### Education 265: Problem Set 4

**Due:** Friday, January 29, 2020 at 1:00 p.m. in the class Canvas folder (Assignments tab 🡪 Problem Set 4). Turn in

* Your answer sheet as a Word document (titled “**LastName\_PS4**”)
* Type your answers in **bold directly** under the questions
* Your do-file (titled “**LastName\_PS4\_Code.do**”).

**Objectives:**

1. We will practice quadratics with STATA (as seen in the other assignment and in lecture)
2. We will attempt to replicate Table 2 from Fryer and Levitt (2006) and practice interpreting coefficients (as we have done before).
3. We will then practice using interactions (something new). Specifically, we will investigate the relation between spring of kindergarten math achievement, gender, and kindergarten-entry characteristics .

**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 same folder. Call the data into Stata (*use for\_ps4.dta, clear*).

1. The first part of this problem set looks at non-linear regressions by replicating\* the Fryer and Levitt (2006) results.

\*Our results will differ slightly because we are not going to weight the regression models. This should be relatively easy, as you have done this for the reading response assignment and we went over this in lecture.

1. **Quadratic**
2. We will (but not yet) run the following two regression models, write *a hypothesis for each (that is, what do you think the coefficients on books and books-squared will be and why do you think so?)*:
   1. Standardized kindergarten entry math achievement (‘zmath1) as predicted by the number of books in the home;

**I believe they should have a positive relationship, but I don’t think the number of books would be a good predictor for test scores anyway.**

* 1. Standardized kindergarten entry math achievement (‘zmath1) as predicted by the number of books in the home and the number of books in the home squared.

**For the same reason as (i), I think the coefficient for number of books and number of books squared would be positive because I believe there should be a positive relationship between entry math scores and number of books, but I don’t think the coefficients would be significant since the number of books variables may not be a good predictor for kindergarten entry math scores.**

1. Now run model 6a (i). and interpret the result in two sentences or less.

**By adding a book to the books in a home of a kindergartener, their math scores would increase by 0.005 SD.**

1. Now run model 6a (ii). and interpret the result in three sentences or less.
   * Make sure to talk about: Is the relationship “U”-shape or an inverted “U”-shape? Where (i.e., how many books) is the maximum/minimum of the function? How do you know that?

**Math = -.65441 + .01504 \* (Books) - .00004 \* (Books ^ 2)**

**Since the coefficient of Books squared is negative, it means that the parabola is inverted U-shaped and there is a maximum. We can calculate the value of books that gives us the highest math score by -b/2a. The formula is derived from finding a solution to dy/dx = 0 equation which in the case of this equation would be .01504 – 2\*(.00004)\*Books = 0. Therefore, the maximum is:**

**x = -b/2a = -(.0150433)/(2 \* -.0000481) = 156**

**We can’t interpret a quadratic the same as linear equations. So by changing the number of books from n to n+1 the change in math score is going to be .01504 + .00004\*(2n+1), so the change in math score is dependent to the number of books which is because we have a quadratic relationship. The coefficient for the squared term is not significant compared to the linear coefficient which can tell that us that the relationship between math score and number of books might be linear.**

1. Report the adjusted R-squared for models 1a & 1b. How do they compare? **It has improved from 1a to 1b, which can suggest that adding the number of books squared as a predictor has improved the model more than it would have been by chance.**
2. Conceptual question: write a research question (and your corresponding hypothesis) that you would use a quadratic instead of a linear regression model.

**Professional athletes’ age and their fitness could have a quadratic relationship. Since they might not be in their best shape at the start of their career and as times by they reach a peak when they are in their best shape and in their prime, and as they start getting older and approaching the final years of their career their fitness level usually starts to decline.**

1. **Tables.** Now, replicate Fryer & Leavitt’s Tables 1 and 2. That is, create a descriptive table that shows the mean and standard deviation of our analysis variables (kindergarten, first, and third grade standardized math scores), in addition to control variables (gender, age in months, birth weight in ounces, continuous SES measure, mom over age 30 at first birth, mom teenager (<= 19 years old) at first birth, mom receives WIC benefits, and number of children’s books) by full sample, White, Black, Hispanic, and Asian American participants .

Next, create a regression table that predicts racial achievement gaps in math over the first four years of school (fall K, spring K, spring 1st, spring 3rd) using only race dummies and using race dummies plus the full set of controls listed above (8 regressions total). Paste you descriptive and regression table into your answer sheet; see Table 1 & 2 at the end of this document for example. My tables are in blue text, yours should be in black text.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table1. Descriptive Statistics of All Analysis Variables by Race | | | | | | | | | | | |
|  | Full sample |  | White |  | Black |  | Hispanic |  | Asian |  | |
|  | mean | SD | mean | SD | mean | SD | mean | SD | mean | SD | |
| Math Test Scores | | | | | | | | | | | |
| Std: Math- Kindergarten Entry | 0.112 | 1.026 | 0.333 | 1.035 | -0.338 | 0.732 | -0.381 | 0.816 | 0.454 | 1.19 | |
| Std: Math- Kindergarten Spring | 0.108 | 1.012 | 0.321 | 1.005 | -0.385 | 0.793 | -0.339 | 0.858 | 0.44 | 1.12 | |
| Std: Math- First Grade Spring | 0.098 | 0.999 | 0.311 | 1 | -0.449 | 0.795 | -0.27 | 0.873 | 0.335 | 1.009 | |
| Std: Math- Third Grade Spring | 0.074 | 0.985 | 0.291 | 0.925 | -0.594 | 0.904 | -0.257 | 0.931 | 0.411 | 0.983 | |
| Child and Family Characteristics | | | | | | | | | | | |
| Age (in months) | 68.433 | 4.254 | 68.799 | 4.254 | 68.03 | 4.158 | 67.674 | 4.205 | 67.646 | 3.935 | |
| Birth Weight (oz) | 111.707 | 20.839 | 113.714 | 20.381 | 104.459 | 21.602 | 110.482 | 21.275 | 107.552 | 17.909 | |
| Female | 0.498 | 0.5 | 0.495 | 0.5 | 0.504 | 0.5 | 0.498 | 0.5 | 0.538 | 0.499 | |
| Number of children's books | 78.52 | 60.124 | 97.496 | 58.736 | 41.452 | 41.337 | 44.093 | 47.268 | 55.57 | 52.656 | |
| Mother over 30 at first birth | 0.135 | 0.342 | 0.162 | 0.369 | 0.057 | 0.231 | 0.08 | 0.272 | 0.224 | 0.418 | |
| Mother a teenager at first birth | 0.221 | 0.415 | 0.145 | 0.352 | 0.461 | 0.499 | 0.332 | 0.471 | 0.083 | 0.276 | |
| Socioeconomic status measure | 0.087 | 0.781 | 0.287 | 0.718 | -0.341 | 0.717 | -0.37 | 0.706 | 0.365 | 0.828 | |
| Mother receives WIC benefits | 0.352 | 0.477 | 0.223 | 0.416 | 0.721 | 0.449 | 0.547 | 0.498 | 0.256 | 0.437 | |
| Observations | 10605 |  | 6524 |  | 1220 |  | 1845 |  | 446 |  | |
| *SD = Standard Deviation, Note. The sample is restricted to those with no missing data on any of the variables presented. Test scores are normalized IRT scores, standardized to a mean of 0 and SD of 1 in the full, unrestricted sample. Standard deviations are only presented for continuous variables.*   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Table 2. Estimated Racial Achievement Gap Over the First Four Years of School Math (Replication of Fryer and Levitt 2006) | | | | | | | | | |  |  |  |  |  |  |  |  |  | |  | K- Entry | Spring K | Spring 1st | Spring 3rd | K- Entry | Spring K | Spring 1st | Spring 3rd | |  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | |  |  |  |  |  |  |  |  |  | | Black | -0.671\*\*\* | -0.706\*\*\* | -0.760\*\*\* | -0.885\*\*\* | -0.126\*\*\* | -0.214\*\*\* | -0.306\*\*\* | -0.412\*\*\* | |  | (0.030) | (0.030) | (0.030) | (0.029) | (0.030) | (0.030) | (0.030) | (0.029) | |  |  |  |  |  |  |  |  |  | | Hispanic | -0.714\*\*\* | -0.660\*\*\* | -0.581\*\*\* | -0.548\*\*\* | -0.213\*\*\* | -0.207\*\*\* | -0.167\*\*\* | -0.129\*\*\* | |  | (0.026) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | |  |  |  |  |  |  |  |  |  | | Asian | 0.121\* | 0.119\* | 0.024 | 0.120\*\* | 0.245\*\*\* | 0.233\*\*\* | 0.123\*\* | 0.195\*\*\* | |  | (0.048) | (0.047) | (0.047) | (0.045) | (0.043) | (0.043) | (0.044) | (0.042) | |  |  |  |  |  |  |  |  |  | | Other | -0.445\*\*\* | -0.417\*\*\* | -0.474\*\*\* | -0.462\*\*\* | -0.164\*\*\* | -0.165\*\*\* | -0.244\*\*\* | -0.228\*\*\* | |  | (0.043) | (0.042) | (0.042) | (0.041) | (0.038) | (0.038) | (0.039) | (0.038) | | **Controls** | | | | | | | | | |  | | | | | | | | | |  |  |  |  |  |  |  |  |  | | Age (in months) |  |  |  |  | 0.062\*\*\* | 0.055\*\*\* | 0.041\*\*\* | 0.022\*\*\* | |  |  |  |  |  | (0.002) | (0.002) | (0.002) | (0.002) | |  |  |  |  |  |  |  |  |  | | Birth Weight (oz) |  |  |  |  | 0.003\*\*\* | 0.003\*\*\* | 0.002\*\*\* | 0.003\*\*\* | |  |  |  |  |  | (0.000) | (0.000) | (0.000) | (0.000) | |  |  |  |  |  |  |  |  |  | | Female |  |  |  |  | -0.003 | -0.037\* | -0.096\*\*\* | -0.188\*\*\* | |  |  |  |  |  | (0.017) | (0.017) | (0.017) | (0.017) | |  |  |  |  |  |  |  |  |  | | Number of children's books |  |  |  |  | 0.002\*\*\* | 0.002\*\*\* | 0.002\*\*\* | 0.001\*\*\* | |  |  |  |  |  | (0.000) | (0.000) | (0.000) | (0.000) | |  |  |  |  |  |  |  |  |  | | Mother over 30 at first birth |  |  |  |  | 0.166\*\*\* | 0.108\*\*\* | 0.090\*\*\* | 0.072\*\* | |  |  |  |  |  | (0.026) | (0.026) | (0.026) | (0.025) | |  |  |  |  |  |  |  |  |  | | Mother a teenager at first birth |  |  |  |  | -0.117\*\*\* | -0.111\*\*\* | -0.107\*\*\* | -0.120\*\*\* | |  |  |  |  |  | (0.023) | (0.023) | (0.023) | (0.022) | |  |  |  |  |  |  |  |  |  | | Socioeconomic status measure |  |  |  |  | 0.349\*\*\* | 0.324\*\*\* | 0.301\*\*\* | 0.332\*\*\* | |  |  |  |  |  | (0.014) | (0.014) | (0.014) | (0.013) | |  |  |  |  |  |  |  |  |  | | Mother receives WIC benefits |  |  |  |  | -0.201\*\*\* | -0.173\*\*\* | -0.166\*\*\* | -0.202\*\*\* | |  |  |  |  |  | (0.022) | (0.022) | (0.022) | (0.021) | |  |  |  |  |  |  |  |  |  | | Constant | 0.333\*\*\* | 0.321\*\*\* | 0.311\*\*\* | 0.291\*\*\* | -4.502\*\*\* | -3.944\*\*\* | -2.902\*\*\* | -1.590\*\*\* | |  | (0.012) | (0.012) | (0.012) | (0.012) | (0.147) | (0.149) | (0.151) | (0.146) | |  |  |  |  |  |  |  |  |  | | Observations | 10605 | 10605 | 10605 | 10605 | 10605 | 10605 | 10605 | 10605 | | R-squared | 0.098 | 0.095 | 0.092 | 0.111 | 0.304 | 0.266 | 0.231 | 0.259 | |  |  |  |  |  |  |  |  |  | | *Standard errors in parentheses* | | |  |  |  |  |  |  | | *Note. Test scores are standardized IRT score. Non-Hispanic Whites are the omitted race category. Standard errors are in parentheses.* | | | | | | | | | | *\* p<0.05 \*\* p<0.01 \*\*\* p<0.001* | | | | | | | | | | | | | | | | | | | |

*Some hints:*

* + **Remember to add tables notes**
  + Remember to properly recode missing values, if needed.
  + Fryer & Levitt used the following dummy variables: Black, Hispanic, Asian, Other race, (White as reference group), Female, Mom is over 30 at first birth, Mom receives WIC benefits, Mom is a teenager at first birth. Some variables will need to be recoded (must be 0, 1 or .) or created.
  + Use the continuous measure of SES (*wksesl*).
  + Birthweight needs to be converted from pounds to ounces.
  + Do not forget to standardize the test scores (i.e., “zmath1”, “zmath2”, “zmath4”, and “zmath5”)
  + Set an analysis sample for this question by generating a variable (“*replicate\_sample*”) that is equal to 1(you should have 10,605 cases) if a case is non-missing on all 4 math test scores, race, and each of the control variables.

1. **Graph**
   1. **Bonus:** In Stata, create 2 *clearly labeled* bar graphs (with error bars) that presents the **gap** between Whites and each of the four other race groups on standardized achievement score of the following:
      1. Spring of third grade math acheivement (model 4)
      2. Spring of third grade math acheivement with controls (model 8)
   2. **Bonus Bonus:** Combine the two bar graphs into one big graph.
2. **Interpreting the results**
   1. When control variables are added to the models, the magnitude of the coefficients on the race dummies is reduced. What does this suggest about the nature of racial achievement gaps?

**It shows that other factors have also influence the race score gap apart from the race itself. For example, from table 1 you can see that black and Hispanic students have lower SES measures on average compared to other ethnicities which can affect our judgment if we don’t have it as a control variable.**

* 1. The gap between Whites and one other race group changes drastically between kindergarten and third grade (especially after controls are added). Which group is this, and what might the change in this coefficient over time suggest?

**The gap between white and black students appears to get larger over time which can suggest that there are other factors apart from the control variables such as schooling conditions, that are affecting the widening of this gap.**

**Congrats on finishing the first part of this problem set, now we will examine interactions!**

1. **Add dummies.** Create dummy variables for mother’s education and *paste your code for this into your answer sheet.* That is, take the “**wkmomed**” variable and generate three mutually exclusive dummy variables: 1) mothers with less than a high school diploma; 2) mothers with a high school diploma, vocational school, or some college; 3) mothers that graduated from college or attended graduate school.

**tab wkmomed**

**replace wkmomed = . if wkmomed < 0**

**gen dmomed1 = .**

**replace dmomed1=0 if wkmomed !=.**

**replace dmomed1=1 if wkmomed < 3**

**tab dmomed1 , m**

**gen dmomed2 = .**

**replace dmomed2=0 if wkmomed !=.**

**replace dmomed2=1 if ( wkmomed >= 3 & wkmomed < 6)**

**tab dmomed2 , m**

**gen dmomed3 = .**

**replace dmomed3=0 if wkmomed !=.**

**replace dmomed3=1 if (wkmomed > 6 & wkmomed != . )**

**tab dmomed3 , m**

1. **Prepare controls.** Some models that we will run have control variables. *Create a global list (i.e., $interactcontrols) of the following variables:.*

* Standardized fall of kindergarten reading achievement “zread1”
* Dummies for race (omit Whites as the comparison group)—dblack, dhispanic, dasian, dother
* Mom’s education dummies that you created (omit mothers with less than a high school diploma as the comparison group)
* The following variables from the Fryer & Levitt controls from Q2: age in months, birthweight in oz, female, number of children’s books, mother over 30 at first birth, mom a teenager at first birth, socioeconomic status, mom received WIC benefits. (notice, we are not adding books2)

1. **Interactions:** Examinegender differences in the association between school-entry characteristics and end-of-kindergarten math scores.
   1. Create an analysis sample variable (“interact\_samp”) that is set to 1 (should have 14,258 cases) if students are non-missing on fall and spring measures of math and reading achievement, and non-missing on the list of control variables in Q5. *Paste your code into the answer sheet.* All models below should be restricted to this “interact\_samp”.

**gen interact\_samp = 0**

**replace interact\_samp = 1 if !mi(zmath1, zmath2, zread1, zread2, race, dfemale, over30, teenager, dwicmom, weight\_o, wksesl, age, p1chlboo)**

**tab interact\_samp**

* 1. Regress spring standardized kindergarten math achievement (zmath2) on fall standardized kindergarten math acheivement (zmath1) *separately for boys and girls* (i.e., run two regression modesl). Concsider the difference or the lack of difference in the “zmath1” coefficient between boys and girls. What can you know from comparing the coefficients and what can’t you know?

**We can understand that there is a difference between boys and girls in the linear ralationship between fall ans spring score, and fall scores is a better predictor for spring for boys, however, we can’t tell how statistically significant their difference is.**

* 1. Run the same model with boys and girls pooled together, but include an interaction between the zmath1 and dfemale. Using the outputted coefficients, write the regression equation (i.e., math equation) for boys and the equation for girls.

**Equation: zmath2 = .026 - .028 \* dfemale + .832 \* zmath1 - .027 \* zmath1Xdfemale**

**Girls: zmath2 = (.026 - .028) + (.832 - .027 ) \* zmath1**

**Boys: zmath2 = .026 + .832 \* zmath1**

* 1. Interpret the coefficient on the interaction term. What does this tell us about gender difference in the relation between fall and spring math scores?

**The coefficient on the interaction term shows the difference of linear relationship between boys and girls. So shows that for every 1 SD increase in fall kindergarten math score, girls’ spring score increase 0.27 SD less than the boys.**

* 1. Run the same model, but add our set of controls (the global you created in Q6) Paste your regression code into your answer sheet.

**reg zmath2 zmath1 $interactcontrols zmath1Xdfemale**

* 1. What happened to the coefficient produced by the interaction term when compared with the model run in 7c? **It decreased**
  2. Run a final model that includes interactions between gender and all of the controls included in 6e (this is the fully interacted model). Then run the joint F-test to determine if the relation between school entry characteristics and spring of kindergarten math achievement systematically differs between boys and girls (*Hint: we did this in lab*). Interpret the F-statistic. (*Hint: you can use the combination of global lists and loop to create the interaction variables includes).*

foreach var in $interactcontrols {

gen `var'\_girls= `var' \* dfemale

}

**F( 15, 14226) = 2.73**

**Prob > F = 0.0003**

**Since the p-value is significant, our hypothesis is true. Therefore, we can hypothesize that the relation between math spring scores, math fall scores, and other characteristics for students, is different significantly for boys and girls.**

* 1. Generate an APA-formatted regression table that includes all 5 regression models run for Q7 (i.e., 2 regressinos from 7b, 1 from 7c, 1 from 7e, and 1 from 7g). Your table should ONLY display the coefficients and standard errors produced by the following variables (when applicable): dfemale, zmath1, and the interaction between dfemale and zmath1. You don’t have to display coefficients and standard errors for all of the controls or the interactions between controls and gender (you can either delete these manually in Excel or use the “keep” option for the *esttab* command). However, you should note in your table in which models these additional controls were added. See the Table 3 as example.

Table3. Regression table of coefficients predicting Math- Kindergarten Spring

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
| Std: Math- Kindergarten Entry | 0.806\*\*\* | 0.832\*\*\* | 0.832\*\*\* | 0.735\*\*\* | 0.736\*\*\* |
|  | (0.007) | (0.006) | (0.006) | (0.008) | (0.009) |
| Female |  |  | -0.028\*\* | -0.033\*\*\* | -0.315 |
|  |  |  | (0.009) | (0.009) | (0.173) |
| Math K entry score x Female |  |  | -0.027\*\* | -0.044\*\*\* | -0.050\*\*\* |
|  |  |  | (0.010) | (0.009) | (0.014) |
| Controls |  |  |  | Inc. | Inc. |
| Interactions between control variables and whether female |  |  |  |  | Inc. |
| Constant | -0.002 | 0.026\*\*\* | 0.026\*\*\* | -0.425\*\*\* | -0.288\* |
|  | (0.006) | (0.007) | (0.007) | (0.087) | (0.118) |
| R-squared | 0.664 | 0.696 | 0.683 | 0.694 | 0.694 |
| Observations | 7056 | 7202 | 14258 | 14258 | 14258 |
| *Standard errors in parentheses* | | |  |  |  |
| *Note. Standard errors in parentheses. Inc. indicates included. Control variables include: child ethnicity, age, birthweight in ounces, number of books in the home, number of books in the home squared, whether mother was over 30 at first birth* | | | | | |
| *\* p<0.05, \*\* p<0.01, \*\*\* p<0.001"* | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1 |  | |  |  | |  |  | |  |  | |  |  | |  |
| *Descriptive Statistics of All Analysis Variables by Race* | | | | | | | | | | | | | | | |
|  | Full Sample | | | White | | | Black | | | Hispanic | | | Asian | | |
|  | M | SD | | M | SD | | M | SD | | M | SD | | M | SD | |
| *Math Test Scores* |  |  | |  |  | |  |  | |  |  | |  |  | |
| Kindergarten Entry | 0.112 | 1.026 | | 0.333 | 1.035 | | -0.338 | 0.732 | | -0.381 | 0.816 | | 0.454 | 1.190 | |
| Spring of Kindergarten | 0.108 | 1.012 | | 0.321 | 1.005 | | -0.385 | 0.793 | | -0.339 | 0.858 | | 0.440 | 1.120 | |
| Spring of First Grade | 0.098 | 0.999 | | 0.311 | 1.000 | | -0.449 | 0.795 | | -0.270 | 0.873 | | 0.335 | 1.009 | |
| Spring of Third Grade | 0.074 | 0.985 | | 0.291 | 0.925 | | -0.594 | 0.904 | | -0.257 | 0.931 | | 0.411 | 0.983 | |
| *Child and Family Characteristics* |  |  | |  |  | |  |  | |  |  | |  |  | |
| Age (in months) | 68.433 | 4.254 | | 68.799 | 4.254 | | 68.030 | 4.158 | | 67.674 | 4.205 | | 67.646 | 3.935 | |
| Birth Weight (oz) | 111.707 | 20.839 | | 113.714 | 20.381 | | 104.459 | 21.602 | | 110.482 | 21.275 | | 107.552 | 17.909 | |
| Female | 0.498 |  | | 0.495 |  | | 0.504 |  | | 0.498 |  | | 0.538 |  | |
| Number of children's books | 78.520 | 60.124 | | 97.496 | 58.736 | | 41.452 | 41.337 | | 44.093 | 47.268 | | 55.570 | 52.656 | |
| Mother over 30 at first birth | 0.135 |  | | .162 |  | | 0.057 |  | | 0.080 |  | | 0.224 |  | |
| Mother a teenager at first birth | 0.221 |  | | .145 |  | | 0.462 |  | | 0.333 |  | | 0.083 |  | |
| Socioeconomic status measure | 0.087 | 0.781 | | 0.287 | 0.718 | | -0.341 | 0.717 | | -0.370 | 0.706 | | 0.365 | 0.828 | |
| Mother receives WIC benefits | 0.352 |  | | .223 |  | | 0.721 |  | | 0.547 |  | | 0.256 |  | |
| Observations | 10605 |  | | 6524 |  | | 1220 |  | | 1845 |  | | 446 | | |
| *Note.* The sample is restricted to those with no missing data on any of the variables presented. Test scores are normalized IRT scores, standardized to a mean of 0 and SD of 1 in the full, unrestricted sample. Standard deviations are only presented for continuous variables. | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2 |  |  | |  | |  | | |  |  | |  | |  | |
| *Estimated Racial Achievement Gap Over the First Four Years of School, Math (Replication of Fryer and Levitt, 2006)* | | | | | | | | | | | | | | | |
|  | K- Entry | | Spring K | | Spring 1st | | Spring 3rd | K- Entry | | | Spring K | | Spring 1st | | Spring 3rd |
|  | (1) | | (2) | | (3) | | (4) | (5) | | | (6) | | (7) | | (8) |
| Black | -0.671\*\*\* | | -0.706\*\*\* | | -0.760\*\*\* | | -0.885\*\*\* | -0.126\*\*\* | | | -0.214\*\*\* | | -0.306\*\*\* | | -0.412\*\*\* |
|  | (0.030) | | (0.030) | | (0.030) | | (0.029) | (0.030) | | | (0.030) | | (0.030) | | (0.029) |
| Hispanic | -0.714\*\*\* | | -0.660\*\*\* | | -0.581\*\*\* | | -0.548\*\*\* | -0.213\*\*\* | | | -0.207\*\*\* | | -0.167\*\*\* | | -0.129\*\*\* |
|  | (0.026) | | (0.025) | | (0.025) | | (0.025) | (0.025) | | | (0.025) | | (0.025) | | (0.025) |
| Asian | 0.121\* | | 0.119\* | | 0.024 | | 0.120\*\* | 0.245\*\*\* | | | 0.233\*\*\* | | 0.123\*\* | | 0.195\*\*\* |
|  | (0.048) | | (0.047) | | (0.047) | | (0.045) | (0.043) | | | (0.043) | | (0.044) | | (0.042) |
| Other | -0.445\*\*\* | | -0.417\*\*\* | | -0.474\*\*\* | | -0.462\*\*\* | -0.164\*\*\* | | | -0.165\*\*\* | | -0.244\*\*\* | | -0.228\*\*\* |
|  | (0.043) | | (0.042) | | (0.042) | | (0.041) | (0.038) | | | (0.038) | | (0.039) | | (0.038) |
| *Controls* |  | |  | |  | |  | 0.062\*\*\* | | | 0.055\*\*\* | | 0.041\*\*\* | | 0.022\*\*\* |
| Age (in months) |  | |  | |  | |  | 0.062\*\*\* | | | 0.055\*\*\* | | 0.041\*\*\* | | 0.022\*\*\* |
|  |  | |  | |  | |  | (0.002) | | | (0.002) | | (0.002) | | (0.002) |
| Birth weight (in ounces) |  | |  | |  | |  | 0.003\*\*\* | | | 0.003\*\*\* | | 0.002\*\*\* | | 0.003\*\*\* |
|  |  | |  | |  | |  | (0.000) | | | (0.000) | | (0.000) | | (0.000) |
| Female |  | |  | |  | |  | -0.003 | | | -0.037\* | | -0.096\*\*\* | | -0.188\*\*\* |
|  |  | |  | |  | |  | (0.017) | | | (0.017) | | (0.017) | | (0.017) |
| Number of kid's books |  | |  | |  | |  | 0.002\*\*\* | | | 0.002\*\*\* | | 0.002\*\*\* | | 0.001\*\*\* |
|  |  | |  | |  | |  | (0.000) | | | (0.000) | | (0.000) | | (0.000) |
| Mother over 30 at first birth |  | |  | |  | |  | 0.166\*\*\* | | | 0.108\*\*\* | | 0.090\*\*\* | | 0.072\*\* |
|  |  | |  | |  | |  | (0.026) | | | (0.026) | | (0.026) | | (0.025) |
| Mother a teenager at first birth |  | |  | |  | |  | -0.117\*\*\* | | | -0.111\*\*\* | | -0.107\*\*\* | | -0.121\*\*\* |
|  |  | |  | |  | |  | (0.023) | | | (0.023) | | (0.023) | | (0.022) |
| Socioeconomic status measure |  | |  | |  | |  | 0.349\*\*\* | | | 0.323\*\*\* | | 0.301\*\*\* | | 0.332\*\*\* |
|  |  | |  | |  | |  | (0.014) | | | (0.014) | | (0.014) | | (0.013) |
| Mother receives WIC benefits |  | |  | |  | |  | -0.201\*\*\* | | | -0.173\*\*\* | | -0.166\*\*\* | | -0.202\*\*\* |
|  |  | |  | |  | |  | (0.022) | | | (0.022) | | (0.022) | | (0.021) |
| Constant | 0.333\*\*\* | | 0.321\*\*\* | | 0.311\*\*\* | | 0.291\*\*\* | -4.502\*\*\* | | | -3.944\*\*\* | | -2.902\*\*\* | | -1.590\*\*\* |
|  | (0.012) | | (0.012) | | (0.012) | | (0.012) | (0.147) | | | (0.149) | | (0.151) | | (0.146) |
| Observations | 10605 | | 10605 | | 10605 | | 10605 | 10605 | | | 10605 | | 10605 | | 10605 |
| R-squared | 0.098 | | 0.095 | | 0.092 | | 0.111 | 0.304 | | | 0.266 | | 0.231 | | 0.259 |
| *Note.* Test scores are standardized IRT score. Non-Hispanic Whites are the omitted race category. Standard errors are in parentheses.  \*p<0.05 \*\* p<0.01 \*\*\* p<0.001 | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3. Sample table template for question 7h. | | |  | |  | |  | |  |
|  | Model 1 | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Math K entry score | XX | XX | | XX | | XX | | XX | |
|  | (XX) | (XX) | | (XX) | | (XX) | | (XX) | |
| Female |  |  | | XX | | XX | | XX | |
|  |  |  | | (XX) | | (XX) | | (XX) | |
| Math K entry score x Female |  |  | | XX | | XX | | XX | |
|  |  |  | | (XX) | | (XX) | | (XX) | |
| Controls |  |  | |  | | Inc. | | Inc. | |
| Interactions between control variables and whether female |  |  | |  | |  | | Inc. | |
| Constant | XX | XX | | XX | | XX | | XX | |
|  | (XX) | (XX) | | (XX) | | (XX) | | (XX) | |
| R-squared | XX | XX | | XX | | XX | | XX | |
| Observations | XXXXX | XXXXX | | XXXXX | | XXXXX | | XXXXX | |
| *Note*. Standard errors in parentheses. Inc. indicates included. Control variables include: child ethnicity, age, birthweight in ounces, number of books in the home, number of books in the home squared, whether mother was over 30 at first birth, whether mother was a teenager at first birth, ECLS-K derived continuous measure of socioeconomic status, whether mother received WIC benefits during pregnancy, mother’s level of education (less than high school diploma is the omitted category). \**p*<.05, \*\**p*<.01, \*\*\**p*<.001. | | | | | | | | | |