Davids_R_Script

David Fraire

5/16/2019

The following regressions show in table form will represent the effects of the Headstart program on the outcome variables of **PPVTat3**, **somecollege**, **hsgrad**. We decided to use a mixture of linear models and logit models to proceed with analysis of the effect of Headstart program on the three aforementioned dependent variables.

Table 1: Headstart effect on PPVT Scores at Age 3

	Dependent variable:			
	PPVTat3			
	(1)	(2)	(3)	(4)
headstart	−6.741***	−6.739***	−2.992***	−3.440**
	(1.054)	(1.055)	(1.010)	(1.677)
Hispanic			−8.504***	−8.274***
			(1.073)	(1.173)
Black			−12.205***	−13.120***
			(0.946)	(1.066)
Male		0.030	0.193	0.702
		(0.852)	(0.783)	(0.777)
hsgrad				3.046***
				(0.785)
FirstBorn				3.721***
				(0.784)
headstart:Black				3.125
				(2.245)
headstart:Hispanic				−1.380
				(2.717)
Constant	25.028***	25.013***	29.140***	25.691***
	(0.477)	(0.645)	(0.671)	(0.875)
Observations	984	984	984	984
R^2	0.040	0.040	0.191	0.228
Adjusted R^2	0.039	0.038	0.188	0.222
Residual Std. Error	13.348 (df = 982)	13.355 (df = 981)	12.270 (df = 979)	12.014 (df = 975)
F Statistic	40.942*** (df = 1; 982)	20.451*** (df = 2; 981)	57.909*** (df = 4; 979)	35.974*** (df = 8; 975)
Note:				*p<0.1; **p<0.05; ***p<0.01

For the previous table, we began by regressing the data for **PPVTat3** on **headstart** using a linear fit. The

effect of the headstart program in this model shows a negative effect ranging from -6.741 to -2.179 PPVT scores and remains significant even after adding variables which are very signficant. The linear model in the last column includes **Hispanic**, **Black**, **Male**, **hsgrad**, **FirstBorn**, and interaction terms for **headstart**. This very significant variables from this regression are **headstart**, **Hispanic**, **Black**, **Male**, **hsgrad**, and **FirstBorn**. The **Black** variable has a very large average effect on the outcome of PPVT scores and is the highest in magnitude. The **headstart** variable also has a significant negative effect, but smaller than those in previous regressions with less covariates. This means that participation in Headstart certainly has a negative effect on the test scores with the data provided in our dataset. Given that the children who participate in Headstart are likely to come from disadvatanged socioeconomic backgrounds, this can be explained in the following models as those who are in Headstart will likely have lower scores to start.

Table 2: Headstart effect on College Enrollment

	Dependent variable:			
	somecollege			
	(1)	(2)	(3)	(4)
headstart	0.486***	0.493***	0.269***	−0.173
	(0.055)	(0.056)	(0.058)	(0.130)
Male		−0.301***	−0.306***	−0.387***
		(0.045)	(0.045)	(0.050)
Black			0.692***	0.684***
			(0.053)	(0.069)
Hispanic			0.482***	0.500***
			(0.059)	(0.073)
LogInc_0to3				0.135***
				(0.032)
headstart:Black				0.399**
				(0.158)
headstart:Hispanic				0.223
				(0.186)
Constant	−1.292***	−1.146***	−1.414***	−2.191***
	(0.025)	(0.032)	(0.040)	(0.344)
Observations	11,470	11,470	11,470	7,485
Log Likelihood	−6,162.935	−6,140.100	−6,049.910	−4,664.154
Akaike Inf. Crit.	12,329.870	12,286.200	12,109.820	9,344.308
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Log Odds of Variable

(Intercept)	headstart
0.275	1.625

Table 4: Log Odds of Variable

(Intercept)	headstart	Male
0.318	1.637	0.740

For the effect of Headstart on college enrollment which is measured by the binary variable **somecollege**, we used a logit model to see the effects on the odds of having gone to college. Simply running a regression of **headstart** on **somecollege** shows that **headstart** has a significant effect on the outcome of having enrolled at a college. The odds of attending college are increased by a factor of 1.625. However, this effect is shown

Table 5: Log Odds of Variable

(Intercept)	headstart	Male	Black	Hispanic
0.243	1.308	0.737	1.997	1.619

Table 6: Log Odds of Variable

(Intercept)	headstart	Male	Black	Hispanic	LogInc_0to3	headstart:Black	headstart:Hispanic
0.112	0.841	0.679	1.982	1.649	1.145	1.491	1.249

to decrease and even become statistically insignificant when controlling for other significant variables such as **Male**, **Black**, **Hispanic**, and **LogInc_0to3**.

The following tables tell us the log odds of each of the regressions showing that **headstart** in the regression with all included covariates has a positive effect of a factor of 1.059 but is not a statistically significant variable after controlling for **Male**, **Black**, **Hispanic**,**LogInc_0to3** and **headstart** interaction terms, all of which are very statistically significant.

Table 7: Log Odds of Variable

(Intercept)	headstart
0.275	1.625

Table 8: Log Odds of Variable

(Intercept)	headstart	Male
0.318	1.637	0.740

Table 9: Log Odds of Variable

(Intercept)	headstart	Male	Black	Hispanic
0.243	1.308	0.737	1.997	1.619

Table 10: Log Odds of Variable

(Intercept)	headstart	Male	Black	Hispanic	LogInc_0to3	headstart:Black	headstart:Hispanic
0.112	0.841	0.679	1.982	1.649	1.145	1.491	1.249

For this regression table we regressed the following variables on **hsgrad** to indicate whether or not a child graduated from highschool, **Male**, **Black**, **Hispanic**, **LogInc_0to3**, **MothEd**. The model uses a logit link funciton which tells us in column that the effect of particpating in the Headstart program will increase the log odds of graduating highschool by .486.

Table 11: Headstart effects on High School Graduation

	Dependent variable:			
	hsgrad			
	(1)	(2)	(3)	(4)
headstart	0.486***	0.493***	0.269***	−0.158
	(0.055)	(0.056)	(0.058)	(0.130)
Male		−0.301***	−0.306***	−0.391***
		(0.045)	(0.045)	(0.050)
Black			0.692***	0.659***
			(0.053)	(0.069)
Hispanic			0.482***	0.528***
			(0.059)	(0.074)
LogInc_0to3				0.084**
				(0.035)
MothED				0.043***
				(0.011)
headstart:Black				0.387**
				(0.158)
headstart:Hispanic				0.206
				(0.186)
Constant	−1.292***	−1.146***	−1.414***	−2.190***
	(0.025)	(0.032)	(0.040)	(0.342)
Observations	11,470	11,470	11,470	7,479
Log Likelihood	−6,162.935	−6,140.100	−6,049.910	−4,650.817
Akaike Inf. Crit.	12,329.870	12,286.200	12,109.820	9,319.634
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 12: Log Odds of Variable

(Intercept)	headstart
0.275	1.625

Table 13: Log Odds of Variable

(Intercept)	headstart	Male
0.318	1.637	0.740

Table 14: Log Odds of Variable

(Intercept)	headstart	Male	Black	Hispanic
0.243	1.308	0.737	1.997	1.619

Table 15: Log Odds of Variable

(Intercept)	headstart	Male	Black	Hispanic	LogInc_0to3	MothED	headstart:Black	headstart:Hispa
0.112	0.854	0.677	1.933	1.695	1.088	1.044	1.473	1.229