ECON 672 Winter 2022 Problem Set #1

Jon Holder

January 13, 2022

Question 1

a. Persico, Postlewaite, and Silverman (2004) are examining the mechanism that influences the relationship between height and earnings gap. They note that previous studies have evaluated the influence of current height. They take advantage of the fact that height changes throughout youth before becoming relatively constant in adulthood. Persico, Postlewaite, and Silverman find that height at 16 is a predictor of future earnings controlling for vocation, family attributes and adult height.

b. Two data sets are evaluated: 1) Britain's National Child Development Survey (NCDS) and the US's National Longitudinal Survey of Youth (NLSY) 1979 cohort.

c. Table 1 shows the raw heights, change in heights and their distributions for white men in the two data sets. Persico, Postlewaite, and Silverman include this data because it is the key parameter of interest. The characteristics and distributions of these heights provide important context for the conclusion of the authors.

d. Table 2 shows various statistics for sample. It includes traditional determinants of wages include individual and family educational attainment, martial status, and number of siblings. Finally, Table 2 contains the ln(wages) which is the dependent variable.

e. Table 3 shows various OLS estimates for effect on ln(wages) using the controls from Table 3. The estimates show a positive return to ln(wages) for full-time white male workers based on reported height during youth. The reduced form specification shows a positive and statistically significant positive return to wages based on adult height. However, the return due to adult height diminishes and is no longer statistically significant when youth height is added to the specification.

f. Persico, Postlewaite, and Silverman explore the relationship between height at various stages of development age 7, 11 and 16 in addition to adult height. The table shows that the height premium is only statistically economically significant at age 16 when all heights and family characteristics are included.

g. The primary conclusion of the paper is that height when a teenager is the primary contributor to adult wage differential rather than adult height. The authors tested for height as indicator of several factors including health, availability of external resources, and intergenerational factors. There appears to be an identified mechanism that associated with height

that produces a future wage premium. The authors suggest this premium is gained through greater social participation of the taller respondents. This premium is present independent of career choices.

Additionally, the wage premium is high enough that receiving growth hormone treatments to increase teen height has a positive return.

Question 2

b. (i) NLSY contained 3 subsamples: a cross-sectional sample that represented noninstitutionalized civilian population, a sample that oversampled civilian Latino, black and other economically disadvantage civilians, and a sample representing the population that was in the military. After the 1984, most members of the military where ineligible for interview so 201 members of that cohort were randomly selected to remain. After the 1990 interview, the economically disadvantaged white population was removed from the sample as noted in the analysis by Persico, Postlewaite, and Silverman (2004).

(ii) The survey collected information about the respondents family background and family formation. They collected information about the respondents education background including high school, college, and/or vocational training as well as a standardized test of general ability, ASVAB. Finally, the survey collected information about labor participation, government program participation, and income and assets.

f. Height data is available in 1981, 1982,1985, 2006, 2008, 2010 2012, 2014. 2016, 2018.

Question 4

a. There are 12,686 observations in the dta file.

b. There in are 6,403 (50.47%) males and 6,283 (49.53%) females in the 1979 sample.

Question 5

Restrict sample to non-poor white men.

Question 6

a. Generate **height81**

(i) HEALTH_HEIGHT_1981 is encoded as height in feet and inches displayed as a 3 digit integer value.

(ii) -5 response means that the person was not interviewed.

(iii) Responses below 0 are error codes. Additionally height cannot be negative.

(iv) A response of ”510” means that respondent is 5 feet and 10 inches tall.

(v) *Generate Stata variable*

(vi) The mean height for white men in 1981 is 70.34 inches or about 5ft 10.3in.

b. Generate **height85**

(i) *Generate Stata variable*

(ii) The mean height of a white man in the 1985 sample is 70.64 inches or 5ft 10.65inches.

(iii) Error coded observations were dropped. Values less than 0 are error coded.

(iv) The mean change in height between 1985 and 1981 is .2889 inches.

c. Generate **age96**

(i) FAM_1B_1979 is the respondent's age in 1979.

(ii) *Generate Stata variable*

d. Generate **income**

(i) Income in truncated for the top 2%. Negative values are error codes but 0 is valid value.

(ii) *Generate Stata Variable*

(iii) There are 2,094 valid observations for income.

(iv) The mean of the valid income is $34,912.86.

e. Generate **hours**

(i) Negative values are error codes but 0 is valid value.

(ii) *Generate Stata Variable*

(iii) There are 2,139 valid observations for hours in the dataset.

(iv) The mean hours worked is 2,222.4.

(v) Generate *fulltime* variable

(vi) 90.04% of the remaining respondents worked fulltime.

f. Generate **lnWage**

(i) *Generate Stata Variable*

(ii) There are 1,908 observations that have non-missing lnWage values.

(iii) The remaining sample size with non-missing values for lnWage height81, height85, and age96 is 1,818.

g. Generate **educ**. This is the highest level of education reported from the 1979,1981, 1985, and 1996 surveys.

(a) Missing values are negative.

(b) *Generate Stata Variable*

(c) The mean number years completed is 12.2 yrs. This corresponds to a little more than a high school

h. Generate **everMarried**

(a) Values less than 0 are error codes. Values 1 or higher indicate the respondent is currently or was previously married and is now either separated, divorced or widowed.

(b) *Generate Stata Variable*

(c) About 78.48% of the sample have been married.

i. Generate **momSchool**

(i) Error codes are negative. A value of 95 means the respondent has an ungraded highest level of education. However, no respondents have this value.

(ii) *Generate Stata Variable*

(iii) The mean level of the mother's education for the sample is 11.9 years.

j. Generate **momSkilled**

(i) This contains classification codes. Codes 1-245 are considered professional/managerial. Negative values are error codes. There are a significant number of error coded values in this parameter, 1,257.

(ii) *Generate Stata Variable*

(iii) The share of mothers in skilled occupations in 1979 is 19.7%.

k. Generate **dadSchool**. Generate **dadSkilled**

(i) The mean level of school for dads is 12.2 years.

(ii) The share of dads in skilled occupations in 1979 is 32.2%

l. Generate **siblings**

m. Generate **finalSample**

(i) *Generate Stata Variable*

(ii) The final sample size is 890 valid responses.

Question 7

	Mean	Median	Standard Deviation	25^{th} Percentile	75^{th} Percentile	Observations
A. United States NLSY Entire Sample						
Height 1981	70.34	71	2.82	69	72	2900
Height 1985	70.64	71	2.78	69	72	2343
Δ1981-1985	-0.29	0	1.44	-1	0	2281
B. Final Estimation Sample						
Height 1981	70.33	71	2.90	69	72	860
Height 1985	70.70	71	2.75	69	72	860
Δ1981-1985	-0.37	0	1.49	-1	0	860

Final sample is full-time employed white males who had non-missing data in the control variables.

Table 1: *Replication of Persico, Postlewaite, and Silverman (2004) Table 1*

Question 8

	Adult Height Median or Below (1)	Adult Height Above Median (2)	Difference (3)
Adult Characteristics			
1981 Height	68.60	72.52	3.92***
	(2.31)	(1.92)	(0.14)
1985 Height	68.81	73.09	4.28***
	(2.04)	(1.27)	(0.11)
Age	34.44	34.48	0.05
	(2.24)	(2.32)	(0.16)
ln(wage/hour)	2.61	2.71	0.010*
	(0.60)	(0.61)	(0.04)
Ever Married(%)	86.67	89.74	0.03
	(34.02)	(30.39)	(0.02)
Family Background			
Mother's years of schooling	12.26	12.49	0.23
	(2.45)	(2.09)	(0.15)
Mother skilled\professional (%)	18.75	21.57	2.83
	(39.07)	(41.19)	(2.76)
Father's years of schooling	12.50	12.86	0.36
	(3.22)	(3.09)	(0.22)
Father skilled\professional (%)	31.88	35.00	3.12
	(46.65)	(47.76)	(3.25)
Observations	480	380	860

Standard errors in parentheses

Full-time employed white males

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: *Replication of Persico, Postlewaite, and Silverman (2004) Table 2.*

Table 2 shows the samples split by above the median height and at or below the median height of 71 inches. The groups are statistically different in both adult and youth height as well as wage. The differences were calculated by regressing the variable of interest on a dummy representing whether the observation was in the above or at/below median height group. The groups are statistically the same in other characteristics.

Question 9

	(1) lnWage	(2) lnWage	(3) lnWage	(4) lnWage
Adult Height (inches)	0.0303***	0.0230**	0.0136	0.00694
	(0.00764)	(0.00747)	(0.0136)	(0.0124)
Youth height (inches)			0.0184	0.0178
			(0.0123)	(0.0113)
Age	0.0271**	0.0279**	0.0240**	0.0249**
	(0.00870)	(0.00849)	(0.00888)	(0.00869)
Mother's years of schooling		0.0129		0.0117
		(0.0125)		(0.0124)
Mother Skilled/Professional		-0.0225		-0.0200
		(0.0575)		(0.0575)
Father's years of schooling		0.0285**		0.0289**
		(0.00887)		(0.00885)
Father Skilled/Professional		0.132**		0.133**
		(0.0478)		(0.0478)
Number of siblings		-0.0146		-0.0147
		(0.0113)		(0.0112)
adj. R^2	0.027	0.086	0.028	0.087
F	12.33	12.57	8.995	11.26
N	860	860	860	860

Standard errors in parentheses

Full-time employed white males

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: *Replication of Persico, Postlewaite, and Silverman (2004) Table 3 for NLSY*

a. The reduced form specification (1) shows similar results to Persico, Postlewaite, and Silverman (2004). There statistically and economically significant returns to wages for taller men. An additional inch will increase wages by 3% holding all else equal. Including family characteristics as controls in regression (2) reduces the return to adult height to 2.3%. It remains economically and statistically significant. Considering teen height in specification (3) further reduces the returns to adult height. Finally, considering teen height and family characteristics in specification (4) reduces the returns to adult height to an economically small value while return to youth height remains about the same as specification (3).

   Unlike Persico, Postlewaite, and Silverman (2004), the returns to teen height are statistically indistinguishable from zero in specification (3) and (4). This maybe to the different samples

evaluated. The authors used a sample of 2,063 responses while I only had 860 valid responses. The largest contributor to sample size difference was the missing data for the mother's job classification. 1,190 respondents had a valid skip (-4) for mother's profession. The general trends in the results are of the same direction as presented in the paper.

References

Persico, Nicola, Andrew Postlewaite, and Dan Silverman. 2004. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *Journal of Political Economy* 112 (5): 1019–1053. ISSN: 00223808, 1537534X. http://www.jstor.org/stable/10.1086/422566.