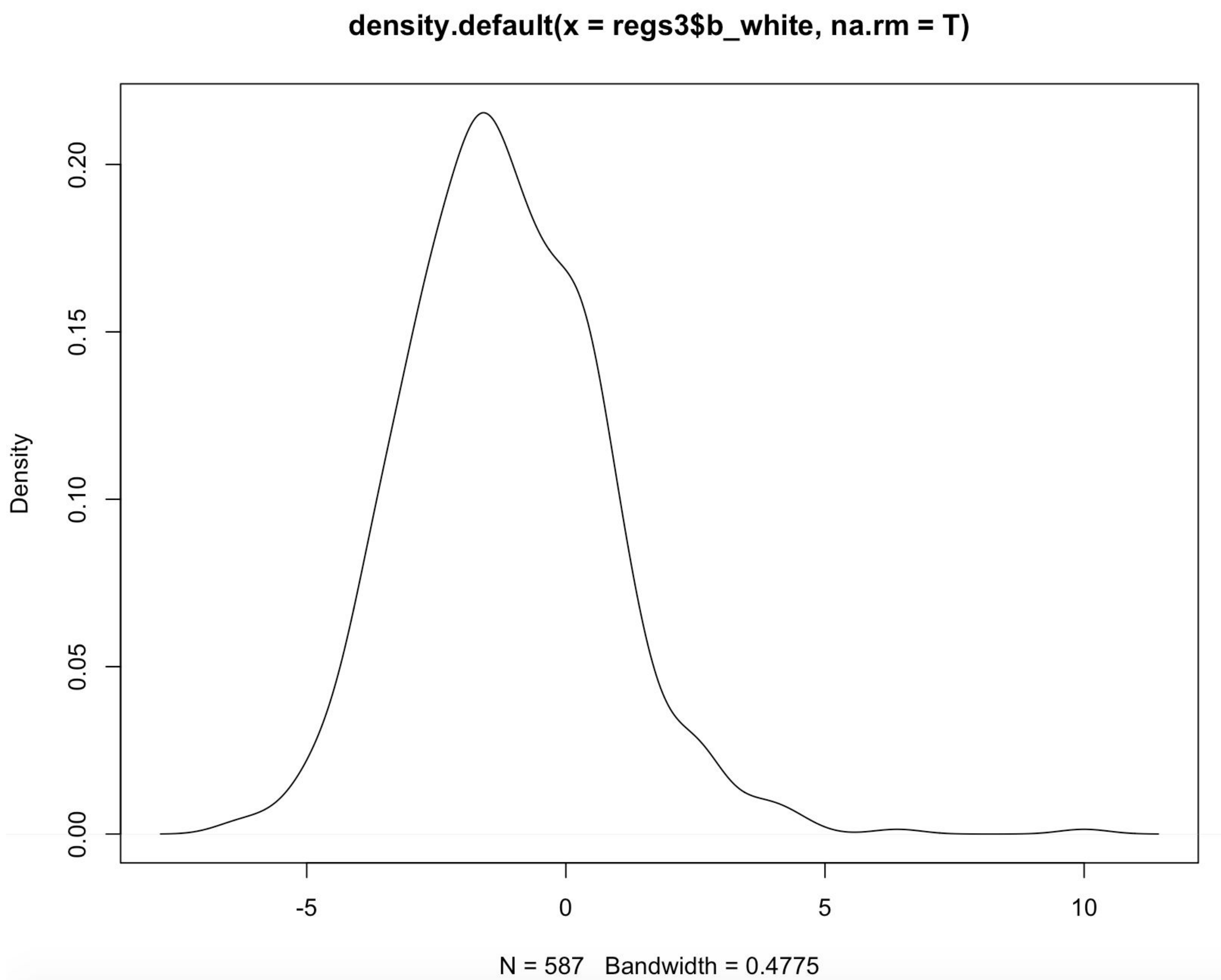
```
. bysort sampyear: reg closeblk white pctblack100 age year

-> sampyear = 11996
note: pctblack100 omitted because of collinearity
note: year omitted because of collinearity

      Source |       SS       df       MS              Number of obs =      51
-------------+------------------------------           F(  2,    48) =    6.75
       Model |  51.3006538     2  25.6503269           Prob > F      =  0.0026
    Residual |  182.346405    48  3.79888344           R-squared     =  0.2196
-------------+------------------------------           Adj R-squared =  0.1870
       Total |  233.647059    50  4.67294118           Root MSE      =  1.9491

------------------------------------------------------------------------------
    closeblk |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |  -2.201945   .5992092    -3.67   0.001    -3.406736   -.9971544
 pctblack100 |  (omitted)
         age |    .001956   .0164679     0.12   0.906     -.031155     .035067
        year |  (omitted)
       _cons |   7.109632   .9121975     7.79   0.000     5.275536    8.943728
------------------------------------------------------------------------------
```

More regressions …

```
> # Run a separate regression for each sampyear, storing intercept, coefficient on
> # white, and the average of the fitted values.
>   # get all unique values of sampyear
> samps <- unique(sub$sampyear)
> regs <- sapply(
+   X = 1:length(samps), # apply the function FUN (below) to each i in 1:length(samps)
+   FUN = function(i){
+     # regression for sampyear i
+     lm <- lm(closeblk ~ white + age, data = sub, sampyear == samps[i])
+     # return intercept, coeff on white, mean of fitted values
+     c(coef(lm)[1:2], mean(lm$fitted))
+ })
>
> rownames(regs) = c("b_const", "b_white", "mean.fitted")
>
> # get the averages over all values of sampyears
> overall <- round(rowMeans(regs, na.rm = T), 2)
> overall
    b_const     b_white mean.fitted
       6.92       -1.86        5.66
```

Each community gets its own regression

Each community has a slope. Here is the distribution.

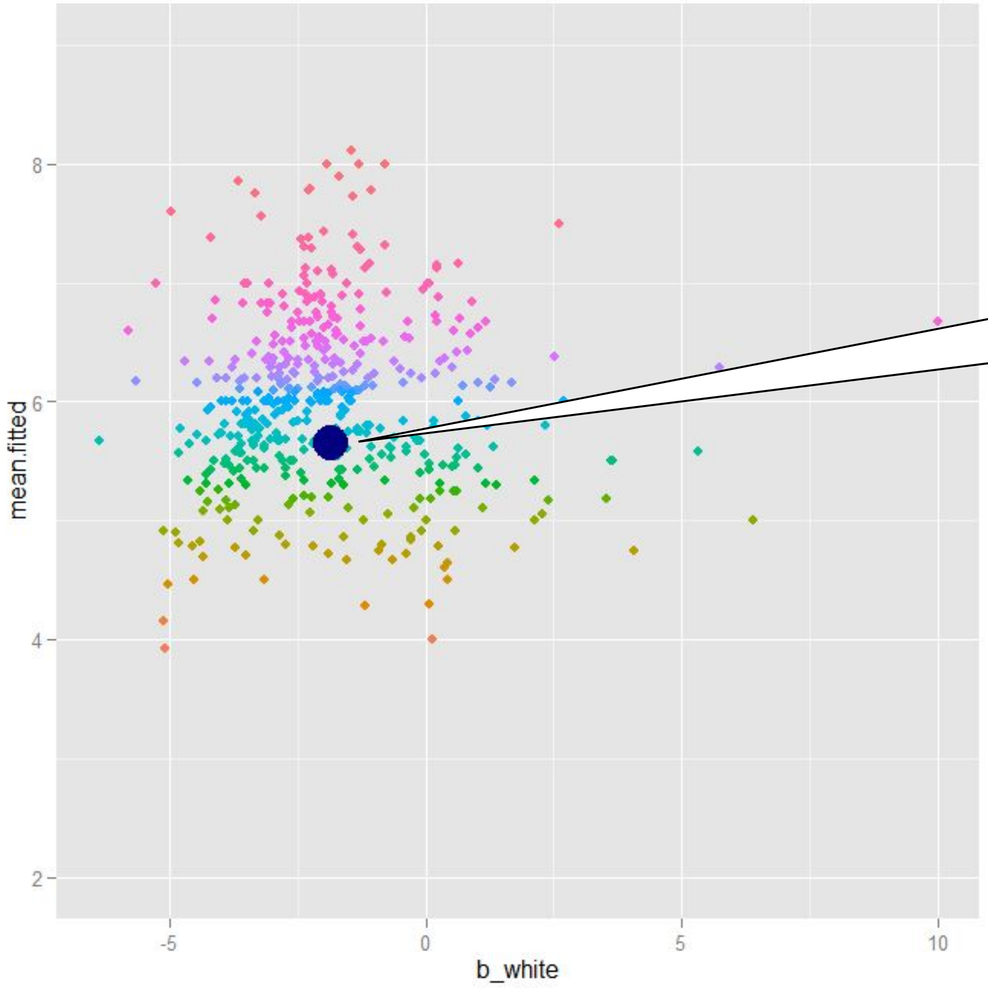


How did I do this graph?

```
# Run a separate regression for each sampyear, storing intercept,
coefficient on
# white, and the average of the fitted values.
# get all unique values of sampyear
samps <- unique(sub$sampyear)
regs <- sapply(
  X = 1:length(samps), # apply the function FUN (below) to each i in
1:length(samps)
  FUN = function(i){
    # regression for sampyear i
    lm <- lm(closeblk ~ white + age, data = sub, sampyear == samps[i])
    # return intercept, coeff on white, mean of fitted values
    c(coef(lm)[1:2], mean(lm$fitted))
  })

rownames(regs) = c("b_const", "b_white", "mean.fitted")

regs2 = as.data.frame(regs)
head(regs)
regs3 = t(regs)
regs3 = as.data.frame(regs3)

plot(density(regs3$b_white, na.rm=T))
```

Or all at once (for the 1st 71 communities)

Fair bit of variability in terms of size of "white" coeff. relative to predicted value of closeness to Blacks (Y axis)



The overall average slope (-2.12) and predicted value of Y (5.35)

How did I do this graph?

```
> # plot coef on white vs. mean fitted for each sampyear
> regs.df <- data.frame(t(regs))
> no_legend <- theme(legend.position = "none")
> g_white_vs_fitted <- ggplot(regs.df, aes(x = b_white, y = mean.fitted,
+                                          color = factor(mean.fitted)))
> g_white_vs_fitted <- g_white_vs_fitted + no_legend
> g_white_vs_fitted + geom_point()
Warning message:
Removed 335 rows containing missing values (geom_point).
>
> # add a larger point showing the overall means
> mean_point <- annotate("point", x = overall[2], y = overall[3], color =
"navyblue", size = 8)
> g_white_vs_fitted + geom_point() + mean_point
Warning message:
Removed 335 rows containing missing values (geom_point).
```

Random intercepts + *random slopes*

A lot going on here. 1) By typing (1 + *white)* before the "| sampyear:" this is like saying, allow *white* to have a random coefficient (slope) for each community

```
> lmer.closeblk2 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 + white |
sampyear), data = sub, REML = FALSE)
> summary(lmer.closeblk2)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.3371   0.5806
          whiteTRUE   0.3589   0.5991   -0.91
 Residual             3.5103   1.8736
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.517456   0.149651   43.55
whiteTRUE   -1.992668   0.064631  -30.83
pct.black    0.008849   0.001472    6.01
age         -0.009742   0.001039   -9.37
n.year       0.037715   0.004010    9.41

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.413
pct.black -0.253  0.364
age       -0.246 -0.073  0.010
```

Random intercepts + *random slopes*

2) The covariance between each community's intercept and slope can be whatever; otherwise, they will be defaulted to making the covariance zero

```
> lmer.closeblk2 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 + white |
sampyear), data = sub, REML = FALSE)
> summary(lmer.closeblk2)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.3371   0.5806
          whiteTRUE   0.3589   0.5991   -0.91
 Residual             3.5103   1.8736
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.517456   0.149651   43.55
whiteTRUE   -1.992668   0.064631  -30.83
pct.black    0.008849   0.001472    6.01
age         -0.009742   0.001039   -9.37
n.year       0.037715   0.004010    9.41

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.413
pct.black -0.253  0.364
age       -0.246 -0.073  0.010
```

Random intercepts + *random slopes*

Interpretation of constant. For a 0 year old Black person in a 0% black community in 1972, on average, they have a 6.51 score, but there is substantial variation around that mean, with a st dev of .5806

```
> lmer.closeblk2 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 + white |
sampyear), data = sub, REML = FALSE)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.3371   0.5806
          whiteTRUE   0.3589   0.5991   -0.91
 Residual             3.5103   1.8736
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.517456   0.149651   43.55
whiteTRUE   -1.992668   0.064631  -30.83
pct.black    0.008849   0.001472    6.01
age         -0.009742   0.001039   -9.37
n.year       0.037715   0.004010    9.41

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.413
pct.black -0.253  0.364
```

Random intercepts + *random slopes*

Interpretation of slope. On average, white people are -1.99 categories less close to Blacks (than Blacks), but there is substantial variation around that average slope, with a st dev of .599

```
> lmer.closeblk2 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 + white |
sampyear), data = sub, REML = FALSE)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.3371   0.5806
          whiteTRUE   0.3589   0.5991   -0.91
 Residual             3.5103   1.8736
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.517456   0.149651   43.55
whiteTRUE   -1.992668   0.064631  -30.83
pct.black    0.008849   0.001472    6.01
age         -0.009742   0.001039   -9.37
n.year       0.037715   0.004010    9.41

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.413
pct.black -0.253  0.364
```

Random intercepts + *random slopes*

Interpretation of corr(white,_cons): There is a very high negative correlation between the constant and the slope for each community (ρ=-.91).

```
> lmer.closeblk2 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 + white |
sampyear), data = sub, REML = FALSE)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.3371   0.5806
          whiteTRUE   0.3589   0.5991   -0.91
 Residual             3.5103   1.8736
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.517456   0.149651   43.55
whiteTRUE   -1.992668   0.064631  -30.83
pct.black    0.008849   0.001472    6.01
age         -0.009742   0.001039   -9.37
n.year       0.037715   0.004010    9.41

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.413
pct.black -0.253  0.364
age       -0.246 -0.073  0.010
n.year    -0.832  0.011 -0.027 -0.067
```

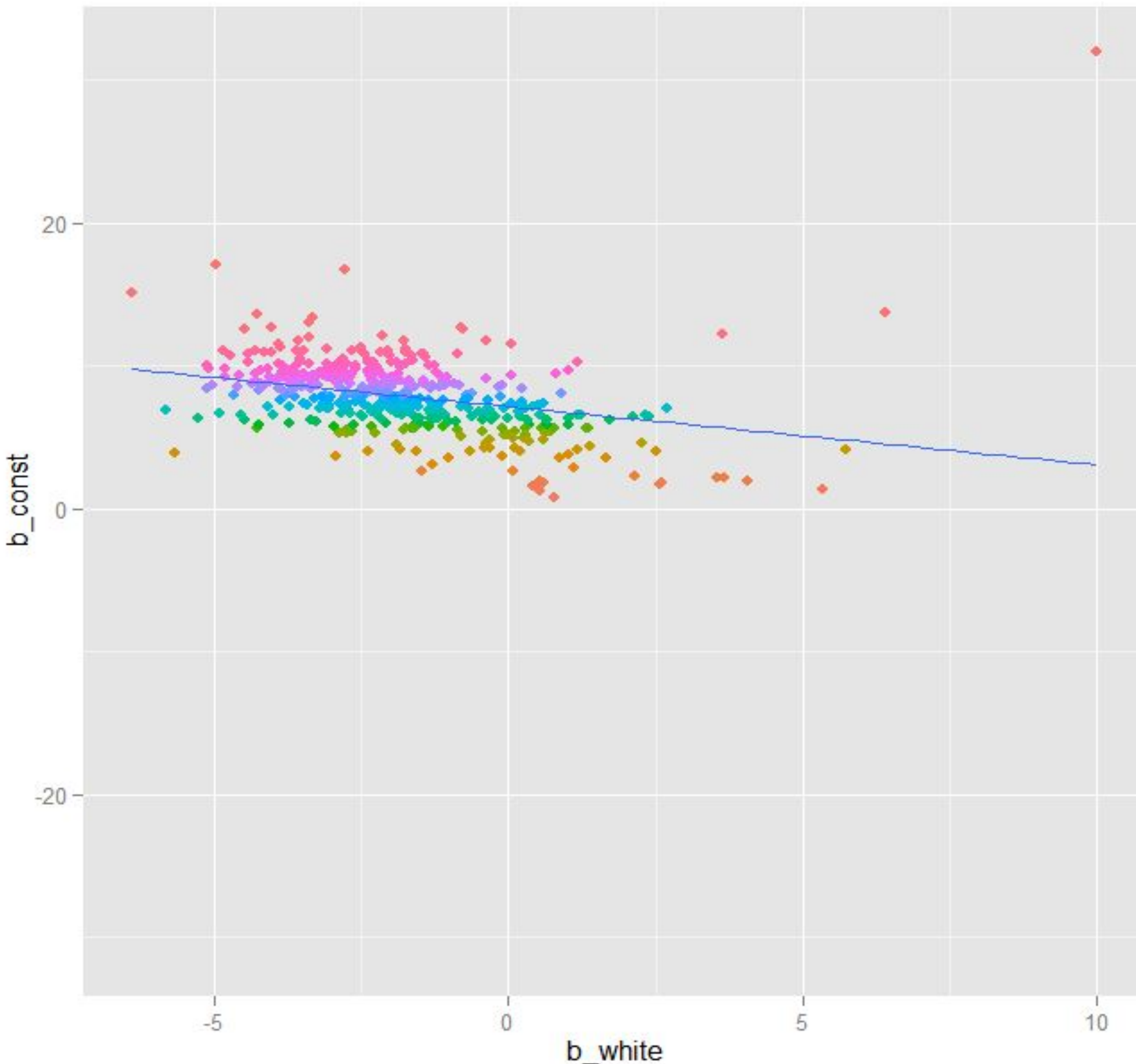
Random intercepts + *random slopes*

Or … corr(white,_cons): Communities with above average closeness to Blacks (in their constant) have below average values on the slope for “white” – and vice versa

```
> lmer.closeblk2 <- lmer(closeblk ~ white + pct.black + age + n.year + (1 + white
| sampyear), data = sub, REML = FALSE)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.3371   0.5806
          whiteTRUE   0.3589   0.5991   -0.91
 Residual             3.5103   1.8736
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
             Estimate Std. Error t value
(Intercept)  6.517456   0.149651   43.55
whiteTRUE   -1.992668   0.064631  -30.83
pct.black    0.008849   0.001472    6.01
age         -0.009742   0.001039   -9.37
n.year       0.037715   0.004010    9.41

Correlation of Fixed Effects:
          (Intr) whTRUE pct.bl age
whiteTRUE -0.413
pct.black -0.253  0.364
age       -0.246 -0.073  0.010
n.year    -0.832  0.011 -0.027 -0.067
```

Random intercepts + *random slopes*

Or … corr(white,_cons) again: In graph form



How did I do this graph?

```
# plot coef on white vs. intercept for each sampyear
g_white_vs_const <- ggplot(regs.df, aes(x = b_white, y = b_const)) +
no_legend

g_white_vs_const + geom_point(aes(color = factor(b_const))) +
stat_smooth(method = "lm", se = F)
```

Are random slopes necessary?

We can determine if random slopes better fit our data by comparing the log-likelihood (LL) from the random intercept model (=-23198) with the LL from the random intercept + random slopes model (LL=-23187)

```
> # Likelihood ratio test via anova()
> anova(lmer.closeblk1, lmer.closeblk2) # look at Chisq stat and p-value
Data: sub
Models:
lmer.closeblk1: closeblk ~ white + pct.black + age + n.year + (1 | sampyear)
lmer.closeblk2: closeblk ~ white + pct.black + age + n.year + (1 + white | sampyear)
               Df   AIC   BIC logLik deviance  Chisq Chi Df Pr(>Chisq)    
lmer.closeblk1  7 46410 46461 -23198    46396                             
lmer.closeblk2  9 46392 46458 -23187    46374 21.827      2  1.821e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Are random slopes necessary?

The random intercept + random slopes model provides a superior fit to the data than the random intercepts model (p<.05)

```
> # Likelihood ratio test via anova()
> anova(lmer.closeblk1, lmer.closeblk2) # look at Chisq stat and p-value
Data: sub
Models:
lmer.closeblk1: closeblk ~ white + pct.black + age + n.year + (1 | sampyear)
lmer.closeblk2: closeblk ~ white + pct.black + age + n.year + (1 + white | sampyear)
               Df   AIC   BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer.closeblk1  7 46410 46461 -23198    46396
lmer.closeblk2  9 46392 46458 -23187    46374 21.827      2  1.821e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Are random slopes necessary?

We could do the test by hand, too

```
> # Or we can do the likelihood ratio test manually to get a sense for how the
> # computation is done
>   # the Chisq statistic is computed as (-2)*loglikelihood(model1) +
2*loglikelihood(model2)
>   # or equivalently as deviance(model1) - deviance(model2)
> chisq <- deviance(lmer.closeblk1) - deviance(lmer.closeblk2)
> chisq # should be the same (up to rouding) as Chisq from anova(lmer.closeblk1,
lmer.closeblk2)
[1] 21.82741
> df <- df.residual(lmer.closeblk1) - df.residual(lmer.closeblk2)
> df # should be the same as Df from anova(lmer.closeblk1, lmer.closeblk2)
[1] 2
> pval <- pchisq(chisq, df, lower.tail = FALSE)
> pval # should be the same (up to rounding) as Pr(>Chisq) from anova(lmer.closeblk1,
lmer.closeblk2)
[1] 1.820695e-05
```

Adding cross-level interactions

- Often we think that how someone's X will affect their Y is dependent on the context they are in
- This suggests an interaction of Level 1 (individuals) with Level 2 (communities)
- In our case: Does the closeness a white person feels for African-Americans depend upon how many African-Americans live around them?

Context matters

```
> lmer.closeblk3 <- lmer(closeblk ~ white*pct.black + age + n.year + (1 + white |
sampyear),
+                        data = sub, REML = FALSE)
> summary(lmer.closeblk3)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ white * pct.black + age + n.year + (1 + white | sampyear)
   Data: sub

     AIC      BIC   logLik deviance df.resid
 46387.7  46461.0 -23183.9  46367.7    11250

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.2968   0.5447
          whiteTRUE   0.3382   0.5816   -0.90
 Residual             3.5123   1.8741
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
                    Estimate Std. Error t value
(Intercept)         6.752638   0.171756   39.32
whiteTRUE          -2.228910   0.109929  -20.28
pct.black           0.001399   0.003149    0.44
age                -0.009667   0.001039   -9.30
n.year              0.036967   0.003992    9.26
whiteTRUE:pct.black  0.009306   0.003545    2.63

Correlation of Fixed Effects:
            (Intr) whTRUE pct.bl age    n.year
whiteTRUE   -0.614
pct.black   -0.546  0.821
age         -0.201 -0.064 -0.019
n.year      -0.746  0.049  0.034 -0.069
```

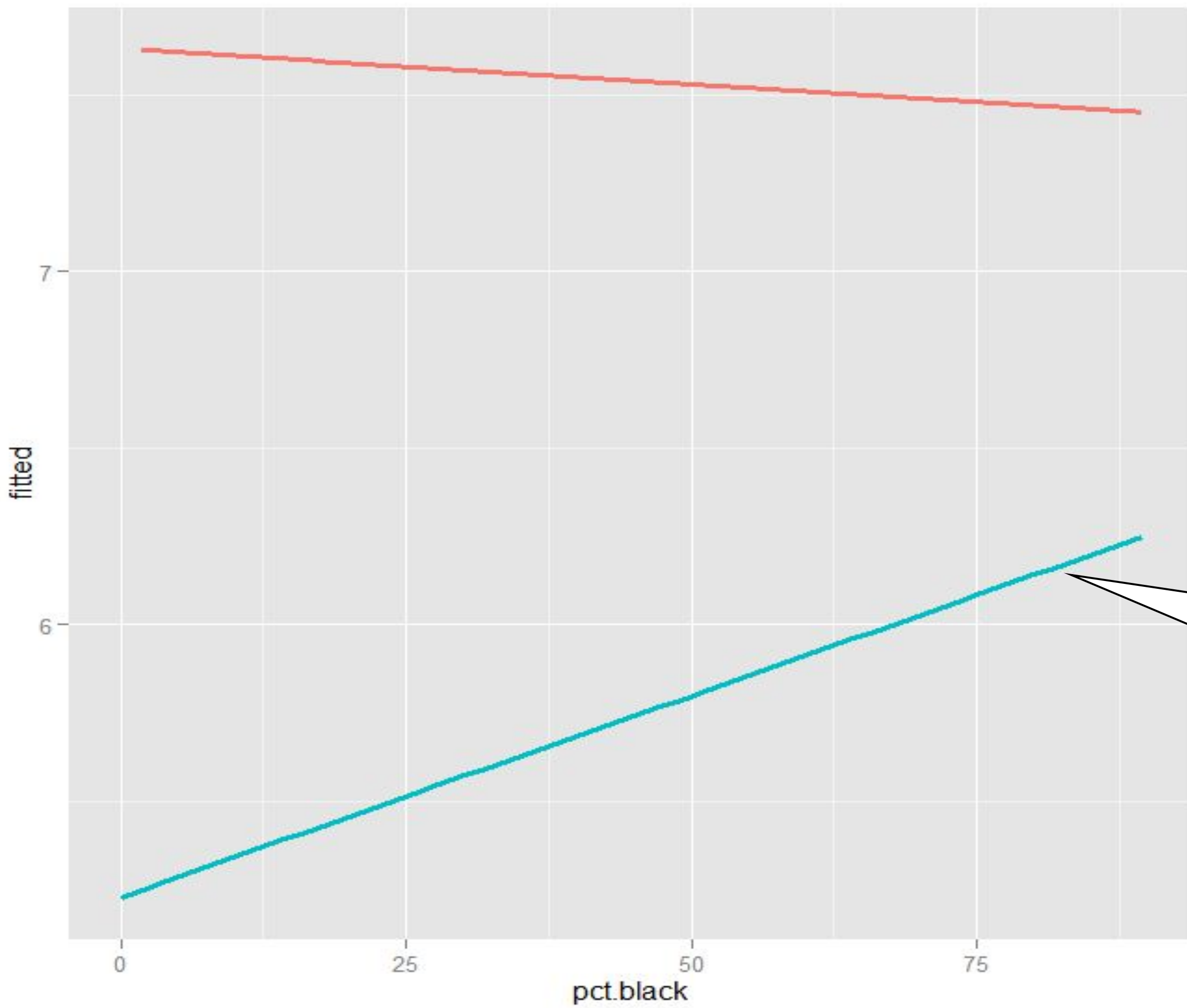
Context matters …

For every percent more blacks a white lives near, s/he moves up 0.0099 more categories (than blacks) of feeling close to African-Americans

```
> summary(lmer.closeblk3)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ white * pct.black + age + n.year + (1 + white | sampyear)
   Data: sub

     AIC      BIC   logLik deviance df.resid
 46387.7  46461.0 -23183.9  46367.7    11250

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.2968   0.5447
          whiteTRUE   0.3382   0.5816   -0.90
 Residual             3.5123   1.8741
Number of obs: 11260, groups:  sampyear, 792

Fixed effects:
                     Estimate Std. Error t value
(Intercept)          6.752638   0.171756   39.32
whiteTRUE           -2.228910   0.109929  -20.28
pct.black            0.001399   0.003149    0.44
age                 -0.009667   0.001039   -9.30
n.year               0.036967   0.003992    9.26
whiteTRUE:pct.black  0.009306   0.003545    2.63

Correlation of Fixed Effects:
```

Context matters

That looks like this …



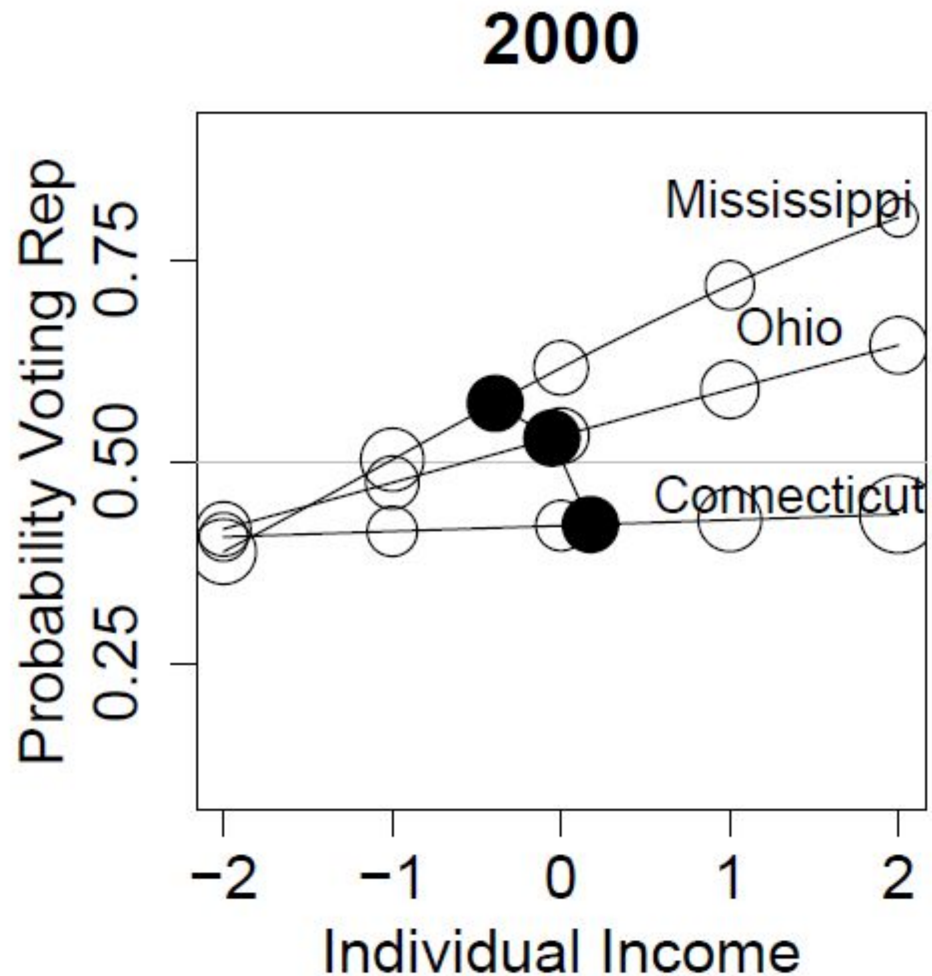
The white slope goes up much faster than the African-American slope does …

How did I do this graph?

```
# visualize the different slopes
plotdata <- data.frame(fitted = fitted(lmer.closeblk3),
                       pct.black = sub$pct.black,
                       white = sub$white)

g_closeblk <- ggplot(plotdata, aes(x = pct.black, y = fitted, group =
white, color = white))
g_closeblk + stat_smooth(method = "lm", se = F, size = 1.25) +
no_legend
```

Remember: Red state, blue state ...

- Here is what it looks like if every state has its own slope



One final thing …

If someone is interviewed by an African-American interviewer, they claim to be 0.61 categories closer to blacks, net of other factors (in STATA)

```
. tab intethn, gen(int)
. rename  int2 int_black
. xtmixed closeblk white pctblack100  whitexpctblack100 age nyear  int_black || sampyear:
white , cov(unstr)

Mixed-effects REML regression                    Number of obs      =      4773
Group variable: sampyear                         Number of groups   =       314
                                                 Obs per group: min =         2
                                                                avg =      15.2
                                                                max =        80
                                                 Wald chi2(6)       =    591.65
Log restricted-likelihood =  -9736.207           Prob > chi2        =    0.0000
------------------------------------------------------------------------------
    closeblk |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |  -2.205772   .1747028   -12.63   0.000    -2.548183   -1.863361
 pctblack100 |  -.0076724   .0053146    -1.44   0.149    -.0180887     .002744
whitexpc~100 |   .0181619   .0058485     3.11   0.002     .0066991    .0296246
         age |  -.0064543   .0015609    -4.14   0.000    -.0095136   -.0033951
       nyear |   .0000418   .0137752     0.00   0.998     -.026957    .0270406
   int_black |   .6180237   .0886897     6.97   0.000     .4441951    .7918522
       _cons |    7.84273   .5157716    15.21   0.000     6.831837    8.853624
------------------------------------------------------------------------------
```

2. Two-stage multilevel models

Two-stage regressions

- Stage 1: Run individual regressions for each unit
- Stage 2: Take the coefficient of interest from Stage 1 and predict it using unit-level Xs

(The only issue is correcting the standard errors in Stage 2)

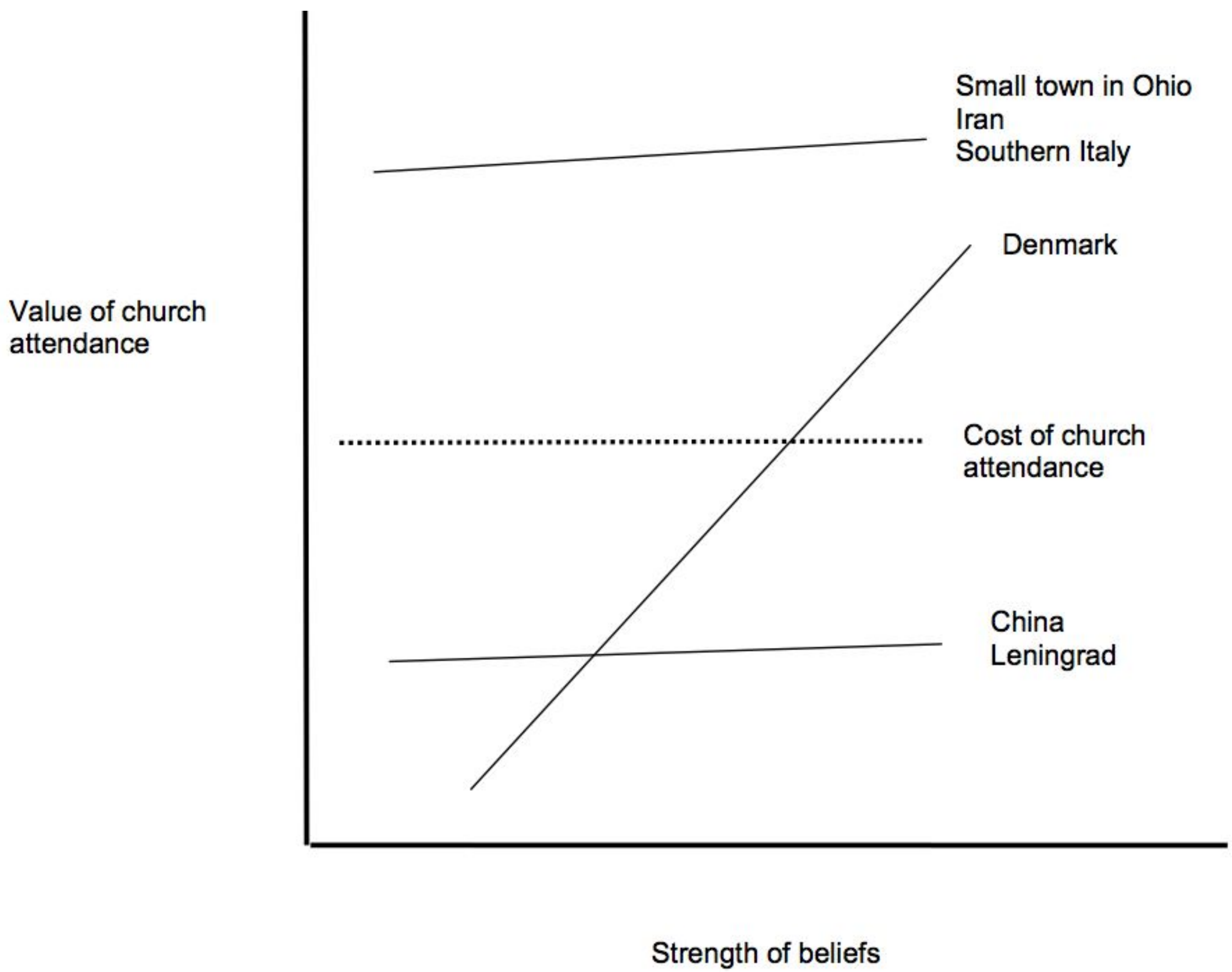
An example

Typically, religion scholars think that there is a constant relationship between religious belief and religious attendance in all societies, where high belief correlates with high attendance

But that may not be true.

Huber, John. "Religious belief, religious participation, and social policy attitudes across countries." *Annual Meetings of the Midwest Political Science Association, Chicago*. 2005.

What if it looked like this?



An example

Stage 1: Run individual regressions of belief on attendance for each country (with n ~ 2000, and get probit coefficients on belief in hell, for instance, ranging from 0.07 in Peru to 2.10 in Denmark. The mean = 0.94 with a standard deviation of .44)

- Stage 2: Take the coefficient of interest (on belief predicting attendance) from Stage 1 and predict it using unit-level Xs ...

Stage 2 regressions

	(1) Believe hell	(2) Believe heaven	(3) God Important	(4) Very spiritual	(5) Believe hell	(6) Very spiritual
GDP (ln)	0.16+ (0.09)	0.14+ (0.08)	0.22* (0.09)	0.25** (0.10)	--	--
Ex-communist	0.48** (0.13)	0.39** (0.11)	0.35** (0.13)	0.47** (0.14)	0.42** (0.13)	0.40** (0.14)
Corruption	-0.11* (0.05)	-0.04 (0.04)	-0.07 (0.06)	-0.13* (0.06)	-0.14** (0.05)	-0.18** (0.06)
State religion (1970)	-0.04 (0.14)	-0.01 (0.12)	-0.02 (0.15)	0.07 (0.16)	-0.05 (0.15)	0.07 (0.17)
Religious pluralism	0.18 (0.34)	0.47 (0.28)	0.70* (0.36)	0.84* (0.38)	0.09 (0.35)	0.70+ (0.40)
Law and Order	0.05 (0.05)	0.07+ (0.04)	0.05 (0.06)	0.07 (0.06)	0.09* (0.05)	0.14** (0.05)
Pct. Orthodox	-0.21 (0.24)	-0.34+ (0.19)	-0.14 (0.24)	-0.12 (0.25)	-0.22 (0.24)	-0.12 (0.27)
Pct. Protestants	0.07 (0.24)	-0.25 (0.21)	0.33 (0.26)	0.34 (0.27)	0.03 (0.25)	0.27 (0.28)
Pct. Muslims	1.14** (0.30)	1.08** (0.25)	0.25 (0.29)	0.69* (0.30)	1.03** (0.30)	0.52 (0.32)
Constant	-0.71** (0.89)	-0.73 (0.76)	-1.48 (0.92)	-1.69 (0.95)	0.76* (0.39)	0.60 (0.44)
N	51	51	51	51	52	51
Adj. R^2	.49	.51	.51	.63	.46	.58

- Being an ex-communist country greatly strengthens the relationship between belief and attendance

- Being a corrupt country has the opposite effect

Final thoughts on this model

- Even Gelman recommends this as a preliminary model
- Standard errors?

Other extensions of this approach to other problems?

Another example

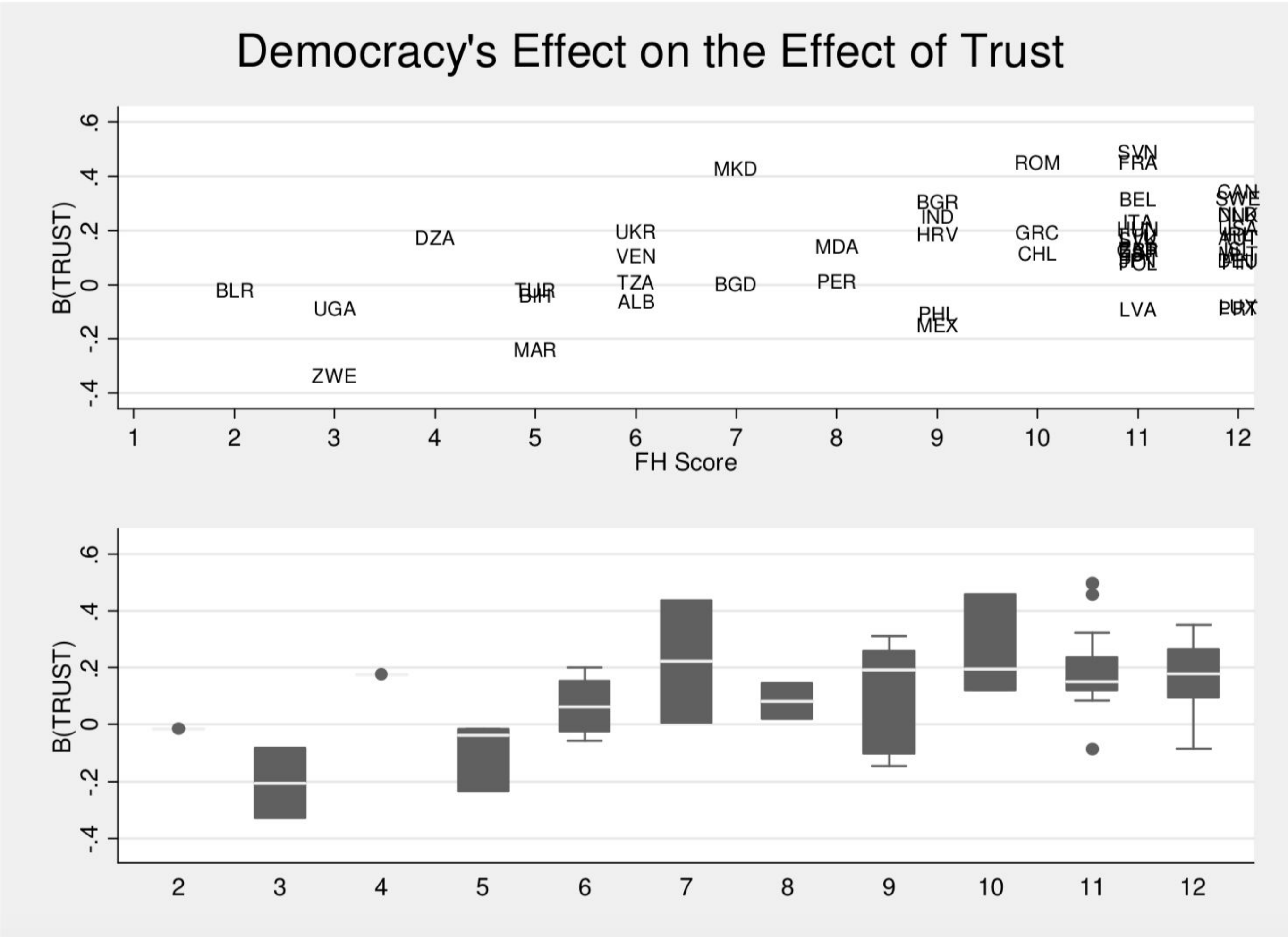
The Democratic Utility of Trust: A Cross-National Analysis

Amaney Jamal and Irfan Nooruddin[1]

The Center for the Study of Democratic Politics

October 5th, 2006

Stage 2 regressions



- Generalized trust's effect on supporting democracy appears to depend in part on how democratic the society already is

Stage 2 regressions

Table 2. Multilevel Model of Democratic Support

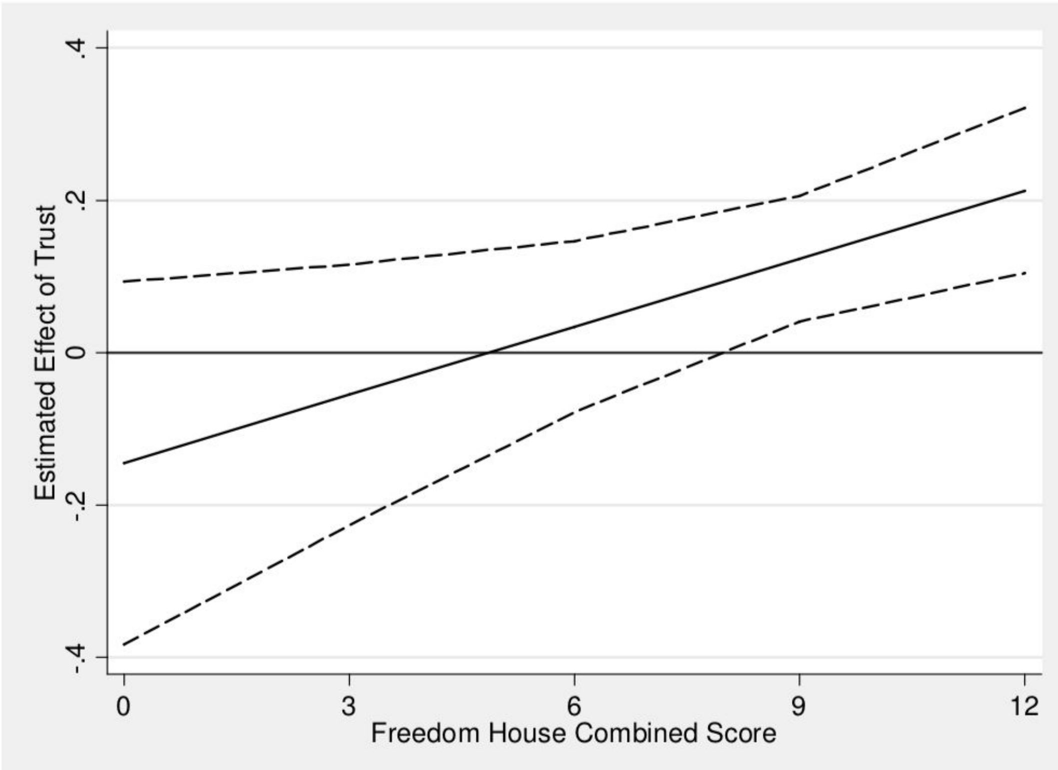
	Model 1 Microlevel Only	Model 2 Country-Level Intercept Effects	Model 3 Full Model with Interaction Effects
Intercept	0.53** (0.06)	0.53** (0.06)	0.54** (0.06)
Individual-Level			
Female	-0.04* (0.02)	-0.04** (0.02)	-0.04** (0.02)
Age	0.001* (0.001)	0.002* (0.001)	0.002** (0.001)
Married	0.04** (0.01)	0.04** (0.01)	0.04** (0.01)
Parent	-0.04** (0.02)	-0.04** (0.02)	-0.04** (0.02)
Less than High School Education	-0.16** (0.03)	-0.16** (0.03)	-0.16** (0.03)
Employed	0.04** (0.01)	0.04** (0.01)	0.03** (0.01)
Religiosity	0.04** (0.01)	0.04** (0.01)	0.04** (0.01)
Teach Children Tolerance	0.04* (0.02)	0.04* (0.02)	0.04** (0.02)
Strong Leader is Good	-0.28** (0.02)	-0.03** (0.02)	-0.28** (0.02)
Traditionalism Scale	0.06** (0.02)	0.06** (0.02)	0.06** (0.02)
Member in Association	0.05** (0.02)	0.05** (0.02)	0.05** (0.02)
Trust	0.14** (0.03)	0.14** (0.03)	0.11** (0.03)
DEMOCRACY			0.03* (0.01)
GDP PER CAPITA			-0.01 (0.02)
RELIGIOSITY			0.01 (0.13)
TOLERANCE			0.11 (0.22)
STRONG LEADER			0.04 (0.05)
TRADITIONALISM			0.11 (0.7)
Country-Level Intercept Effects			
Democracy		-0.03 (0.02)	-0.04** (0.02)
GDP per capita (Log)		0.01 (0.04)	0.01 (0.04)
Religiosity (Mean)		-0.07 (0.23)	-0.07 (0.23)
Tolerance (Mean)		0.46 (0.34)	0.42 (0.36)
Strong Leader (Mean)		-0.18 (0.11)	-0.19 (0.11)
Traditionalism (Mean)		0.08 (0.11)	0.02 (0.11)
Variance Components: Remaining between-country variance)			
Intercept	0.078**	0.072**	0.069**
Percent explained		7.7	11.5
Trust	0.032**	0.032**	0.019**
Percent explained			40.6

Note : Entries are restricted maximum likelihood unstandardized coefficients estimated with HLM 6.0.
See 'Data and Variables' section for details of coding. N = 33858.

- In Model 3, there is a cross-level interaction between trust (at the individual level) and the democracy score of the country. The interaction effect is 0.03*.

Stage 2 regressions

Figure 2. Democracy's Effect on the Effect of Trust



- "Figure 2 plots the estimated effect of *Trust* on *Support for Democracy* at different levels of the Freedom House score. We also plot the estimated 95% confidence intervals. Given that most of the countries in the World Values Survey are relatively free, it is to be expected that the effect is more precisely estimated in the upper range of the X-axis. Substantively, Figure 2 tells us the following: In free societies, social trust bolsters support for democratic government, but it appears to have the opposite effect in unfree societies. And in societies in the middle-to-low range of the Freedom House scale, social trust has a statistically insignificant effect on democratic attitudes. Quite clearly the effect of being trusting on support for democracy is greater in freer societies."

Back to “Slopes as outcomes”

Recall our earlier example. A neighborhood characteristic (% African-American) moderates the relationship between between being white and feeling close to African-Americans. Higher concentrations of African-Americans generate a steeper slope for whites predicting closeness.

```
> summary(lmer.closeblk3)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ white * pct.black + age + n.year + (1 + white | sampyear)
   Data: sub

Fixed effects:
                    Estimate Std. Error t value
(Intercept)         6.752638   0.171756   39.32
whiteTRUE          -2.228910   0.109929  -20.28
pct.black           0.001399   0.003149    0.44
age                -0.009667   0.001039   -9.30
n.year              0.036967   0.003992    9.26
whiteTRUE:pct.black  0.009306   0.003545    2.63
```

BTW, Grand-mean centering or group-mean centering?

Kelley, Jonathan, et al. "Group-mean-centering independent variables in multi-level models is dangerous." *Quality & Quantity* 51.1 (2017): 261-283.

Bell, Andrew, Kelvyn Jones, and Malcolm Fairbrother. "Understanding and misunderstanding group mean centering: a commentary on Kelley et al.'s dangerous practice." *Quality & quantity* 52.5 (2018): 2031-2036.

The “Within-Between” Model

The "Within-Between" Model

What if we could get the best of both worlds? Fixed effects + random effects? You can. All you need to do is include unit-level means (at Level 2) plus deviations from those means at the individual level (Level 1)

1st- Simple OLS

Overall, whites register as 1.38 points lower on closeness to African-Americans

```
> summary(lm1)

Call:
lm(formula = closeblk ~ white, data = sub7)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.74581    0.03839  175.71   <2e-16 ***
whiteTRUE   -1.38432    0.04357  -31.78   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.01 on 12270 degrees of freedom
Multiple R-squared:  0.07603, Adjusted R-squared:  0.07596
F-statistic:  1010 on 1 and 12270 DF,  p-value: < 2.2e-16


> AIC(lm1)
[1] 51969.27
```

2nd- Fixed Effects

Within the same community, whites register as 1.24 points lower on closeness to African-Americans

```
> summary(lm2)

Call:
lm(formula = closeblk ~ white + as.factor(sampyear), data = sub7)

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  6.3611857  0.2497357  25.472  < 2e-16 ***
whiteTRUE                   -1.2431860  0.0477057 -26.059  < 2e-16 ***
as.factor(sampyear)11998     0.0608722  0.3088452   0.197 0.843756
Etc. etc.

Residual standard error: 1.97 on 11479 degrees of freedom
Multiple R-squared:  0.1696,  Adjusted R-squared:  0.1123
F-statistic:  2.96 on 792 and 11479 DF,  p-value: < 2.2e-16

> AIC(lm2)
[1] 52240.92
```

3rd- Equivalent of Fixed Effects in MLM

Within the same community, whites register as 1.24 points lower on closeness to African-Americans

```
> summary(lm3)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + (1 | sampyear)   Data: sub7

     AIC      BIC   logLik deviance df.resid
 52075.2  52104.9 -26033.6  52067.2    12268

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.2883   0.5369
 Residual             3.8920   1.9728
Number of obs: 12272, groups:  sampyear, 792

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.64768    0.02726  207.20
d.white     -1.24319    0.04776  -26.03

Correlation of Fixed Effects:
        (Intr)
d.white 0.000
```

4th- Within-Between

Within the same community, whites register as 1.24 points lower on closeness to African-Americans. Net of race, a person who lives in a 0% white community vs. someone who lives in a 100% white community leads to a 2 point drop in closeness to African-Americans

```
> summary(lm4)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + white.m + (1 | sampyear)   Data: sub7

     AIC      BIC   logLik deviance df.resid
 51828.2  51865.3 -25909.1  51818.2    12267

Random effects:
 Groups   Name        Variance Std.Dev.
 sampyear (Intercept) 0.1491   0.3861
 Residual             3.8815   1.9701
Number of obs: 12272, groups:  sampyear, 792

Fixed effects:
            Estimate Std. Error t value
(Intercept)   7.2877     0.1002   72.74
d.white      -1.2432     0.0477  -26.06
white.m      -2.0696     0.1232  -16.79
```

```
Correlation of Fixed Effects:
        (Intr) d.whit
d.white  0.000
white.m -0.972  0.000
```

5th- Within-Between + Random Slopes

All of the previous results, but now, we see that the "within community" slope on white is on average -2.05, but the standard deviation is 0.87, suggesting meaningful variation around that slope

```
> summary(lm5)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + white.m + (1 + d.white | sampyear)  Data: sub7

     AIC      BIC   logLik deviance df.resid
 51711.0  51762.9 -25848.5  51697.0    12265

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.1566   0.3957
          d.white     0.7649   0.8746   -0.66
 Residual             3.7723   1.9422
Number of obs: 12272, groups:  sampyear, 792

Fixed effects:
            Estimate Std. Error t value
(Intercept)  7.27914    0.09807   74.23
d.white     -1.25443    0.06338  -19.79
white.m     -2.05578    0.12059  -17.05
```

5th- Within-Between + Random Slopes

This model (AIC=51,711) fits 100+ AIC points better than the previous one (AIC= 51,828.2)

```
> summary(lm5)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + white.m + (1 + d.white | sampyear)  Data: sub7

     AIC       BIC   logLik deviance df.resid
 51711.0  51762.9 -25848.5  51697.0    12265
Etc., etc.

Fixed effects:
            Estimate Std. Error t value
(Intercept)  7.27914    0.09807   74.23
d.white     -1.25443    0.06338  -19.79
white.m     -2.05578    0.12059  -17.05

Correlation of Fixed Effects:
        (Intr) d.whit
d.white -0.170
white.m -0.971  0.117
```

This is helpful to do MLM interactions

The Stata Journal (2013)
13, Number 1, pp. 65–76

Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models

Reinhard Schunck
Department of Sociology
University of Bielefeld
Bielefeld, Germany
reinhard.schunck@uni-bielefeld.de

Abstract. Correlated random-effects (Mundlak, 1978, *Econometrica* 46: 69–85; Wooldridge, 2010, *Econometric Analysis of Cross Section and Panel Data* [MIT Press]) and hybrid models (Allison, 2009, *Fixed Effects Regression Models* [Sage]) are attractive alternatives to standard random-effects and fixed-effects models because they provide within estimates of level 1 variables and allow for the inclusion of level 2 variables. I discuss these models, give estimation examples, and address

This is also helpful for MLM interactions

Why You Should Always Include a Random Slope for the Lower-Level Variable Involved in a Cross-Level Interaction

Public | 0 | ...

Contributors: Jan Paul Heisig, Merlin Schaeffer

Date created: 2018-01-11 08:36 AM | Last Updated: 2018-07-16 06:01 AM

Category: Project

Description: Mixed effects multilevel models are often used to investigate cross-level interactions, a specific type of context effect that may be understood as an upper-level variable moderating the association between a lower-level predictor and the outcome. We argue that multilevel models involving cross-level interactions should always include random slopes on the lower-level components of those interactions. Failure to do so will usually result in severely anti-conservative statistical inference. Monte Carlo simulations and illustrative empirical analyses highlight the practical relevance of the issue. Using European Social Survey data, we examine a total 30 cross-level interactions. Introducing a random slope term on the lower-level variable involved in a cross-level interaction, reduces the absolute t-ratio by 31% or more in three quarters of cases, with an average reduction of 42%. Many practitioners seem to be unaware of these issues. Roughly half of the cross-level interaction estimates published in the European Sociological Review between 2011 and 2016 are based on models that omit the crucial random slope term. Detailed analysis of the associated test statistics suggests that many of the estimates would not meet conventional standards of statistical significance if estimated using the correct specification. This raises the question how much robust evidence of cross-level interactions sociology has actually produced over the past decades.

6th- Plus Interaction

You need to calculate the within-between interactions in this round-about way...

```
sub7$ww.int = sub7$d.white * sub7$white.m

tempww = sub7 %>%
  dplyr::group_by(sampyear) %>%
  dplyr::summarise_at(vars(ww.int), funs(mean(., na.rm=TRUE))) %>%
  rename_at(vars(ww.int), ~ paste0(., ".m"))

sub8 = inner_join(sub7, tempww, by="sampyear")
sub8$d.ww.int = sub8$ww.int - sub8$ww.int.m


lm6 <- lmer(closeblk ~ d.white + white.m + d.ww.int + ww.int.m + (1 + d.white |
sampyear),
            data = sub8, REML = FALSE)
```

6th- Plus Interaction, Cont'd

Does a person's whiteness interact with their community's level of whiteness to lead to varying level of closeness to African-Americans?

```
> summary(lm6)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + white.m + d.ww.int + ww.int.m + (1 + d.white |
    sampyear)
   Data: sub8

     AIC      BIC   logLik deviance df.resid
 51713.3  51780.1 -25847.7  51695.3    12263

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.1565   0.3956
          d.white     0.7658   0.8751   -0.67
 Residual             3.7723   1.9422
Number of obs: 12272, groups:  sampyear, 792

Fixed effects:
              Estimate Std. Error t value
(Intercept)  7.287e+00  1.003e-01  72.684
d.white     -1.496e+00  2.705e-01  -5.531
white.m     -2.065e+00  1.235e-01 -16.722
d.ww.int     3.450e-01  3.761e-01   0.917
ww.int.m    -8.882e+14  9.785e+14  -0.908
```

6th- Plus Interaction, Cont'd

The interaction of whiteness*(community)whiteness = D.ww.int = 0.345(n.s.). Technically, it says that being more white in a whiter area leads to more closeness to African-Americans, but not even close statistically

```
> summary(lm6)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + white.m + d.ww.int + ww.int.m + (1 + d.white |
    sampyear)
   Data: sub8

     AIC      BIC   logLik deviance df.resid
 51713.3  51780.1 -25847.7  51695.3    12263

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 sampyear (Intercept) 0.1565   0.3956
          d.white     0.7658   0.8751   -0.67
 Residual             3.7723   1.9422
Number of obs: 12272, groups:  sampyear, 792

Fixed effects:
              Estimate Std. Error t value
(Intercept)  7.287e+00  1.003e-01  72.684
d.white     -1.496e+00  2.705e-01  -5.531
white.m     -2.065e+00  1.235e-01 -16.722
d.ww.int     3.450e-01  3.761e-01   0.917
ww.int.m    -8.882e+14  9.785e+14  -0.908
```

6th- Plus Interaction, Cont'd

You can see the fit on this is worse than the model without interactions. 2 points higher AIC in this model.

```
> summary(lm6)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: closeblk ~ d.white + white.m + d.ww.int + ww.int.m + (1 + d.white |
    sampyear)
   Data: sub8

     AIC      BIC   logLik deviance df.resid
 51713.3  51780.1 -25847.7  51695.3    12263



Correlation of Fixed Effects:
         (Intr) d.whit whit.m d.int
d.white  -0.225
white.m  -0.972  0.217
d.int     0.191 -0.972 -0.195
ww.int.m  0.095  0.004 -0.103 -0.002

fit warnings:
Some predictor variables are on very different scales: consider rescaling
```