### Education 265: Problem Set 8 (wk9)

**Due:** Friday, March 5, 2020 at 1:00 p.m. in the class Canvas folder (Assignments tab 🡪 Problem Set 8).

* Your answer sheet as a **Word document** (titled “**LastName\_PS7**”)
* Type your answers in **bold directly** under the questions
* Your do-file (titled “**LastName\_PS8\_Code.do**”).

**Objectives:** Inspired by ECCRN and Duncan (2003) childcare quality paper, we will run 3 sets of models: level models, simple change, and residualized change. We won’t go as far as run instantenous change models or accumulated inputs model. We will examine the relationship between time spent on academic work and academic achivement.

1. Run level model, simple change model, and resdualized change models
2. Use fixed effets

**Note:** Almost all of the data processing for this problem set has been done during lab. We will be using the outcomes and predictor variables that we created and use the same set of control variables for our problem set. It should take you less than 30 minutes to add the required the code for question 2. If it is taking you more than an hour, let me know.

1. **Set up your do-file**. Start by setting the path directory (use globals or just the *cd* command) so that any data files are saved to the right folder. Call the data into Stata (*use* ***for****\_****ps8.dta****, clear*).
2. **Create a complete regression table** with 12 columns with all the notes,so that a naïve reader skimming your paper (and knows regression) can understand the table. (This will be worth a full point—half the PS). I have provided you with a table shell with the correct N and R2 below.

*Some hints*:

* 1. *You need to set the global for the control variables (like we did in lab)*
  2. *The code you need for this question is nearly identical to the code we ran for the lab example.*
  3. *You will have to create a new outcome variable for the simple change models (like we did in lab)*
  4. *Your table title, column headers, row headers, and table notes should be completley understandable and self constaned (like we have done in previous problem sets but I’m explicitly grading you on this so make this precise, accurate, and complete.)*
  5. *You should tell your reader what they need to know about the outcome, the predictor, and the control vriables. For example, you created a new outcome variable for simple change models-your reader should know about this or they will assume all the columns use the same outcome.*
  6. *You will also have to be clear about your fixed effects as there can be many school fixed effects and you chose a specific one (as talked about in lab)*
  7. *You have to talk about standard errors, too. Don’t worry about explaining clustering SE, just say that you did it.*

1. **Interpret the results**
   1. Interpret the main predictor estimate in column (6)

**Every 1 SD deviation increase in minutes spent weekly on math and reading in spring of kindergarten is associated with a 0.030 SD decrease in the academic achievement growth from spring of kindergarten to 1st grade, controlling for race, gender, age, mother’s age at first birth, SES measures, mother’s eligibility for WIC, and mother’s education level, with a significant p-value of <0.001.**

* 1. What happens to the coefficient when you add schol fixed effects? See columns (6) and (7).

**By adding the school fixed effects, the coefficient increases from -0.030 to -0.001 and the standard error increases from 0.006 to 0.009, and the coefficient isn’t statistically significant anymore with the p-value of 0.899.**

* 1. What does the change or the lack of change from adding school fixed effects suggest about the relationship between first-grade academic achivement and kindergarten time spent on academic work?

**The change in the coefficient means that taking differences among different schools into account and looking at the relationship of weekly time spent and achievemnts in first grade within each school would make the relationship less negative and the coefficient unsignificant which can suggest a difference among the population of schools in general.**

* 1. Having run this analysis and looking across all the models you have run, what “conclusion” do you come to about the relationship between first grade academic time and spring of first grade achievement? (2-5 sentences)

**First grade academic time has although small, a significant coefficient in predicting spring of first grade without the fixed effects for schools, adding fixed effects makes the coefficient insignificant which can show us that coefficients were explainable by differences among schools to some extent. However, the negative coefficient specifically in simple models is not intuitive at all and doesn’t fit with our preception of the relationship, since we expect achievement to have a positive relationship with the weekly time spent. So this might be due to the fact that the teacher reported values for time spent are not accurate and teachers may not have a good interpretation on that. It might also be a result of lack of effective control variables or also omitted variable bias in the variables and experiment. Another issue might be due to the grouping of time spent in 30 minutes intervals that may impact the relationship between them.**

* 1. Think about what questions/concerns/hypotheses you had as you answered the previous quesiton. Discuss **either**
     1. additional analyses you can run with the class dataset to address any of those questions, concerns, or hypotheses; and what you expect to find **OR**
     2. additional analyses you can run with the class dataset to elaborate on the relationship between time on academic work and academic achivement in any interesting way; and what you expect to find

**I believe my answer to d and f cover my answer to this question as well!**

* 1. **(Bonus)** Run one of those analyses, and report on what you found.

**To explore a potential missjudgement from teachers in time reports, I experimented with teacher’s years experience teaching kindergarten and its potential effect on the relationship. I added an interaction term interacting acamdemic time spend with whether the teacher has higher than 10 years experience teaching kindergarten. Results show that there is a difference although very tiny between the judgement of teachers with higher than 10 years experience with other teachers which can be a sign of a potential teacher bias in this experiments. More detail on my analysis is available in the code file.**

**I also added a new variable which is the change in teacher reported time spent from spring of kindergarten to 1st grade and explored a potential change in the relationship, which I believe would be the same as instantaneous change model. The relationship was again associated with some change in the coefficient, and the relationship change from negative to positive.**

**Therefore, my conclusion is the models here might be susceptible to bias or lack of enough control variables, while the choice of the model can also be inflential in capturing the relationship between two different variables.**

1. What does clustering standard errors at the school level appear to do, if anything? You do not have to describe conceptually what’s happening, just descsribe what is change, if anything is changing.

**Coefficients stay the same, however standard errors change.**

**use 3 significant digits after the decimal (no more, no less)**

Table 1. Coefficients from Regression of Academic Achievement Outcome on Academic Time Spent In Spring of Kindergarten, in Three Different Logitudinal Models with also Inclusion of Control Variables and School Fixed Effects

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Level Models | | | | Simple Change Models | | | | Residualized Change Models | | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Academic Time - Kindergarten (standardized) | -0.047\*\*\* | 0.019\* | -0.010 | -0.010 | -0.032\*\*\* | -0.030\*\*\* | -0.001 | -0.001 | -0.035\*\*\* | -0.019\*\*\* | -0.003 | -0.003 |
| Spring of Kindergarten Achievement |  |  |  |  |  |  |  |  | 0.810\*\*\* | 0.763\*\*\* | 0.776\*\*\* | 0.776\*\*\* |
| *Controls* |  | Inc. | Inc. | Inc. |  | Inc. | Inc. | Inc. |  | Inc. | Inc. | Inc. |
| *School Fixed Effects* |  |  | Inc. | Inc. |  |  | Inc. | Inc. |  |  | Inc. | Inc. |
| *Cluster standard error* |  |  |  | Inc. |  |  |  | Inc. |  |  |  | Inc. |
| Observations | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 |
| R-squared | 0.002 | 0.242 | 0.370 | 0.370 | 0.003 | 0.022 | 0.220 | 0.220 | 0.675 | 0.688 | 0.749 | 0.749 |

*Note*. \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Control variables are race, gender, age, mother’s age at first birth, SES measures, mother’s eligibility for WIC, and mother’s education level.

Table 1. Coefficients from Regression of Academic Achievement Outcome on Academic Time Spent In Spring of Kindergarten, in Three Different Logitudinal Models with also Inclusion of Control Variables and School Fixed Effects

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Level Models | | | | Simple Change Models | | | | Residualized Change Models | | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Academic Time - Kindergarten (standardized) | -0.047\*\*\* | 0.019\* | -0.010 | -0.010 | -0.032\*\*\* | -0.030\*\*\* | -0.001 | -0.001 | -0.035\*\*\* | -0.019\*\*\* | -0.003 | -0.003 |
|  | (0.010) | (0.009) | (0.014) | (0.015) | (0.006) | (0.006) | (0.009) | (0.010) | (0.006) | (0.006) | (0.009) | (0.009) |
| Spring of Kindergarten Achievement |  |  |  |  |  |  |  |  | 0.810\*\*\* | 0.763\*\*\* | 0.776\*\*\* | 0.776\*\*\* |
|  |  |  |  |  |  |  |  |  | (0.006) | (0.007) | (0.007) | (0.011) |
| *Controls* |  | Inc. | Inc. | Inc. |  | Inc. | Inc. | Inc. |  | Inc. | Inc. | Inc. |
| *School Fixed Effects* |  |  | Inc. | Inc. |  |  | Inc. | Inc. |  |  | Inc. | Inc. |
| *Cluster standard error* |  |  |  | Inc. |  |  |  | Inc. |  |  |  | Inc. |
| Observations | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 | 9440 |
| R-squared | 0.002 | 0.242 | 0.370 | 0.370 | 0.003 | 0.022 | 0.220 | 0.220 | 0.675 | 0.688 | 0.749 | 0.749 |

*Note*. \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Control variables are race, gender, age, mother’s age at first birth, SES measures, mother’s eligibility for WIC, and mother’s education level.

* I didn’t know if I have to include standard errors or not so I made 2 tables one with standard errors and one without them