

A graphical exploration of fixed effects panel models

Matthew Alampay Davis

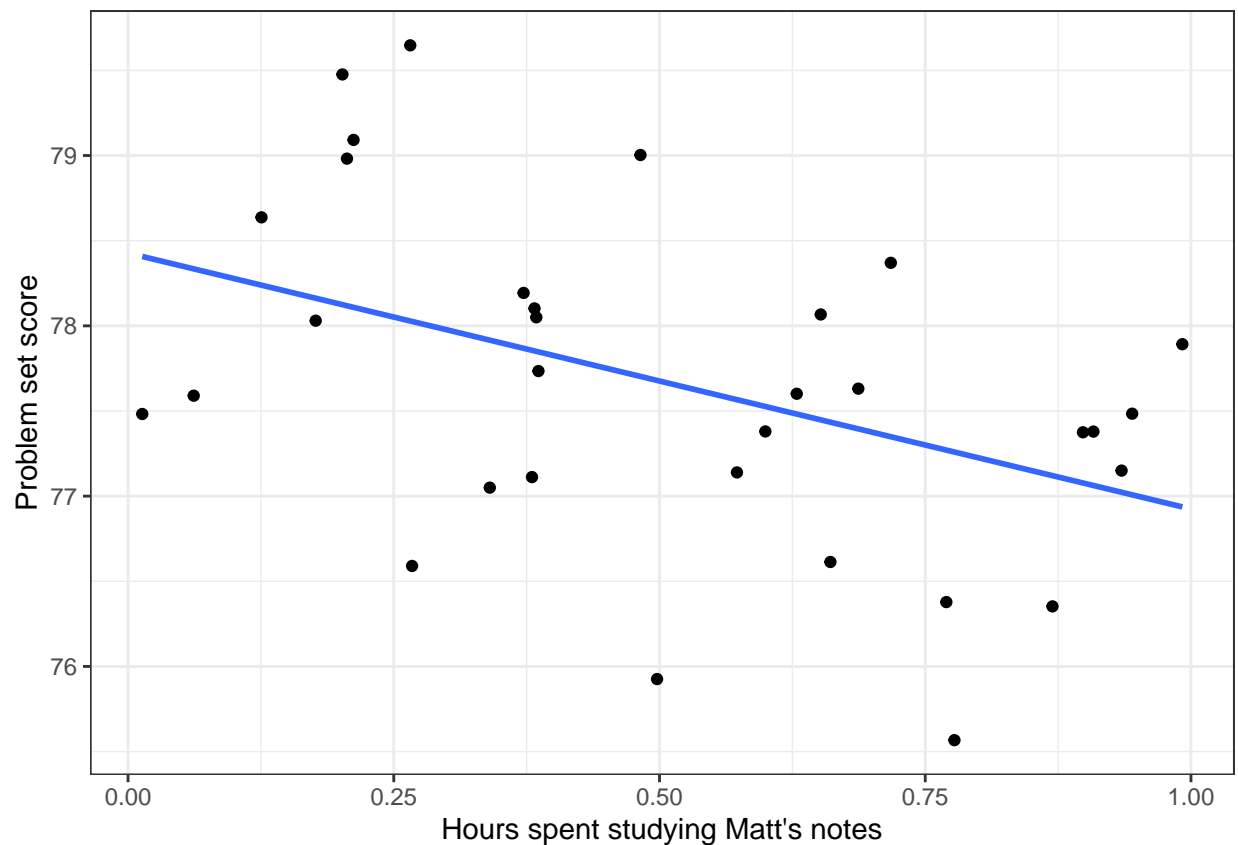
December 17, 2021

This note is just to demonstrate what fixed effects panel models do and how they differ from pooled OLS estimation. If you find it just adds to your confusion, then that's my fault and feel free to disregard. I've omitted most code but what is visible doesn't matter; you only need to follow the discussion and the graphs.

Suppose I have 32 datapoints. This (fake) dataset has two variables: $Notes_i$ is the number of hours student i spent studying my recitation notes and $Score_i$ is the score they received in the corresponding problem set:

```
ggplot(test, aes(x = study, y = score)) +  
  theme_bw() +  
  geom_smooth(method = 'lm', se = F) +  
  ylab('Problem set score') + xlab("Hours spent studying Matt's notes") +  
  geom_point()
```

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



This plot seems to suggest that studying my notes actually makes students perform worse. Indeed, when we run a basic regression, we get a negative effect significant at the 5% level:

```
lm(score ~ study, test) %>% summary

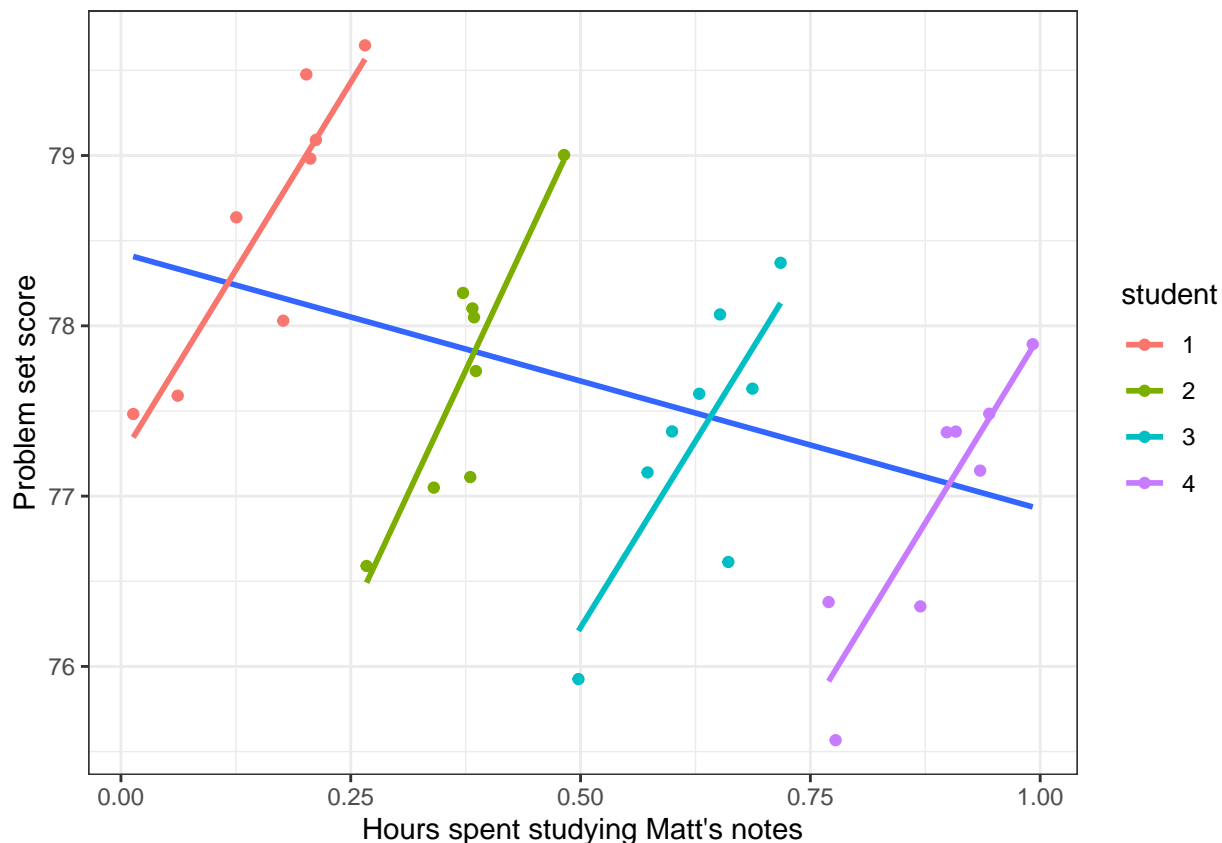
##
## Call:
## lm(formula = score ~ study, data = test)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7537 -0.7504  0.1636  0.5120  1.6187
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  78.4273     0.3283  238.867  <2e-16 ***
## study       -1.5029     0.5623   -2.673   0.012 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8957 on 30 degrees of freedom
## Multiple R-squared:  0.1923, Adjusted R-squared:  0.1654
## F-statistic: 7.143 on 1 and 30 DF,  p-value: 0.01205
```

This regression suggests that for every additional hour a student spends studying my notes, their expected problem set score falls by 1.5 points. Obviously, this interpretation doesn't make sense given that I'm such a good TA. How did this happen?

It turns out that our data isn't cross-sectional after all. The 32 observations actually correspond to just four students and their performances on the eight problem sets. Thus we can say the dataset has a panel structure where we have $T = 8$ periods corresponding to the eight problem sets and $N = 4$ students as different entities. If we color the points by student, we get the following:

```
ggplot() +
  theme_bw() +
  geom_smooth(data = test, aes(x = study, y = score), method = 'lm', se = F) +
  ylab('Problem set score') + xlab("Hours spent studying Matt's notes") +
  geom_point(data = test, aes(x = study, y = score, color = student)) +
  geom_smooth(data = test, aes(x = study, y = score, color = student), se = F, method = 'lm')

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



Now the data is starting to make some sense: each student individually is seeing positive performance gains from reading my notes (as seen in the similarly positive sloped student-specific lines of best fit) but if we regarded all the points as independently drawn as before, we'd infer a negative effect (as seen in the blue line of best fit) because we'd be actually be running a pooled OLS regression of panel data. This gives misleading estimates if there are student-specific omitted factors that need to be accounted for. Fixed effects panel models account for these even without having exhaustive data on those potentially relevant variables. Formally, these individual-specific omitted factors are contained in the α_i term here:

$$Score_{it} = \alpha_i + \beta Notes_{it} + u_{it}$$

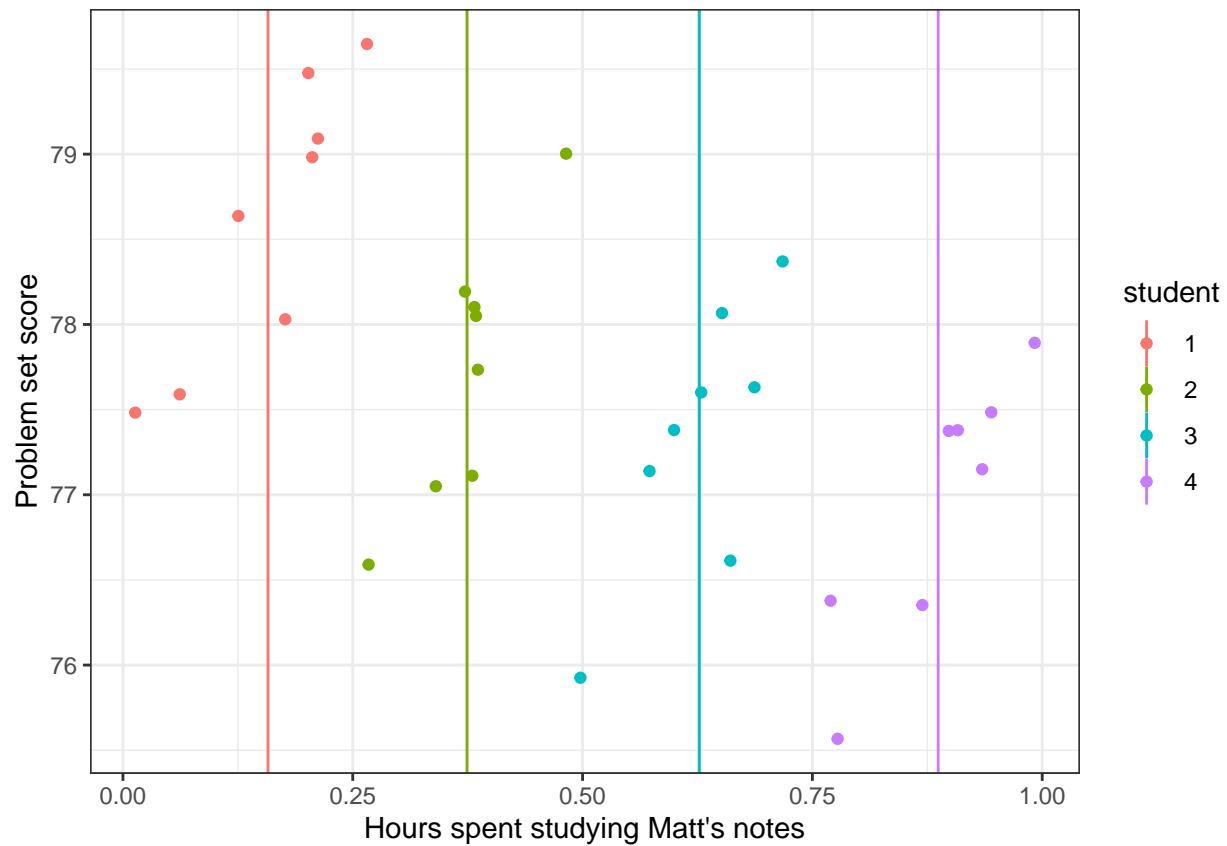
For example, α_4 might absorb the negative effect of Student 4 having a particularly annoying roommate who prevents Student 4 from focusing on his work. Even if different students experience similar benefits to studying my excellent notes (meaning that reading my notes have the same per-hour positive effect β on their problem set score), systematic differences in student-level characteristics could give each linear relationship a different intercept α_i for each student i .

Fixed effects models allow us to account for these unobservable characteristics simply by entity demeaning. The lecture and textbook guides us through the algebraic justification for this, but basically the individual effects are differenced away in the demeaning process so that the identifying variation comes from deviations from entity means: this is why it's often called the “within” estimator. We can visualize what demeaning variables accomplishes graphically here for intuition.

First, let us mark the mean time spent studying my *Notes* for each student:

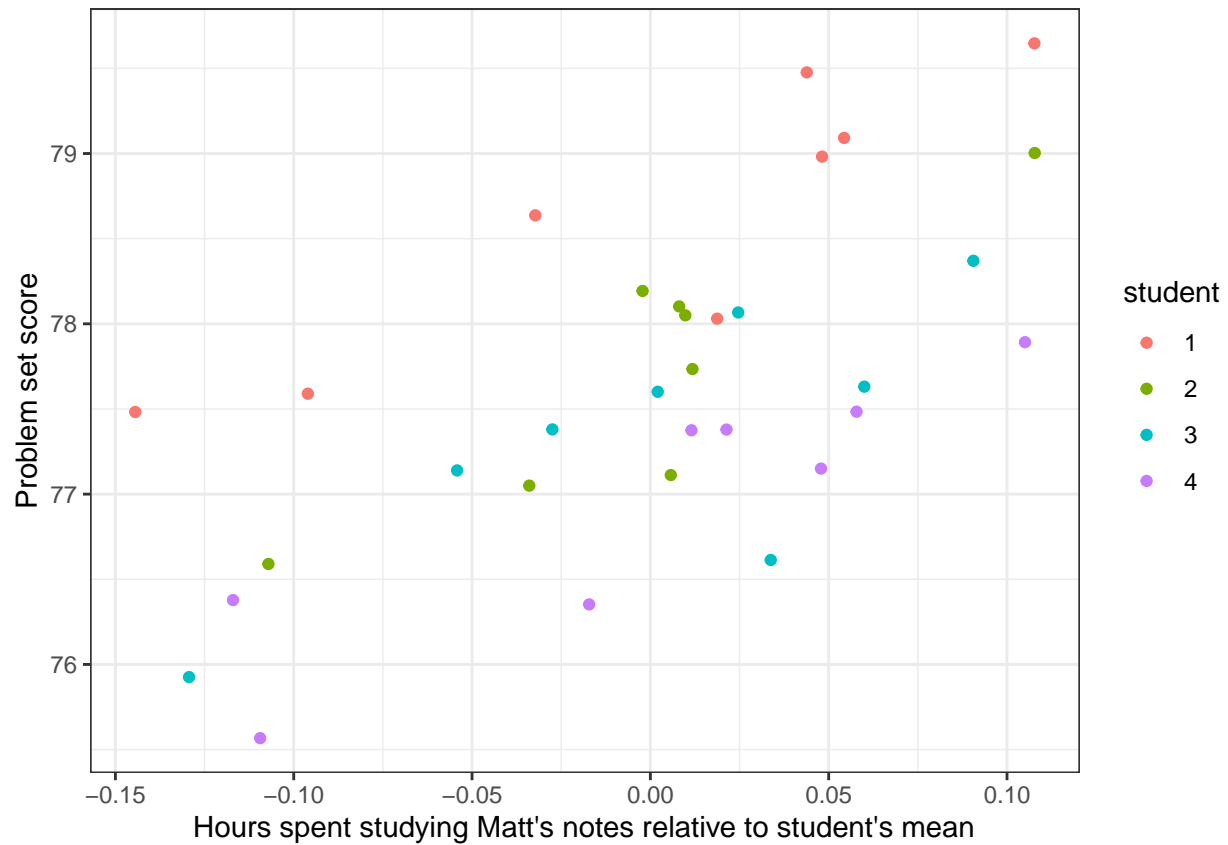
```
ggplot(test, aes(x = study, y = score, color = student)) +  
  theme_bw() +
```

```
ylab('Problem set score') + xlab("Hours spent studying Matt's notes") +
geom_point() + geom_vline(aes(xintercept = mean.study, color = student))
```



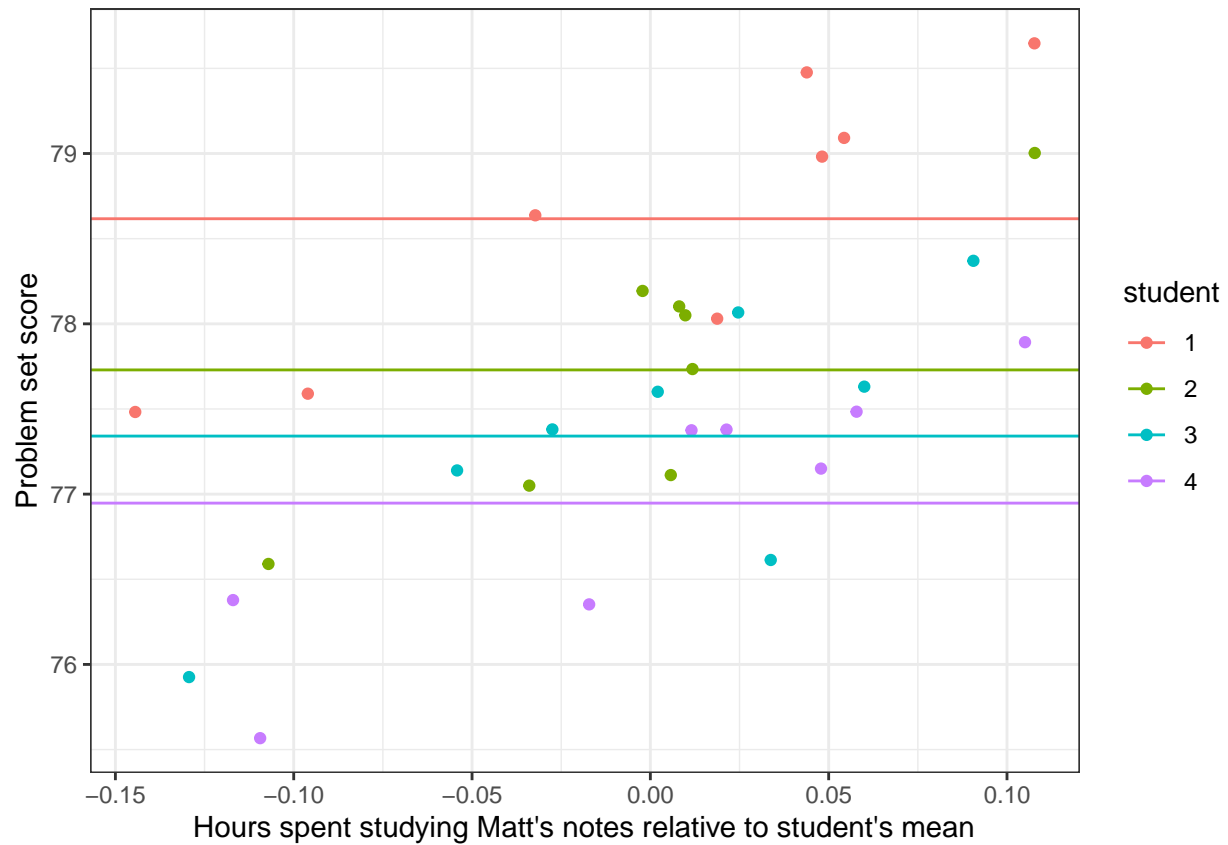
Then, for each observation, subtract their mean hours from their student-specific mean hours so that each student's hours are mean zero:

```
ggplot(test, aes(x = study-mean.study, y = score, color = student)) +
  theme_bw() +
  ylab('Problem set score') + xlab("Hours spent studying Matt's notes relative to student's mean") +
  geom_point()
```



Now for the dependent variable. Let's mark the mean problem set score for each student:

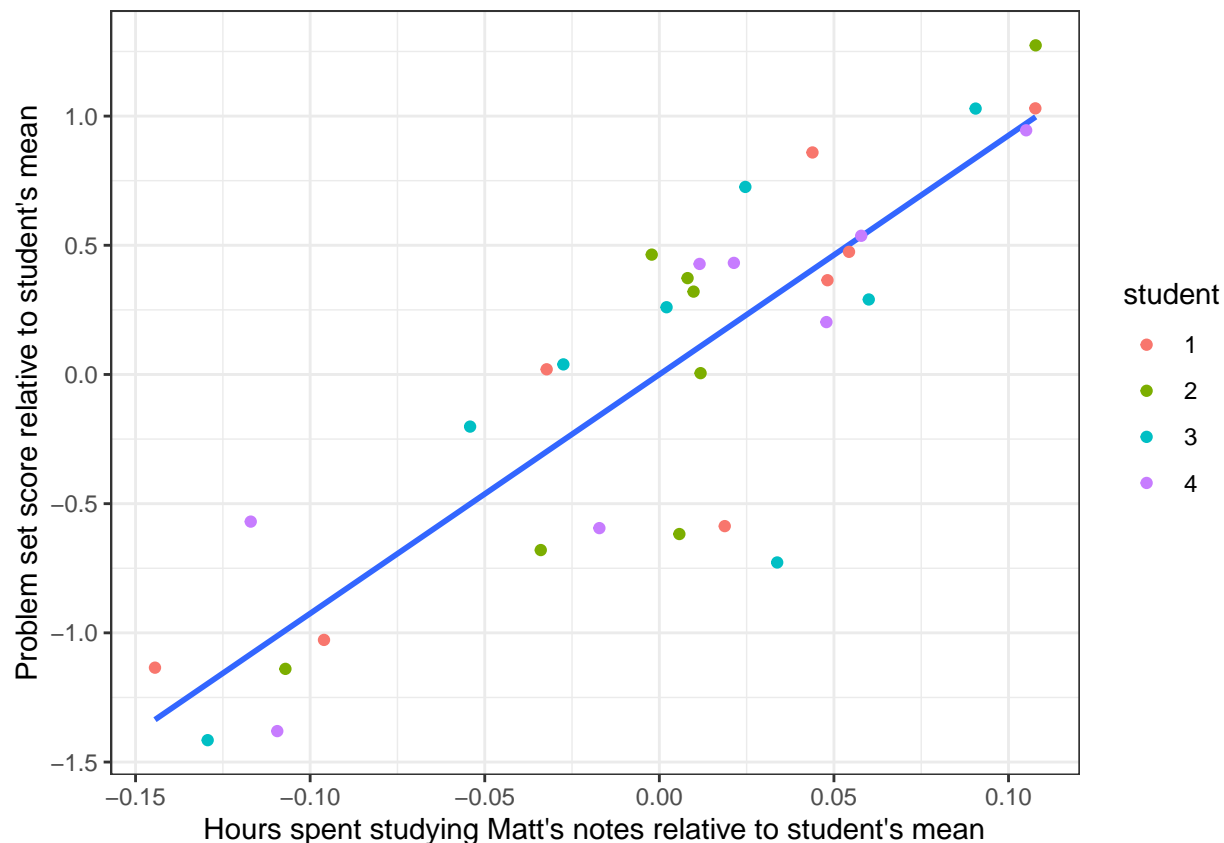
```
ggplot(test, aes(x = study-mean.study, y = score, color = student)) +
  theme_bw() +
  ylab('Problem set score') + xlab("Hours spent studying Matt's notes relative to student's mean") +
  geom_point() + geom_hline(aes(yintercept = mean.score, color = student))
```



Then as before, for each observation, subtract their student-specific mean score so the vertical axis is now relative to their student-specific means:

```
ggplot(test, aes(x = study-mean.study, y = score-mean.score)) +
  theme_bw() +
  ylab("Problem set score relative to student's mean") + xlab("Hours spent studying Matt's notes relative to student's mean") +
  geom_smooth(method = 'lm', se = F) +
  geom_point(aes(color = student))
```

'geom_smooth()' using formula 'y ~ x'



Aha. Now when we run the OLS regression on this demeaned data, we can get a sensible estimate of the effect of studying my notes.

```
test %<>% mutate(score.demeaned = score-mean.score,
                  study.demeaned = study-mean.study)
lm(score.demeaned ~ study.demeaned, data = test) %>% summary
```

```
##
## Call:
## lm(formula = score.demeaned ~ study.demeaned, data = test)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03981 -0.22407  0.01793  0.29445  0.51302
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -8.304e-16  6.880e-02   0.000      1
## study.demeaned  9.250e+00  9.982e-01   9.267 2.61e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3892 on 30 degrees of freedom
## Multiple R-squared:  0.7411, Adjusted R-squared:  0.7325
## F-statistic: 85.87 on 1 and 30 DF, p-value: 2.614e-10
```

We might also prefer to run this regression as a panel model so that we can also estimate the fixed effects

themselves:

```
lm(score ~ study + factor(student), data = test) %>% summary

##
## Call:
## lm(formula = score ~ study + factor(student), data = test)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03981 -0.22407  0.01793  0.29445  0.51302
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    77.1571     0.2205  349.929 < 2e-16 ***
## study           9.2503     1.0522   8.791 2.09e-09 ***
## factor(student)2 -2.8902     0.3065  -9.429 4.94e-10 ***
## factor(student)3 -5.6166     0.5346 -10.505 4.88e-11 ***
## factor(student)4 -8.4136     0.7941 -10.596 4.04e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4103 on 27 degrees of freedom
## Multiple R-squared:  0.8475, Adjusted R-squared:  0.8249
## F-statistic: 37.51 on 4 and 27 DF,  p-value: 1.173e-10
```

The coefficients in the two regressions agree: for every hour spent studying my notes, a student's expected problem set score increases by 9.25 points and that this estimate is very significant. Note however that the first regression has standard errors that are a bit too small because when we manually demean the variables ourselves, we are exploiting information given by the student variable: effectively, we are manually using the student's identity as a control variable.

Also note the coefficients on the fixed effects. Student 1, the highest performing student, is the omitted case (who has an expected problem set score of 77.2). The fixed effect estimates give the relative expected problem set scores of the other students with students 2-4 decreasing in expected problem set score, corresponding to systematic differences in implied intercepts. The p-values indicate these intercepts are significant, meaning there is strong evidence to reject the null hypothesis that the other students are expected to receive the same problem set score as Student 1 for the same level of studying.

Note that all the above analysis describes entity fixed effects. We could perform a similar exercise to capture time fixed effects μ_t if we suspect the presence of time-specific (i.e., problem-set specific) characteristics of each observation. For example, a particularly difficult problem set would cause trouble for all students relative to the average problem set while still increasing with exposure to my notes. And of course we could include both entity and time fixed effects together to account for both sources of systematic difference.

Hopefully this was helpful but if not, feel free to disregard. The main point is that if anyone does poorly on the exam, even if it might seem like it's my fault it's really not. Best of luck!