Longitudinal Final Report

2025-04-26

```r
library(haven)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyr)
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

```r
library(lme4)
```

```
## Loading required package: Matrix

##
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```r
library(nlme)
```

```
##
## Attaching package: 'nlme'
```

```
## The following object is masked from 'package:lme4':
##
##       lmList


## The following object is masked from 'package:dplyr':
##
##       collapse
```

```r
library(performance)
library(geepack)
library(ggplot2)
# library(afex) # will mask lmer

df <- read.csv("/Users/kangyiyuan/Desktop/2480 Final Project/data_final.csv")

head(df)
```

```
##      PID time discrimination smoke_ever marijuana_bi marijuana_occ cocaine
## 1 7039    7              2          0            0             0       0
## 2 7039    9              1          0            0             0       0
## 3 7039   11              1          0            0             0       0
## 4 7039   13              2          0            0             0       0
## 5 7040    7              3          0            1             0       0
## 6 7040    9              2          0            1             2       0
##   chronic_con  bmi dietpill amphet barbit tranquil steroids mental_worry race
## 1           0 25.1        0      0      0        0        0            3    1
## 2           1 27.3        0      0      0        0        0            3    1
## 3           1 25.8        0      0      0        0        0            5    1
## 4           1 21.5        0      0      0        0        0            6    1
## 5           0 22.1        0      0      0        0        0            4    1
## 6           0 21.9        0      0      0        0        0            5    1
```

```r
all_vars <- df %>% dplyr::select(-PID, -time) %>% names()

# Check time-variant and tme-invariant
is_time_variant <- function(var) {
  tmp <- df %>% dplyr::select(PID, time, all_of(var))
  var_name <- var
  var_change <- tmp %>%
    group_by(PID) %>%
    summarize(var_sd = sd(.data[[var_name]], na.rm = TRUE)) %>%
    pull(var_sd)

  mean(var_change > 0, na.rm = TRUE) > 0.08 # over 0.08 will be time variant
}

time_variant_vars <- all_vars[sapply(all_vars, is_time_variant)]
time_invariant_vars <- setdiff(all_vars, time_variant_vars)

time_variant_vars
```

```
## [1] "discrimination" "smoke_ever"     "marijuana_bi"   "marijuana_occ"
```

```
##  [5] "cocaine"         "chronic_con"    "bmi"           "dietpill"
##  [9] "amphet"          "barbit"         "tranquil"      "steroids"
## [13] "mental_worry"
```

```
time_invariant_vars
```

```
## [1] "race"
```

```r
# Variable modification:
# Binary change
# Smoke treated as binary
df <- df %>%
  mutate(smoke_ever_clean = case_when(
    smoke_ever %in% c(1,0) ~ 1,    # Yes
    smoke_ever == 5  ~ 0,    # never
  ))

# new variable to indicate whether the observation is a current smoker or not
df <- df %>%
  mutate(current_smoker = case_when (
    smoke_ever == 0 ~ 1,
    smoke_ever %in% c(1, 5) ~ 0
  ))

# race change
problematic_pid <- df %>%
  group_by(PID) %>%
  summarise(n_race = n_distinct(race)) %>%
  filter(n_race > 1) %>%
  pull(PID)

# delete those PID
df <- df %>%
  filter(!(PID %in% problematic_pid))

# race change, merge race number 3-7 into the other category
df <- df %>%
  mutate(race_clean = case_when(
    race == 1 ~ 1,
    race == 2 ~ 2,
    race %in% c(3, 4, 5, 7) ~ 3
  ))

# convert times to wave
df <- df %>%
  mutate(wave = case_when(
    time == 7 ~ 0,
    time == 9 ~ 1,
    time == 11 ~ 2,
    time == 13 ~ 3
    ))
```

data preparation

```r
exposure <- "discrimination"
main_out <- "mental_worry"
level_2 <- "race_clean"

df$race_clean <- as.factor(df$race_clean)

sub_vars <- c("smoke_ever_clean",
              "current_smoker",
              "marijuana_bi",
              "cocaine",
              "dietpill",
              "amphet",
              "barbit",
              "tranquil",
              "steroids",
              "chronic_con",
              "bmi"
               )

sub_vars_d <- c("discrimination",
                 "smoke_ever_clean",
                "current_smoker",
                "marijuana_bi",
                "cocaine",
                "dietpill",
                "amphet",
                "barbit",
                "tranquil",
                "steroids",
                "chronic_con",
                "bmi"
                 )

main_cov <- "discrimination"

df$race_labeled <- factor(df$race,
  levels = c(1, 2, 3, 4, 5, 7),
  labels = c(
    "1 White",
    "2 Black, AA or Negro",
    "3 American Indian or Alaska Native",
    "4 Asian",
    "5 Native Hawaiian or Pacific Islander",
    "7 Some other race"
  )
)

df$race_clean_labeled <- factor(df$race_clean,
  levels = c(1, 2, 3),
  labels = c(
    "1 White",
```

```
    "2 Black, AA or Negro",
    "3 Other"
  )
)
```

factorize variables

OLS.

```r
library(broom)
library(purrr)

# identify different kinds of outcome
# categorical or numerical

# OLS with wave
run_ols_with_aic <- function(sub_var) {
  fm <- paste(main_out, "~ wave  +", sub_var)
  model <- lm(as.formula(fm), data = df)

  coef_info <- tidy(model) %>% filter(term == sub_var)

  tibble(
    substance = sub_var,
    estimate = coef_info$estimate,
    std_error = coef_info$std.error,
    p_value = coef_info$p.value,
    AIC = AIC(model),
    BIC = BIC(model)
  )
}

ols_results_with_aic <- map_dfr(sub_vars, run_ols_with_aic)

model_comparisons <- ols_results_with_aic %>%
  mutate(sig = case_when(
    p_value < 0.001 ~ "***",
    p_value < 0.01  ~ "**",
    p_value < 0.05  ~ "*",
    p_value < 0.1   ~ ".",
    TRUE ~ ""
  ))

print(model_comparisons)
```

```
## # A tibble: 11 x 7
##    substance        estimate std_error  p_value   AIC   BIC sig
##    <chr>               <dbl>     <dbl>    <dbl> <dbl> <dbl> <chr>
## 1 smoke_ever_clean   0.522     0.0739 2.29e-12 6861. 6883. "***"
## 2 current_smoker     0.475     0.0864 4.40e- 8 6880. 6902. "***"
```

```
##  3 marijuana_bi    0.325    0.0716  6.09e- 6 6889. 6912. "***"
##  4 cocaine         0.468    0.117   7.01e- 5 6894. 6916. "***"
##  5 dietpill        0.269    0.0994  6.89e- 3 6903. 6925. "**"
##  6 amphet          0.262    0.107   1.46e- 2 6904. 6926. "*"
##  7 barbit          0.505    0.161   1.71e- 3 6900. 6922. "**"
##  8 tranquil        0.543    0.121   7.44e- 6 6890. 6912. "***"
##  9 steroids        0.359    0.174   3.98e- 2 6906. 6928. "*"
## 10 chronic_con     0.607    0.110   4.08e- 8 6880. 6902. "***"
## 11 bmi             0.00974  0.00614 1.13e- 1 6907. 6930. ""
```

General OLS only with wave

```
worry_ols <- lm(mental_worry ~ wave, data = df)
summary(worry_ols)
```

```
##
## Call:
## lm(formula = mental_worry ~ wave, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5719 -1.4946 -0.4946  1.4281  3.5054
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.57189    0.05976  59.773   <2e-16 ***
## wave        -0.02575    0.03194  -0.806     0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.542 on 1862 degrees of freedom
## Multiple R-squared:  0.0003489,  Adjusted R-squared:  -0.0001879
## F-statistic: 0.6499 on 1 and 1862 DF,  p-value: 0.4202
```

```
cor(df$mental_worry,df$wave)
```

```
## [1] -0.01867981
```

```
cov2cor(vcov(worry_ols))
```

```
##             (Intercept)       wave
## (Intercept)   1.0000000 -0.8017837
## wave         -0.8017837  1.0000000
```

```
AIC(worry_ols)
```

```
## [1] 6907.902
```

```
BIC(worry_ols)
```

```
## [1] 6924.494
```

```
# individual growth plot (select 12 random individual)
set.seed(100)
race1_ids <- df %>%
  filter(race_clean_labeled == "1 White") %>%
  distinct(PID) %>%
  sample_n(5) %>%
  pull(PID)

race2_ids <- df %>%
  filter(race_clean_labeled == "2 Black, AA or Negro") %>%
  distinct(PID) %>%
  sample_n(4) %>%
  pull(PID)

race3_ids <- df %>%
  filter(race_clean_labeled == "3 Other") %>%
  distinct(PID) %>%
  sample_n(3) %>%
  pull(PID)
ids <- c(race1_ids, race2_ids, race3_ids)

df %>%
  filter(PID %in% ids) %>%
  # plot the empirical plots
  ggplot(mapping = aes(x = wave, y = mental_worry)) +
  geom_point() +
  geom_line() +
  facet_wrap(~PID) +
  labs(x = "Wave", y = "Worry Score",
       title = "The Empirical Growth Plots for 12 Participants")
```

The Empirical Growth Plots for 12 Participants
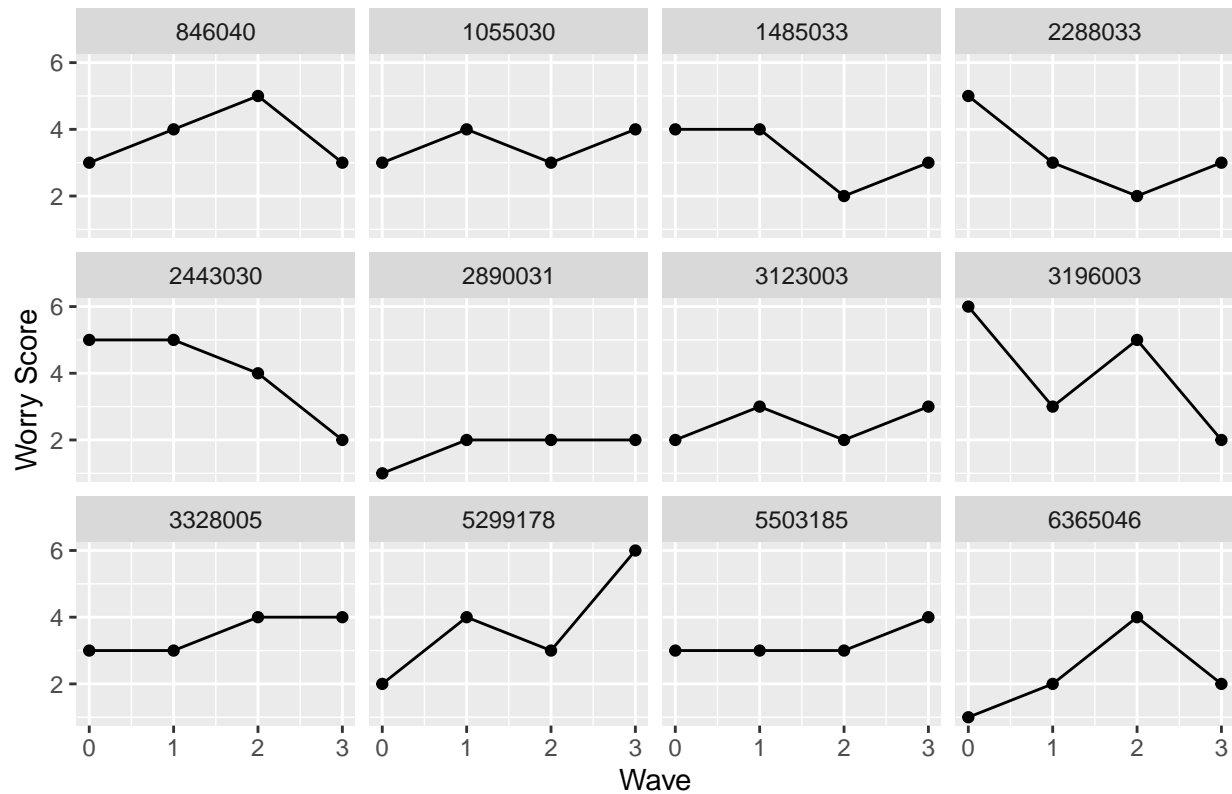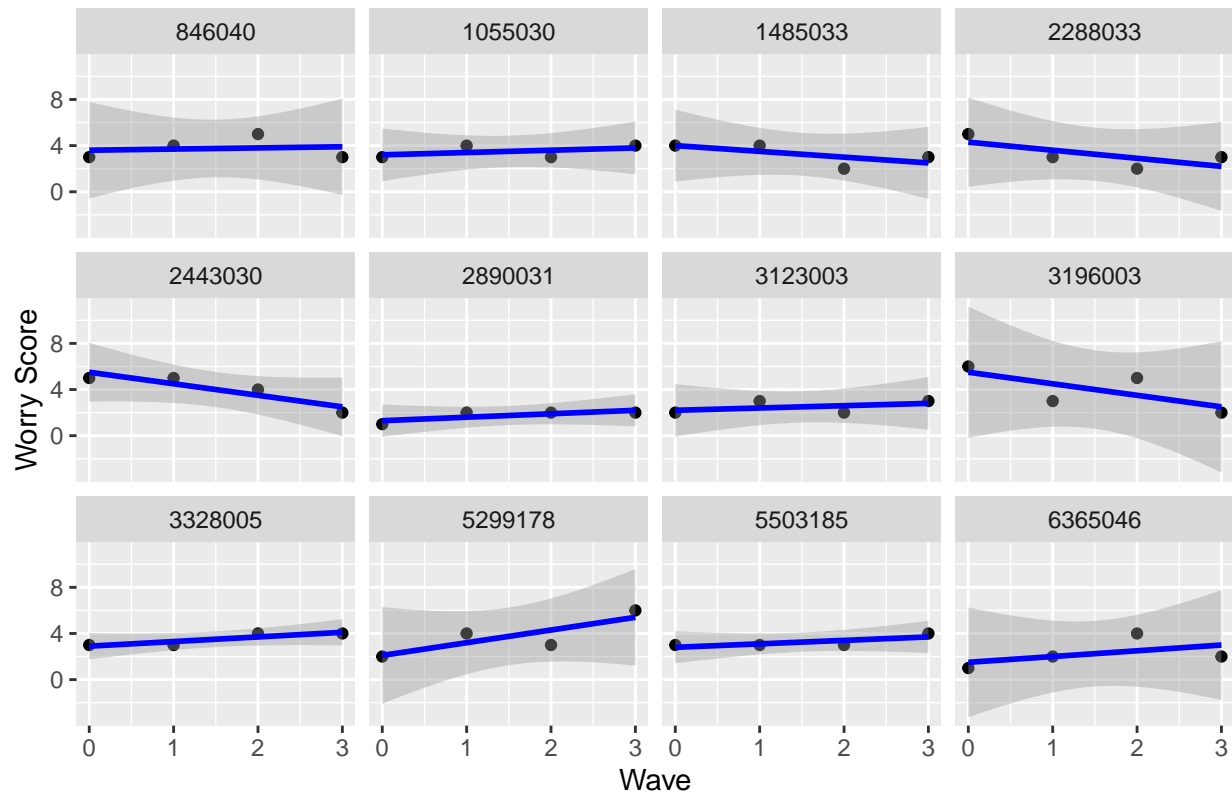


```
df %>%
    filter(PID %in% ids) %>%
# plot the empirical plots
    ggplot(mapping = aes(x = wave, y = mental_worry)) + geom_point() + facet_wrap("PID") +
    geom_smooth(method = "lm", se = T, color = "blue") +
    labs(x = "Wave", y = "Worry Score",
        title = "The Empirical Growth Plots for 12 Participants")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

The Empirical Growth Plots for 12 Participants


```
#
individual_ols <- df %>%
  group_by(PID) %>%
  do(model = lm(mental_worry ~ wave, data = .))

interceptols <- slopeols <- NULL

for(i in 1:nrow(individual_ols)){
  interceptols[i] <- individual_ols[[2]][[i]][["coefficients"]][1]
  slopeols[i] <- individual_ols[[2]][[i]][["coefficients"]][2]
}

summary(interceptols)[4]
```

```
##     Mean
## 3.571888
```

```
summary(slopeols)[4]
```

```
##        Mean
## -0.02575107
```

```
cor(interceptols, slopeols)
```

```
## [1] -0.5959579
```

```r
var(interceptols)
```

```
## [1] 2.05338
```

```r
var(slopeols)
```

```
## [1] 0.3177871
```

```r
# report in table

library(ggplot2)

ggplot(df, mapping = aes(x = wave, y = mental_worry)) +
  geom_smooth(aes(group = as.factor(PID)), method = "lm", se = F, color =
                "black") + geom_smooth(method = "lm", se = F, color = "red") +
  labs(x = "Wave", y = "Worry Score",
       title = "The Individual OLS Plots with Mean Trajectory.")
```

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```


The Individual OLS Plots with Mean Trajectory.

```r
df %>%
  filter(PID %in% ids) %>%
  ggplot(mapping = aes(x = wave, y = mental_worry)) +
```
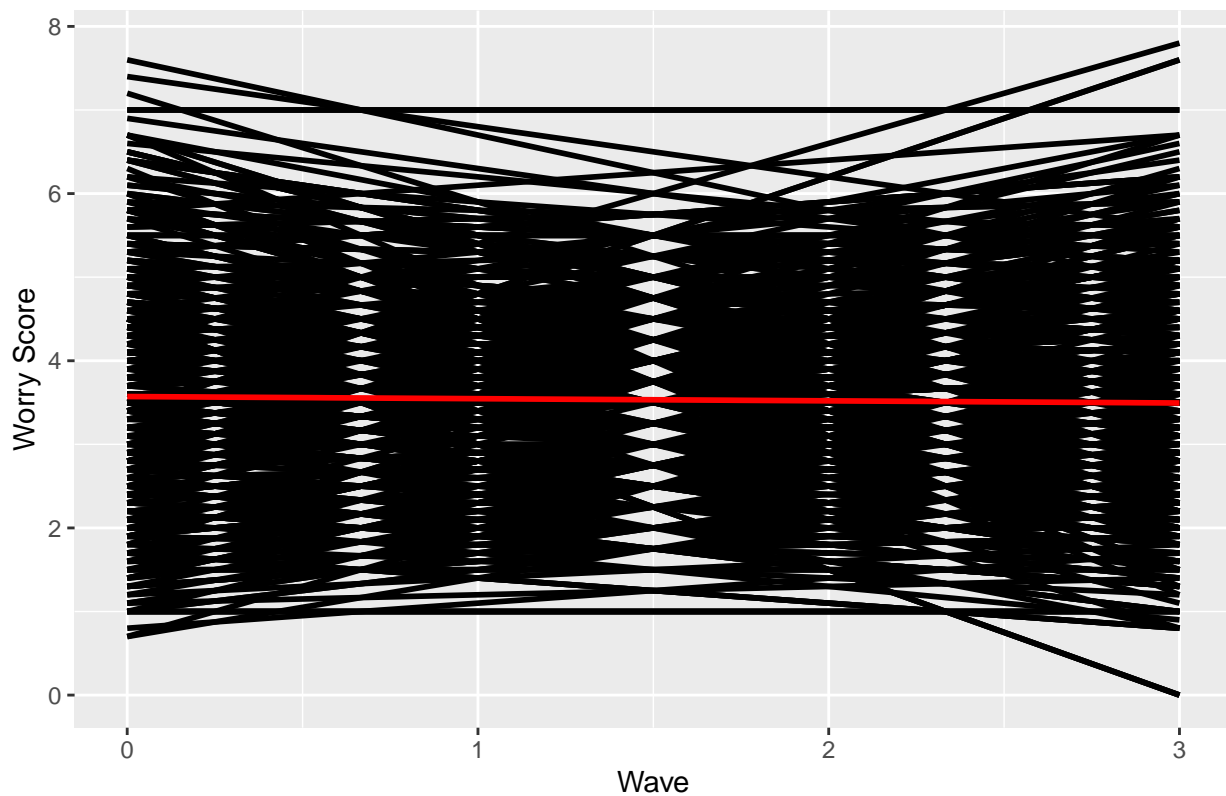
```
  geom_smooth(aes(group = as.factor(PID)), method = "lm", se = F, color =
                "black") + geom_smooth(method = "lm", se = F, color = "red") +
  labs(x = "Wave", y = "Worry Score",
       title = "The Individual OLS of the 12 Plots with Mean Trajectory.")
```

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```

The Individual OLS of the 12 Plots with Mean Trajectory.
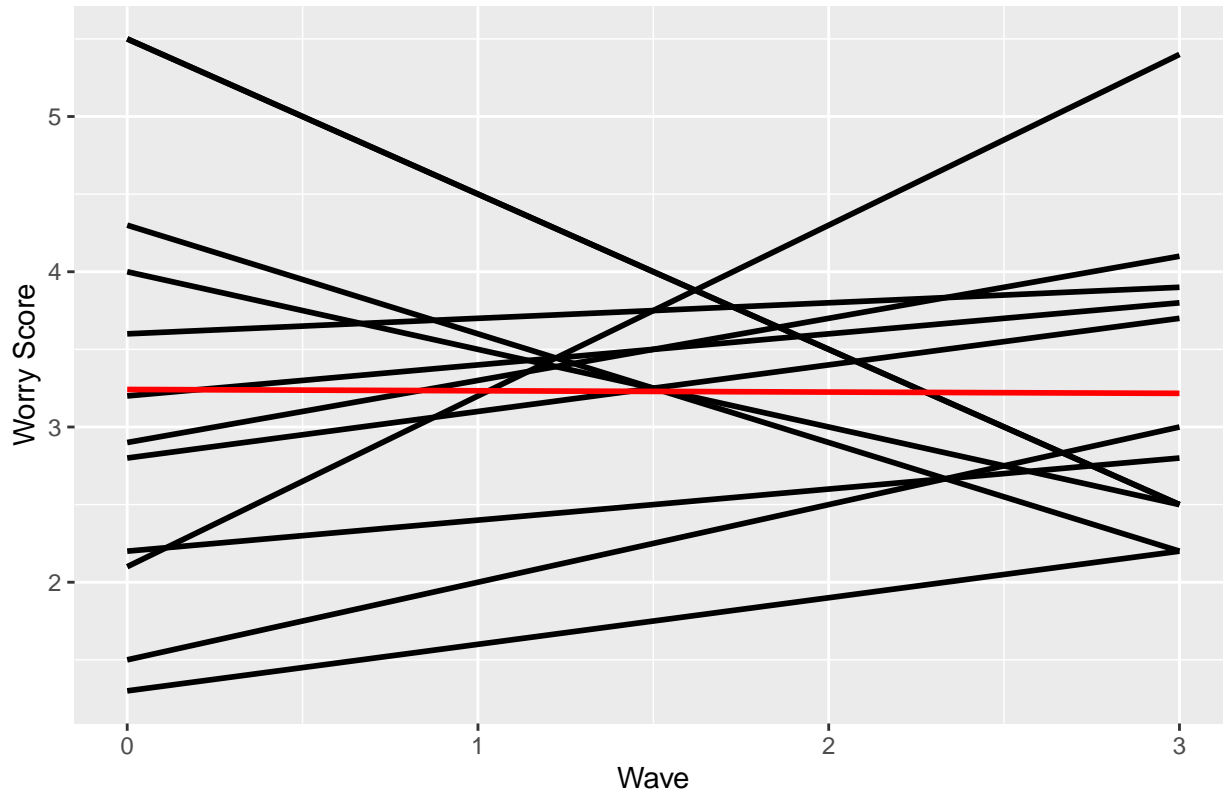
On average, worry score remains stable over time.

Panel plot

```
library(ggplot2)
ggplot(df, aes(x = wave, y = mental_worry, group = PID)) +
  geom_line(alpha = 0.2) +
  stat_smooth(aes(group = 1), method = "lm", se = TRUE, color = "blue") +
  theme_minimal() +
  facet_wrap(~ race_clean_labeled)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```
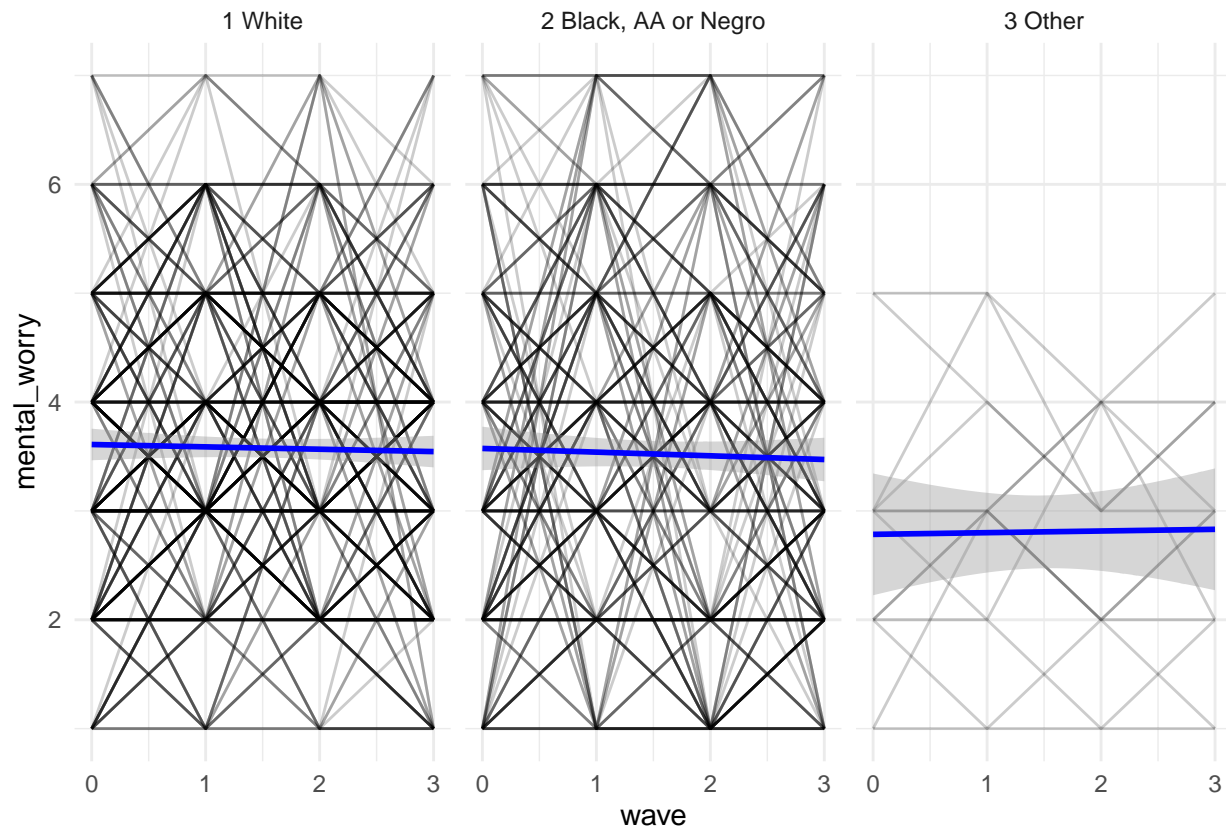
```
# OLS plots by race
df %>%
    filter(PID %in% ids) %>%
ggplot(aes(x = wave, y = mental_worry, group = PID, color = factor(PID))) +
  geom_line(linewidth = 0.5) +
  geom_point() +
  theme_minimal() +
  facet_wrap(~ race_clean_labeled) +
  stat_smooth(data = df, aes(group = race_clean_labeled), method = "lm",
            color = "black", se = FALSE) +
  labs(color = "Participant ID", x = "Wave", y = "Worry Score",
       title = "The Empirical Plots by Race") +
  scale_x_continuous(breaks = 0:3)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

The Empirical Plots by Race
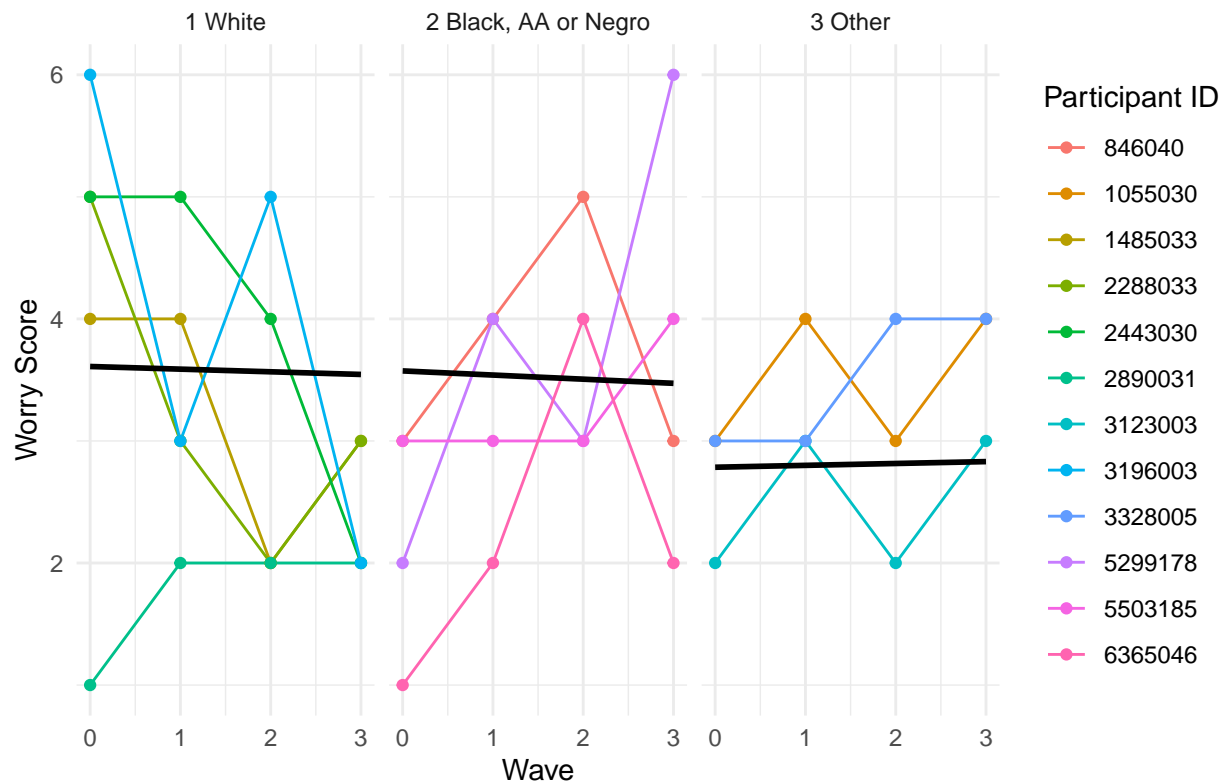
lmer: model building

```r
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':
##
##     lmer

## The following object is masked from 'package:stats':
##
##     step
```

```r
#ols_currentsmoker <-lm(mental_worry ~ wave + current_smoker, data = df_byid)
#summary(ols_currentsmoker)

# model.smoke.ever <- lme(mental_worry ~ wave + discrimination + smoke_ever_clean, data = df, random= ~
# summary(model.smoke.ever)

# # unconditional mean model
# model.a <- lmer(mental_worry ~ 1 + (1|PID), data = df, REML = FALSE)
```

```
# summary(model.a)
#
# # uncon growth by race: level-2 variability small
# model.d <- lmer(mental_worry ~ wave + (wave|race_clean_labeled), data = df, REML = FALSE)
# summary(model.d)

# unconditional growth model
model.wave <- lmer(mental_worry ~ wave + (wave|PID), data = df, REML = FALSE)
summary(model.wave)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ wave + (wave | PID)
##    Data: df
##
##       AIC       BIC    logLik -2*log(L)  df.resid
##    6546.9    6580.1   -3267.4    6534.9      1858
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6766 -0.5793 -0.0636  0.5535  3.3540
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 1.12924  1.0627
##           wave        0.05432  0.2331   -0.35
##  Residual             1.31395  1.1463
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)  3.57189    0.06631 465.99431  53.867   <2e-16 ***
## wave        -0.02575    0.02609 465.97512  -0.987    0.324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## wave -0.596
```

```
icc(model.wave)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.447
##   Unadjusted ICC: 0.447
```

```
# + discrimination
model.dis <- lmer(mental_worry ~ wave + discrimination +
                  (wave|PID), data = df, REML = FALSE)
summary(model.dis)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ wave + discrimination + (wave | PID)
##    Data: df
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##   6514.1   6552.8  -3250.1    6500.1      1857
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5577 -0.5948 -0.0716  0.5507  3.3683
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 1.023    1.0112
##           wave        0.046    0.2145   -0.31
##  Residual             1.317    1.1474
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                 Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)      2.98949    0.11698 1431.44448  25.555  < 2e-16 ***
## wave            -0.02027    0.02578  466.77144  -0.786    0.432
## discrimination   0.22409    0.03753 1855.77472   5.971 2.81e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) wave
## wave        -0.359
## discrimintn -0.834  0.036
```

```r
icc(model.dis)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.427
##   Unadjusted ICC: 0.418
```

```r
# + mari_bi
model.dis.mari <- lmer(mental_worry ~ wave + discrimination +
                       marijuana_bi + (wave|PID),
                   data = df, REML = FALSE)
summary(model.dis.mari)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ wave + discrimination + marijuana_bi + (wave |
##     PID)
##    Data: df
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##   6508.1   6552.4  -3246.1    6492.1      1856
```

```
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6124 -0.5887 -0.0572  0.5607  3.3469
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 1.00278  1.0014
##           wave        0.04446  0.2108   -0.30
##  Residual             1.31529  1.1469
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                 Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)      2.92954    0.11843 1415.63973  24.737  < 2e-16 ***
## wave            -0.03257    0.02607  487.44547  -1.249  0.21217
## discrimination   0.21174    0.03772 1856.79229   5.613 2.28e-08 ***
## marijuana_bi     0.22274    0.07866 1681.69559   2.832  0.00468 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) wave   dscrmn
## wave        -0.318
## discrimintn -0.795  0.055
## marijuana_b -0.175 -0.167 -0.121
```

```r
icc(model.dis.mari)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.424
##   Unadjusted ICC: 0.413
```

```r
# + ever smoke
model.dis.mari.esmoke <- lmer(mental_worry ~ wave + discrimination +
                                marijuana_bi + smoke_ever_clean + (wave|PID),
                     data = df, REML = FALSE)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00223131 (tol = 0.002, component 1)
```

```r
summary(model.dis.mari.esmoke)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula:
## mental_worry ~ wave + discrimination + marijuana_bi + smoke_ever_clean +
##     (wave | PID)
##    Data: df
##
##       AIC      BIC   logLik -2*log(L) df.resid
```

```
##     6497.3     6547.1    -3239.6     6479.3        1855
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6939 -0.5794 -0.0605  0.5558  3.4428
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 0.96149  0.9806
##           wave        0.04794  0.2190   -0.30
##  Residual             1.31065  1.1448
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)        2.85822    0.11917 1370.02368  23.984  < 2e-16 ***
## wave              -0.03226    0.02617  485.53283  -1.233 0.218254
## discrimination     0.20648    0.03763 1856.69687   5.487 4.65e-08 ***
## marijuana_bi       0.15421    0.08055 1769.61934   1.914 0.055731 .
## smoke_ever_clean   0.33736    0.09332 1266.11505   3.615 0.000312 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) wave   dscrmn mrjn_b
## wave        -0.315
## discrimintn -0.778  0.055
## marijuana_b -0.129 -0.161 -0.107
## smok_vr_cln -0.157  0.003 -0.048 -0.237
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.00223131 (tol = 0.002, component 1)
```

```r
icc(model.dis.mari.esmoke)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.417
##   Unadjusted ICC: 0.401
```

```r
# + race
model.dis.mari.esmoke.race <- lmer(mental_worry ~ wave + discrimination +
                                 marijuana_bi + smoke_ever_clean +
                                 (wave|PID) + race_clean_labeled,
                    data = df, REML = FALSE)
summary(model.dis.mari.esmoke.race)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula:
## mental_worry ~ wave + discrimination + marijuana_bi + smoke_ever_clean +
##     (wave | PID) + race_clean_labeled
##    Data: df
##
```

```
##       AIC      BIC    logLik -2*log(L)  df.resid
##    6495.7   6556.6   -3236.9    6473.7      1853
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6954 -0.5774 -0.0621  0.5553  3.4343
##
## Random effects:
##  Groups    Name         Variance Std.Dev. Corr
##  PID       (Intercept) 0.94550  0.9724
##            wave        0.04794  0.2190   -0.30
##  Residual              1.31057  1.1448
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                                    Estimate Std. Error        df t value
## (Intercept)                         2.89021    0.12799 1259.11695  22.581
## wave                               -0.03248    0.02617  485.58273  -1.241
## discrimination                      0.20469    0.03759 1856.20406   5.445
## marijuana_bi                        0.15734    0.08055 1771.18751   1.953
## smoke_ever_clean                    0.33780    0.09312 1259.60780   3.628
## race_clean_labeled2 Black, AA or Negro  -0.01987    0.10394  464.96203  -0.191
## race_clean_labeled3 Other          -0.73315    0.31036  462.43368  -2.362
##                                    Pr(>|t|)
## (Intercept)                         < 2e-16 ***
## wave                               0.215180
## discrimination                     5.86e-08 ***
## marijuana_bi                       0.050935 .
## smoke_ever_clean                   0.000297 ***
## race_clean_labeled2 Black, AA or Negro 0.848478
## race_clean_labeled3 Other          0.018576 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) wave   dscrmn mrjn_b smk_v_ r_BAoN
## wave       -0.289
## discrimintn -0.719  0.055
## marijuana_b -0.139 -0.161 -0.108
## smok_vr_cln -0.162  0.003 -0.049 -0.235
## rc__2B,AAoN -0.362 -0.012 -0.018  0.058  0.046
## rc_cln_lb3O -0.132  0.003  0.025 -0.015  0.006  0.142
```

```r
icc(model.dis.mari.esmoke.race)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.413
##   Unadjusted ICC: 0.394
```

```r
# Discussion
first_modelpoi <- glm(formula = mental_worry ~ wave, family = poisson(link = "log"), data = df)
summary(first_modelpoi)
```

```
## 
## Call:
## glm(formula = mental_worry ~ wave, family = poisson(link = "log"),
##     data = df)
## 
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.273121   0.020539  61.986   <2e-16 ***
## wave        -0.007288   0.011022  -0.661    0.508
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for poisson family taken to be 1)
## 
##     Null deviance: 1303.6  on 1863  degrees of freedom
## Residual deviance: 1303.1  on 1862  degrees of freedom
## AIC: 6987
## 
## Number of Fisher Scoring iterations: 4
```

```
final_modelpoi <- glm(formula = mental_worry ~ wave + discrimination + marijuana_bi + smoke_ever_clean
summary(final_modelpoi)
```

```
## 
## Call:
## glm(formula = mental_worry ~ wave + discrimination + marijuana_bi +
##     smoke_ever_clean + race_clean_labeled, family = poisson(link = "log"),
##     data = df)
## 
## Coefficients:
##                                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)                       1.034555   0.041807  24.746  < 2e-16
## wave                             -0.007971   0.011114  -0.717  0.47324
## discrimination                    0.073609   0.012976   5.673 1.40e-08
## marijuana_bi                      0.029883   0.026999   1.107  0.26837
## smoke_ever_clean                  0.114446   0.027410   4.175 2.97e-05
## race_clean_labeled2 Black, AA or Negro -0.007895   0.025345  -0.312  0.75541
## race_clean_labeled3 Other        -0.227030   0.084428  -2.689  0.00717
## 
## (Intercept)                      ***
## wave
## discrimination                   ***
## marijuana_bi
## smoke_ever_clean                 ***
## race_clean_labeled2 Black, AA or Negro
## race_clean_labeled3 Other        **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for poisson family taken to be 1)
## 
##     Null deviance: 1303.6  on 1863  degrees of freedom
## Residual deviance: 1227.0  on 1857  degrees of freedom
## AIC: 6920.8
```

```
##
## Number of Fisher Scoring iterations: 4
```

```
first_model_quasi <- glm(formula = mental_worry ~ wave, family = quasipoisson(), data = df)
summary(first_model_quasi)
```

```
##
## Call:
## glm(formula = mental_worry ~ wave, family = quasipoisson(), data = df)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.273121   0.016846  75.573   <2e-16 ***
## wave        -0.007288   0.009040  -0.806     0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 0.6727554)
##
##     Null deviance: 1303.6  on 1863  degrees of freedom
## Residual deviance: 1303.1  on 1862  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```
final_model_quasi <- glm(formula = mental_worry ~ wave + discrimination + marijuana_bi + smoke_ever_clea
summary(final_model_quasi)
```

```
##
## Call:
## glm(formula = mental_worry ~ wave + discrimination + marijuana_bi +
##     smoke_ever_clean + race_clean_labeled, family = quasipoisson(),
##     data = df)
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        1.034555   0.033343  31.028   < 2e-16
## wave                              -0.007971   0.008864  -0.899 0.368623
## discrimination                     0.073609   0.010349   7.113 1.62e-12
## marijuana_bi                       0.029883   0.021533   1.388 0.165367
## smoke_ever_clean                   0.114446   0.021860   5.235 1.83e-07
## race_clean_labeled2 Black, AA or Negro -0.007895   0.020214  -0.391 0.696149
## race_clean_labeled3 Other         -0.227030   0.067335  -3.372 0.000763
##
## (Intercept)                        ***
## wave
## discrimination                     ***
## marijuana_bi
## smoke_ever_clean                   ***
## race_clean_labeled2 Black, AA or Negro
## race_clean_labeled3 Other          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## (Dispersion parameter for quasipoisson family taken to be 0.6360752)
## 
##      Null deviance: 1303.6  on 1863  degrees of freedom
## Residual deviance: 1227.0  on 1857  degrees of freedom
## AIC: NA
## 
## Number of Fisher Scoring iterations: 4
```

```r
model_list <- list(model.wave, model.dis, model.dis.mari, model.dis.mari.esmoke,
                   model.dis.mari.esmoke.race)
aic_tab <- function(model){
  return(summary(model)$AICtab)
}

models_aic_tab <- sapply(model_list, aic_tab)
colnames(models_aic_tab) <- c("model.wave", "model.dis", "model.dis.mari",
                              "model.dis.mari.esmoke","model.dis.mari.esmoke.race")

models_aic_tab
```

```
##            model.wave model.dis model.dis.mari model.dis.mari.esmoke
## AIC          6546.890  6514.108       6508.121              6497.297
## BIC          6580.073  6552.821       6552.365              6547.072
## logLik      -3267.445 -3250.054      -3246.060             -3239.649
## -2*log(L)    6534.890  6500.108       6492.121              6479.297
## df.resid     1858.000  1857.000       1856.000              1855.000
##            model.dis.mari.esmoke.race
## AIC                          6495.728
## BIC                          6556.563
## logLik                      -3236.864
## -2*log(L)                    6473.728
## df.resid                     1853.000
```

```r
# model.wave icc
icc.b <- icc(model.wave)$ICC_adjusted
icc_total.b <- sum(as.data.frame(VarCorr(model.wave),comp="Variance")$vcov[-3])
icc_intercept.b <- as.data.frame(VarCorr(model.wave),comp="Variance")$vcov[1]/icc_total.b
icc_wave.b <- as.data.frame(VarCorr(model.wave),comp="Variance")$vcov[2]/icc_total.b
icc_residual.b <- as.data.frame(VarCorr(model.wave),comp="Variance")$vcov[4]/icc_total.b

# model.dis icc
icc.dis <- icc(model.dis)$ICC_adjusted
icc_total.dis <- sum(as.data.frame(VarCorr(model.dis),comp="Variance")$vcov[-3])
icc_intercept.dis <- as.data.frame(VarCorr(model.dis),comp="Variance")$vcov[1]/icc_total.dis
icc_wave.dis <- as.data.frame(VarCorr(model.dis),comp="Variance")$vcov[2]/icc_total.dis
icc_residual.dis <- as.data.frame(VarCorr(model.dis),comp="Variance")$vcov[4]/icc_total.dis

# model.dis.mari
icc.dis.mari <- icc(model.dis.mari)$ICC_adjusted
icc_total.dismari <- sum(as.data.frame(VarCorr(model.dis.mari),
                                       comp="Variance")$vcov[-3])
icc_intercept.dismari <- as.data.frame(VarCorr(model.dis.mari),
                                       comp="Variance")$vcov[1]/icc_total.dismari
```

```r
icc_wave.dismari <- as.data.frame(VarCorr(model.dis.mari),
                                  comp="Variance")$vcov[2]/icc_total.dismari
icc_residual.dismari <- as.data.frame(VarCorr(model.dis.mari),
                                      comp="Variance")$vcov[4]/icc_total.dismari


# model.dis.mari.esmoke
icc.dis.mari.esmoke <- icc(model.dis.mari.esmoke)$ICC_adjusted
icc_total.dis.mari.esmoke <- sum(as.data.frame(VarCorr(model.dis.mari.esmoke),
                                               comp="Variance")$vcov[-3])
icc_intercept.dis.mari.esmoke <- as.data.frame(VarCorr(model.dis.mari.esmoke),
                                               comp="Variance")$vcov[1]/icc_total.dis.mari.esmoke
icc_wave.dis.mari.esmoke <- as.data.frame(VarCorr(model.dis.mari.esmoke),
                                          comp="Variance")$vcov[2]/icc_total.dis.mari.esmoke
icc_residual.dis.mari.esmoke <- as.data.frame(VarCorr(model.dis.mari.esmoke),
                                              comp="Variance")$vcov[4]/icc_total.dis.mari.esmoke


# model.dis.mari.esmoke + race
icc.dis.mari.esmoke.race <- icc(model.dis.mari.esmoke.race)$ICC_adjusted
icc_total.dis.mari.esmoke.race <-
  sum(as.data.frame(VarCorr(model.dis.mari.esmoke.race),comp="Variance")$vcov[-3])
icc_intercept.dis.mari.esmoke.race <-
  as.data.frame(VarCorr(model.dis.mari.esmoke.race),
                comp="Variance")$vcov[1]/icc_total.dis.mari.esmoke
icc_wave.dis.mari.esmoke.race <-
  as.data.frame(VarCorr(model.dis.mari.esmoke.race),
                comp="Variance")$vcov[2]/icc_total.dis.mari.esmoke
icc_residual.dis.mari.esmoke.race <-
  as.data.frame(VarCorr(model.dis.mari.esmoke.race),
                comp="Variance")$vcov[4]/icc_total.dis.mari.esmoke


icc_table <- data.frame (
  Model = c("model.wave", "model.dis", "model.dis.mari", "model.dis.mari.esmoke",
            "model.dis.mari.esmoke.race"),
  icc_intercept = c(icc_intercept.b, icc_intercept.dis, icc_intercept.dismari,
                    icc_intercept.dis.mari.esmoke, icc_intercept.dis.mari.esmoke.race),
  icc_wave = c(icc_wave.b, icc_wave.dis, icc_wave.dismari,
               icc_wave.dis.mari.esmoke, icc_wave.dis.mari.esmoke.race),
  icc_residual = c(icc_residual.b, icc_residual.dis, icc_residual.dismari,
                   icc_residual.dis.mari.esmoke, icc_residual.dis.mari.esmoke.race),
  icc_total_adjusted = c(icc.b, icc.dis, icc.dis.mari,
                         icc.dis.mari.esmoke, icc.dis.mari.esmoke.race)
)

print(icc_table)
```

```
##                       Model icc_intercept   icc_wave icc_residual
## 1                 model.wave     0.4521463 0.02174827    0.5261054
## 2                  model.dis     0.4287358 0.01928794    0.5519763
## 3             model.dis.mari     0.4244526 0.01881676    0.5567307
## 4      model.dis.mari.esmoke     0.4144198 0.02066309    0.5649171
## 5 model.dis.mari.esmoke.race     0.4075277 0.02066326    0.5648827
##   icc_total_adjusted
```

22

```
## 1          0.4466821
## 2          0.4268750
## 3          0.4237545
## 4          0.4170820
## 5          0.4134289
```

Summary Table 1

Parameter	Model 1 (Unconditional growth)	Model 2 (+ Discrimination)	Model 3 (+ Marijuana)	Model 4 (+ Ever smoking)	Model 5 (Final) (+ Race)
Fixed Effects					
Intercept	3.572 (0.066)	2.990 (0.117)***	2.930 (0.118)***	2.858 (0.119)***	2.890 (0.128)***
Wave	-0.026 (0.026)	-0.020 (0.026)	-0.033 (0.026)	-0.032 (0.026)	-0.033 (0.026)
Discrimination	-	0.224 (0.038)***	0.212 (0.038)***	0.207 (0.038)***	0.205 (0.038)***
Marijuana use	-	-	0.2227 (0.079)**	0.154 (0.081).	0.157 (0.081).
Ever Smoking	-	-	-	0.337 (0.093)***	0.338 (0.093)***
Race (ref: White)	-	-	-	-	-
Black, AA, or Negro	-	-	-	-	-0.020 (0.104)
Other	-	-	-	-	-0.733 (0.310)*
Random Effects					
Residual (Level 1)	1.314 (1.146)	1.317 (1.147)	1.315 (1.147)	1.311 (1.145)	1.311 (1.145)
Intercept (Level 2)	1.129 (1.063)	1.023 (1.011)	1.003 (1.001)	0.962 (0.981)	0.986 (0.972)
Wave (Level 2)	0.054 (0.233)	0.046 (0.214)	0.045 (0.211)	0.048 (0.219)	0.048 (0.219)
Model Fit					
AIC	6547	6514	6508	6497	**6496**
BIC	6580	6553	6552	**6547**	6557
Log-Likelihood	-3267	-3250	-3246	-3240	**-3237**
Deviance	6535	6500	6492	6479	**6474**

AA = African-American.   ******* $p < .001$, ****** $p < .01$, ***** $p < .05$, **.** $p < .1$.

Summary Table ICC

Model	ICC Intercept	ICC Wave	ICC Residual	ICC Total Adjusted
Model 1 (Unconditional growth)	0.452	0.0217	0.526	0.447
Model 2 (+ Discrimination)	0.429	0.0193	0.552	0.427
Model 3 (+ Marijuana)	0.424	0.0188	0.557	0.424
Model 4 (+ Ever Smoke)	0.414	0.0207	0.565	0.417
Model 5 (Final: + Race)	**0.408**	**0.0207**	**0.565**	**0.413**

```r
# model + interactions
model.inter1 <- lmer(mental_worry ~ wave + discrimination*wave +
                    marijuana_bi + smoke_ever_clean +
                    (wave|PID)+ race_clean_labeled,
                    data = df, REML = FALSE)
summary(model.inter1)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ wave + discrimination * wave + marijuana_bi +
##     smoke_ever_clean + (wave | PID) + race_clean_labeled
##    Data: df
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##   6496.6   6562.9  -3236.3    6472.6      1852
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6707 -0.5760 -0.0630  0.5672  3.4387
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 0.93971  0.9694
##           wave        0.04719  0.2172   -0.29
##  Residual             1.31043  1.1447
```

```
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                                     Estimate Std. Error         df t value
## (Intercept)                          2.77967    0.16387 1011.92316  16.963
## wave                                 0.04479    0.07637  859.69108   0.586
## discrimination                       0.24793    0.05496 1068.39706   4.511
## marijuana_bi                         0.15428    0.08059 1770.98465   1.914
## smoke_ever_clean                     0.33605    0.09312 1260.35333   3.609
## race_clean_labeled2 Black, AA or Negro -0.01990  0.10395  464.89947  -0.191
## race_clean_labeled3 Other           -0.73261    0.31040  462.37756  -2.360
## wave:discrimination                 -0.02991    0.02778  922.62007  -1.077
##                                     Pr(>|t|)
## (Intercept)                         < 2e-16 ***
## wave                                0.55771
## discrimination                      7.17e-06 ***
## marijuana_bi                        0.05573 .
## smoke_ever_clean                    0.00032 ***
## race_clean_labeled2 Black, AA or Negro 0.84828
## race_clean_labeled3 Other           0.01868 *
## wave:discrimination                 0.28181
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) wave   dscrmn mrjn_b smk_v_ r_BAoN rc__3O
## wave        -0.664
## discrimintn -0.840  0.699
## marijuana_b -0.085 -0.090 -0.101
## smok_vr_cln -0.115 -0.016 -0.047 -0.234
## rc__2B,AAoN -0.282 -0.005 -0.013  0.058  0.046
## rc_cln_lb3O -0.105  0.003  0.018 -0.015  0.006  0.142
## wv:dscrmntn  0.625 -0.940 -0.730  0.037  0.018  0.001 -0.002
```

```
icc(model.inter1)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.413
##   Unadjusted ICC: 0.394
```

```
model.inter2 <- lmer(mental_worry ~ wave + discrimination +
                       marijuana_bi*wave + smoke_ever_clean + (wave|PID) +
                       race_clean_labeled,
                     data = df, REML = FALSE)
summary(model.inter2)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ wave + discrimination + marijuana_bi * wave +
##     smoke_ever_clean + (wave | PID) + race_clean_labeled
##    Data: df
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##   6497.7   6564.1  -3236.9    6473.7      1852
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6951 -0.5777 -0.0618  0.5551  3.4337
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 0.94573  0.9725
##           wave        0.04795  0.2190   -0.30
##  Residual             1.31052  1.1448
## Number of obs: 1864, groups:  PID, 466
```

```
##
## Fixed effects:
##                                    Estimate Std. Error       df t value
## (Intercept)                        2.892e+00  1.310e-01  1.237e+03  22.073
## wave                              -3.431e-02  3.638e-02  5.885e+02  -0.943
## discrimination                     2.048e-01  3.765e-02  1.856e+03   5.441
## marijuana_bi                       1.516e-01  1.128e-01  1.024e+03   1.344
## smoke_ever_clean                   3.377e-01  9.312e-02  1.259e+03   3.627
## race_clean_labeled2 Black, AA or Negro -1.979e-02  1.040e-01  4.650e+02  -0.190
## race_clean_labeled3 Other         -7.333e-01  3.104e-01  4.624e+02  -2.363
## wave:marijuana_bi                  3.879e-03  5.365e-02  7.593e+02   0.072
##                                    Pr(>|t|)
## (Intercept)                        < 2e-16 ***
## wave                               0.346011
## discrimination                       6e-08 ***
## marijuana_bi                       0.179351
## smoke_ever_clean                   0.000298 ***
## race_clean_labeled2 Black, AA or Negro 0.849127
## race_clean_labeled3 Other          0.018554 *
## wave:marijuana_bi                  0.942384
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) wave   dscrmn mrjn_b smk_v_ r_BAoN rc__3O
## wave       -0.352
## discrimintn -0.689  0.001
## marijuana_b -0.246  0.403 -0.116
## smok_vr_cln -0.159  0.006 -0.049 -0.164
## rc__2B,AAoN -0.351 -0.017 -0.018  0.033  0.046
## rc_cln_lb3O -0.131  0.008  0.024 -0.005  0.006  0.142
## wave:mrjn_b  0.214 -0.695  0.055 -0.700 -0.006  0.012 -0.009
```

```r
icc(model.inter2)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.413
##   Unadjusted ICC: 0.394
```

```r
model.inter3 <- lmer(mental_worry ~ wave + discrimination + marijuana_bi +
                     smoke_ever_clean*wave + (wave|PID)+ race_clean_labeled,
                     data = df, REML = FALSE)
summary(model.inter3)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula:
## mental_worry ~ wave + discrimination + marijuana_bi + smoke_ever_clean *
##     wave + (wave | PID) + race_clean_labeled
##    Data: df
##
##      AIC       BIC    logLik -2*log(L)  df.resid
##   6494.7    6561.1   -3235.3    6470.7      1852
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7115 -0.5752 -0.0676  0.5628  3.3993
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 0.93992  0.9695
##           wave        0.04779  0.2186   -0.29
##  Residual             1.30816  1.1437
## Number of obs: 1864, groups:  PID, 466
##
```

```
## Fixed effects:
##                                      Estimate Std. Error          df t value
## (Intercept)                           2.84999    0.12991 1219.41364  21.938
## wave                                  0.00081    0.03238  520.53626   0.025
## discrimination                        0.20122    0.03761 1856.33870   5.350
## marijuana_bi                          0.15474    0.08051 1770.72055   1.922
## smoke_ever_clean                      0.48314    0.12485  794.61135   3.870
## race_clean_labeled2 Black, AA or Negro -0.01680   0.10391  465.11833  -0.162
## race_clean_labeled3 Other            -0.73459    0.31023  462.34856  -2.368
## wave:smoke_ever_clean                -0.09619    0.05517  588.90796  -1.744
##                                      Pr(>|t|)
## (Intercept)                           < 2e-16 ***
## wave                                  0.980050
## discrimination                        9.9e-08 ***
## marijuana_bi                          0.054766 .
## smoke_ever_clean                      0.000118 ***
## race_clean_labeled2 Black, AA or Negro 0.871609
## race_clean_labeled3 Other            0.018302 *
## wave:smoke_ever_clean                0.081734 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) wave   dscrmn mrjn_b smk_v_ r_BAoN rc__3O
## wave       -0.333
## discrimintn -0.697  0.013
## marijuana_b -0.133 -0.142 -0.107
## smok_vr_cln -0.237  0.395 -0.072 -0.188
## rc__2B,AAoN -0.359  0.000 -0.019  0.058  0.045
## rc_cln_lb3O -0.130  0.001  0.025 -0.015  0.003  0.142
## wv:smk_vr_c  0.177 -0.590  0.053  0.019 -0.667 -0.016  0.002
```

```r
icc(model.inter3)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.414
##   Unadjusted ICC: 0.394
```

Discussion

Factor wave for discussion

```r
df_mod <- df %>%
  mutate(fac_wave = factor(wave))

model.wave.f <- lmer(mental_worry ~ fac_wave + (wave|PID),
                     data = df_mod, REML = FALSE)
summary(model.wave.f)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ fac_wave + (wave | PID)
##    Data: df_mod
##
##       AIC      BIC   logLik -2*log(L) df.resid
##    6528.7   6572.9  -3256.3    6512.7     1856
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7505 -0.5830 -0.0551  0.5487  3.3564
##
```

```
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 1.15092  1.073
##           wave        0.06051  0.246    -0.36
##  Residual             1.28299  1.133
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##               Estimate Std. Error          df t value Pr(>|t|)
## (Intercept)    3.43562    0.07227   646.10094  47.539  < 2e-16 ***
## fac_wave1      0.29399    0.07507  1161.71008   3.916 9.53e-05 ***
## fac_wave2      0.12661    0.07763  1315.05967   1.631    0.103
## fac_wave3     -0.03004    0.08170   551.31404  -0.368    0.713
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) fc_wv1 fc_wv2
## fac_wave1 -0.545
## fac_wave2 -0.564  0.517
## fac_wave3 -0.570  0.512  0.557
```

```r
model.dis.f <- lmer(mental_worry ~ fac_wave + discrimination + (wave|PID),
                    data = df_mod, REML = FALSE)
summary(model.dis.f)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ fac_wave + discrimination + (wave | PID)
##    Data: df_mod
##
##      AIC       BIC    logLik -2*log(L)  df.resid
##   6494.0    6543.8   -3238.0    6476.0      1855
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7678 -0.5884 -0.0497  0.5679  3.3715
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 1.04449  1.0220
##           wave        0.05261  0.2294   -0.33
##  Residual             1.28339  1.1329
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                  Estimate Std. Error          df t value Pr(>|t|)
## (Intercept)       2.83523    0.12078  1518.65634  23.474  < 2e-16 ***
## fac_wave1         0.31363    0.07504  1153.55354   4.179 3.14e-05 ***
## fac_wave2         0.14477    0.07726  1322.48445   1.874   0.0612 .
## fac_wave3        -0.01090    0.08083   554.65603  -0.135   0.8928
## discrimination    0.22877    0.03732  1858.17182   6.130 1.07e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) fc_wv1 fc_wv2 fc_wv3
## fac_wave1   -0.357
## fac_wave2   -0.362  0.516
## fac_wave3   -0.364  0.512  0.551
## discrimintn -0.811  0.043  0.038  0.039
```

```r
# + mari_bi
model.dis.mari.f <- lmer(mental_worry ~ fac_wave + discrimination +
                         marijuana_bi + (wave|PID),
                    data = df_mod, REML = FALSE)
summary(model.dis.mari.f)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: mental_worry ~ fac_wave + discrimination + marijuana_bi + (wave |
##     PID)
##    Data: df_mod
##
##       AIC       BIC    logLik -2*log(L)  df.resid
##    6490.1    6545.4   -3235.0    6470.1      1854
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8078 -0.5838 -0.0531  0.5594  3.3559
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 1.0248   1.0123
##           wave        0.0507   0.2252   -0.32
##  Residual             1.2845   1.1333
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                  Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)       2.78970    0.12183 1503.78551  22.899  < 2e-16 ***
## fac_wave1         0.29133    0.07560 1168.60565   3.853 0.000123 ***
## fac_wave2         0.11185    0.07836 1356.24793   1.427 0.153669
## fac_wave3        -0.04263    0.08168  576.42604  -0.522 0.601929
## discrimination    0.21807    0.03754 1858.60985   5.809 7.39e-09 ***
## marijuana_bi      0.19161    0.07862 1682.69019   2.437 0.014907 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) fc_wv1 fc_wv2 fc_wv3 dscrmn
## fac_wave1  -0.332
## fac_wave2  -0.326  0.525
## fac_wave3  -0.331  0.520  0.562
## discrimintn -0.778  0.057  0.059  0.058
## marijuana_b -0.148 -0.121 -0.173 -0.160 -0.123
```

```r
icc(model.dis.mari.f)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates
##   the conditional random effect variances.
##   Solution: Respecify fixed structure!
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.444
##   Unadjusted ICC: 0.430
```

```r
# + ever smoke
model.dis.mari.esmoke.f <- lmer(mental_worry ~ fac_wave + discrimination +
                                marijuana_bi + smoke_ever_clean + (wave|PID),
                  data = df_mod, REML = FALSE)
summary(model.dis.mari.esmoke.f)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula:
## mental_worry ~ fac_wave + discrimination + marijuana_bi + smoke_ever_clean +
##     (wave | PID)
##    Data: df_mod
##
##       AIC       BIC    logLik -2*log(L)  df.resid
##    6479.4    6540.2   -3228.7    6457.4      1853
```

```
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8910 -0.5756 -0.0532  0.5545  3.4504
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  PID      (Intercept) 0.98232  0.9911
##           wave        0.05408  0.2326   -0.32
##  Residual             1.28051  1.1316
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)        2.71917    0.12252 1463.10904  22.194  < 2e-16 ***
## fac_wave1          0.28906    0.07554 1172.56480   3.827 0.000137 ***
## fac_wave2          0.11336    0.07842 1350.01713   1.446 0.148549
## fac_wave3         -0.04301    0.08197  573.12684  -0.525 0.599955
## discrimination     0.21296    0.03745 1858.37458   5.686 1.51e-08 ***
## marijuana_bi       0.12448    0.08047 1770.04027   1.547 0.122049
## smoke_ever_clean   0.33420    0.09300 1274.18589   3.594 0.000338 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) fc_wv1 fc_wv2 fc_wv3 dscrmn mrjn_b
## fac_wave1   -0.329
## fac_wave2   -0.324  0.525
## fac_wave3   -0.327  0.521  0.564
## discrimintn -0.763  0.057  0.058  0.057
## marijuana_b -0.106 -0.115 -0.168 -0.154 -0.109
## smok_vr_cln -0.152 -0.008  0.005 -0.001 -0.048 -0.235
```

```r
icc(model.dis.mari.esmoke.f)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates
##    the conditional random effect variances.
##    Solution: Respecify fixed structure!
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.434
##   Unadjusted ICC: 0.415
```

```r
# + race
model.dis.mari.esmoke.race.f <- lmer(mental_worry ~ fac_wave +
                                discrimination + marijuana_bi +
                                smoke_ever_clean + (wave|PID) +
                                race_clean_labeled,
                     data = df_mod, REML = FALSE)
summary(model.dis.mari.esmoke.race.f)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##    method [lmerModLmerTest]
## Formula:
## mental_worry ~ fac_wave + discrimination + marijuana_bi + smoke_ever_clean +
##     (wave | PID) + race_clean_labeled
##    Data: df_mod
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##   6477.9   6549.8  -3225.9    6451.9      1851
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8919 -0.5736 -0.0538  0.5567  3.4419
```

```
##
## Random effects:
##  Groups    Name        Variance Std.Dev. Corr
##  PID       (Intercept) 0.96656  0.9831
##            wave        0.05409  0.2326   -0.32
##  Residual              1.28041  1.1316
## Number of obs: 1864, groups:  PID, 466
##
## Fixed effects:
##                                     Estimate Std. Error        df t value
## (Intercept)                          2.75252    0.13106 1348.63771  21.003
## fac_wave1                            0.28855    0.07553 1172.71284   3.820
## fac_wave2                            0.11269    0.07842 1350.16650   1.437
## fac_wave3                           -0.04367    0.08197  573.18984  -0.533
## discrimination                       0.21119    0.03742 1857.94557   5.644
## marijuana_bi                         0.12758    0.08046 1771.62281   1.586
## smoke_ever_clean                     0.33457    0.09280 1267.64013   3.605
## race_clean_labeled2 Black, AA or Negro  -0.02301    0.10391  464.92992  -0.221
## race_clean_labeled3 Other           -0.73062    0.31025  462.39567  -2.355
##                                     Pr(>|t|)
## (Intercept)                          < 2e-16 ***
## fac_wave1                           0.000140 ***
## fac_wave2                           0.150956
## fac_wave3                           0.594381
## discrimination                     1.91e-08 ***
## marijuana_bi                       0.113026
## smoke_ever_clean                   0.000324 ***
## race_clean_labeled2 Black, AA or Negro 0.824873
## race_clean_labeled3 Other          0.018943 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) fc_wv1 fc_wv2 fc_wv3 dscrmn mrjn_b smk_v_ r_BAoN
## fac_wave1  -0.304
## fac_wave2  -0.299  0.525
## fac_wave3  -0.302  0.521  0.564
## discrimintn -0.707  0.057  0.059  0.058
## marijuana_b -0.117 -0.116 -0.168 -0.154 -0.111
## smok_vr_cln -0.157 -0.009  0.005 -0.002 -0.049 -0.233
## rc__2B,AAoN -0.352 -0.009 -0.012 -0.012 -0.018  0.058  0.046
## rc_cln_lb3O -0.129  0.003  0.003  0.003  0.024 -0.015  0.006  0.142
```

```r
icc(model.dis.mari.esmoke.race.f)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates
##   the conditional random effect variances.
##   Solution: Respecify fixed structure!
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.430
##   Unadjusted ICC: 0.408
```

Use correlation - can't rule out other possibilities

```r
df_mod_base <- df_mod %>%
  filter(wave == 0)
df_mod_1 <- df_mod %>%
  filter(wave == 1)
df_mod_2 <- df_mod %>%
  filter(wave == 2)
df_mod_3 <- df_mod %>%
```

```
  filter(wave == 3)

df_mentalworry <- data.frame(mw_base = df_mod_base$mental_worry,
                             mw_1 = df_mod_1$mental_worry,
                             mw_2 = df_mod_2$mental_worry,
                             mw_3 = df_mod_3$mental_worry)

cor(df_mentalworry)
```

```
##           mw_base      mw_1      mw_2      mw_3
## mw_base 1.0000000 0.4211421 0.4089502 0.3084425
## mw_1    0.4211421 1.0000000 0.4639576 0.4249891
## mw_2    0.4089502 0.4639576 1.0000000 0.4507795
## mw_3    0.3084425 0.4249891 0.4507795 1.0000000
```