Problem Set 5

Spencer Papay - ssp2170

2/22/2017

Question 1

a) Use natural log, as the production parameters can be estimated by a regression of the log of production on the logs of capital, labor, and materials. $lnQ = \beta_0 + \beta_1 lnK + \beta_2 lnL + \beta_3 lnM + u$

b) Test the hypothesis that $\beta_1 + \beta_2 + \beta_3 = 1$ at an appropriate significance level, say 5%. Using 1 as the right hand side implies constant returns to scale. Constant returns to scale implies $f(aK, aL, aM) = af(K, L, M)$. It indicates that an increase in the capital, labor, and materials by an equal factor, a, will increase production by the same factor a.

c) We can test for β_1 if we rewrite the equaion as $Q = \lambda(KLM)^{\beta_1} + e_u$. This identity holds true because just as seen in b), the same returns to scale can carry across all variables. In this case, we would test the hypothesis if $\beta_1 = 1$.

Hence, to test the returns to scale hypothesis, one sets these restrictions on the parameters and transform the regression model. You go about by substituting the restrictions into the equation in order to obtain a reparametrized equation. OLS applied to the new equation will yield restricted estimates.

Question 2

Find complete Stargazer table on following pages for summary of Models 1 through 10

a)

```
mean(birthweight$weight<2500)
```

```
## [1] 0.0559
```

~5.6% of babies weigh less than 2500g at birth.

b) Model 1:

```
m1 <- lm(weight~boy, data=birthweight)
```

Boys are about 116g heavier than girls at birth, and this difference has a high degree of precision with a 5g SE.

c) Model 2:

```
m2 <- lm(weight~boy+smoke, data=birthweight)
```

Boys are on average 116g heavier than girls, holding smoking status constant. Babies whose mothers have smoked during pregnancy are on average 241g lighter than those of mothers who didn't, holding gender constant. The coefficient on "boy" does not change much when controlling for smoking, suggesting that these two variables are independent.

d) Model 3:

```
m3 <- lm(weight~boy+smoke+boy:smoke, data=birthweight)
```

The effect of mothers who smoke during pregnancy is predicted to be about 26g larger for boys than girls. This estimate is not statistically significant at the 5% level, however.

e) Model 4:

```r
m4 <- lm(weight~mom_age, data=birthweight)
```

An increase of one year in mother's age is associated with a 9.6g increase in baby weight.

f) Model 5:

```r
m5 <- lm(weight~mom_age+smoke, data=birthweight)
```

Holding smoking during pregnancy constant, one additional year of mother's age is associated with an average 8.6g increase in baby's weight. If mother's age is held constant, smoking during pregnancy has an estimated reduction in birth weight by 229g.

g) Model 6:

```r
m6 <- lm(weight~mom_age+smoke+mom_age:smoke, data=birthweight)
```

Smoking is associated with a reduction in birth weight by $114.53g - 13.04g \times$ Mom's age in years. This interaction is very significant, and estimates that smoking is more harmful to the baby if the mother is older.

h) Model 7:

```r
m7 <- lm(weight~mom_age+I(mom_age^2)+I(mom_age^3), data=birthweight)
```

A plot of the above model reveals that weight is increasing in mother's age from 18 to ~30 years. However, above this level, the relationship levels out.

i)

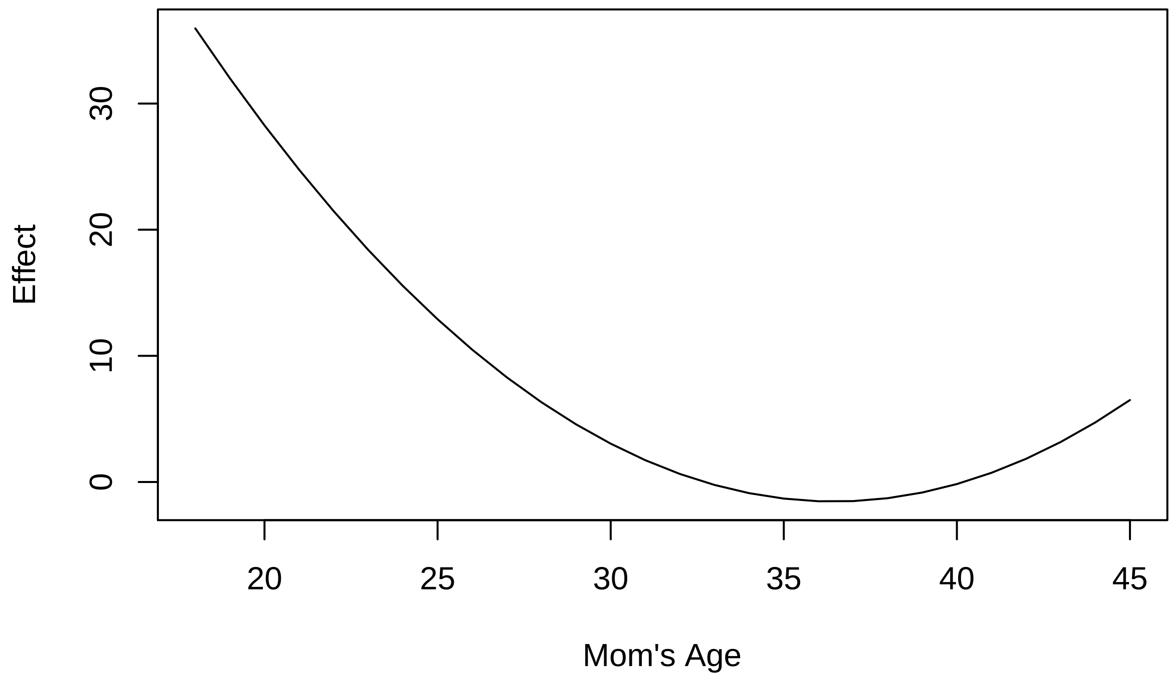
```r
linearHypothesis(m7, c("I(mom_age^2)=0", "I(mom_age^3)=0"), vcov=hccm(m7, type="hc1"))
```

```
## Linear hypothesis test
##
## Hypothesis:
## I(mom_age^2) = 0
## I(mom_age^3) = 0
##
## Model 1: restricted model
## Model 2: weight ~ mom_age + I(mom_age^2) + I(mom_age^3)
##
## Note: Coefficient covariance matrix supplied.
##
##   Res.Df Df      F    Pr(>F)
## 1  49998
## 2  49996  2 85.777 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The F-test rejects the null hypothesis that the coefficients on the second- and third-degree terms are jointly equal to 0 at a reasonable significance level. Therefore, investigating a nonlinear effect of mother's age would behoove our research.

j)

```r
plot(18:45, diff(predict(m7,data.frame(mom_age=18:46))), type="l", xlab="Mom's Age", ylab="Effect")
```



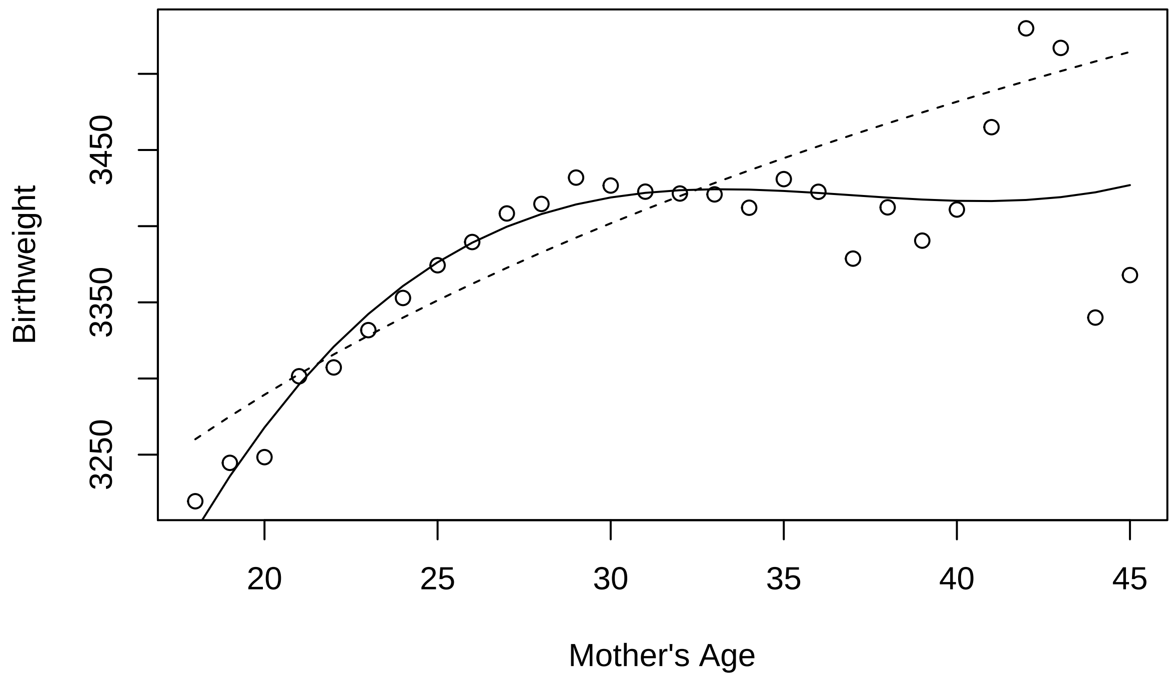
The plot reveals that the estimated effect of a one-year increase in age is about 36g for an 18 year-old mother, decreasing over age until about 35, at which point the difference reaches 0g. The effect increases again with age after this point.

k) Model 8:

```
m8 <- lm(weight~log(mom_age),data=birthweight)
```

This model estimates that a 1% increase in mother's age (percent because this is log) is associated with a 2.7g increase in birth weight. We can plot this in ggplot and see the results:

```
plot(18:45,as.vector(by(birthweight$weight, as.factor(birthweight$mom_age),mean)),
      xlab="Mother's Age",ylab="Birthweight")
    lines(18:45, predict(m7,data.frame(mom_age=18:45)), lty=1)
    lines(18:45, predict(m8,data.frame(mom_age=18:45)), lty=2)
```



This plot reveals that the estimated regression lines from the polynomial better fits the average weight of babies of mothers aged 18 through 45.

l) Model 9:

```
m9 <- lm(weight~mom_age + I(mom_age^2) + I(mom_age^3)
         + educ + black + married + boy + smoke, data=birthweight)
```

As the question asks, if we take the effect of smoking as the relationship of interest, many of the other variables can be taken as controls. The gender is arguably independent from the other variables in the model, but does seem to have a causal effect. After controlling for several factors, smoking still seems to be associated with a decrease in average birthweight of about 215g. However, it remains possible that other variables are omitted from this regression that are not in this dataset, for example socioeconomic status/lifestyle choices as well as drinking; both of these may be correlated with smoking.

m)

$68.55g - 34.03g = \boxed{34.52g}$

n) Model 10:

```
m10 <- lm(weight ~ mom_age + I(mom_age^2)+ I(mom_age^3)
             + boy + smoke + black + educ +
            married + black:smoke, data = birthweight)
```

The effect of mothers who smoke during pregnancy is predicted to be about 16g larger for black mothers than non-black mothers. This result is not significant at even the 10% level. This would suggest, but more evidence would be needed to say conclusively, that the effects of a mother smoking are independent on a mother's race.

```r
stargazer(m1, m2, m3, m4, m5, type="latex", header=FALSE,
          title="Effects on Birthweight - Models 1-5",
dep.var.caption = "Baby's Birthweight in g",
omit.stat = c("f"),
float.env = "sidewaystable", intercept.bottom = FALSE,
covariate.labels = c("Constant", "Boy", "Smoke",
                     "Boy×Smoke", "Mom's age"),se=list(se1,se2,se3,se4,se5))
```

```r
stargazer(m6, m7, m8, type="latex", header=FALSE,
          title="Effects on Birthweight - Models 6-8",
dep.var.caption = "Baby's Birthweight in g",
omit.stat = c("f"),
intercept.bottom = FALSE, covariate.labels = c("Constant", "Mom's age",
                                               "Smoke", "Mom's Age×Smoke",
                                               "Mom's age$^2$", "Mom's age$^3$",
                                               "log(Mom's age)"),se=list(se6,se7,se8))
```

```r
stargazer(m9, m10, type="latex", header=FALSE,
          title="Effects on Birthweight - Models 9-10",
dep.var.caption = "Baby's Birthweight in g",
omit.stat = c("f"),
intercept.bottom = FALSE, covariate.labels = c("Constant", "Mom's age",
                                               "Mom's age$^2$", "Mom's age$^3$",
                                               "High School", "Some College", "College",
                                               "Black", "Married", "Black×Smoke",
                                               "Boy",  "Smoke"), se=list(se9,se10))
```

Table 1: Effects on Birthweight - Models 1-5

	Baby's Birthweight in g				
	weight				
	(1)	(2)	(3)	(4)	(5)
Constant	3,310.564***	3,342.003***	3,343.764***	3,106.455***	3,167.970***
	(3.520)	(3.595)	(3.707)	(12.549)	(12.646)
Boy	116.688***	116.978***	113.562***		
	(5.033)	(4.980)	(5.318)		
Smoke		−241.764***	−255.306***		−229.127***
		(7.568)	(10.662)		(7.705)
Boy×Smoke			26.154*		
			(15.122)		
Mom's age				9.640***	8.489***
				(0.453)	(0.451)
Observations	50,000	50,000	50,000	50,000	50,000
R^2	0.011	0.031	0.031	0.010	0.028
Adjusted R^2	0.011	0.031	0.031	0.009	0.028
Residual Std. Error	563.381 (df = 49998)	557.463 (df = 49997)	557.451 (df = 49996)	563.692 (df = 49998)	558.422 (df = 49997)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: Effects on Birthweight - Models 6-8

	Baby's Birthweight in g		
	weight		
	(1)	(2)	(3)
Constant	3,119.952***	1,625.009***	2,458.024***
	(13.356)	(239.369)	(39.995)
Mom's age	10.228***	148.867***	
	(0.477)	(26.150)	
Smoke	114.530***		
	(36.815)		
Mom's Age×Smoke	−13.043***		
	(1.393)		
Mom's age^2		−4.071***	
		(0.927)	
Mom's age^3		0.037***	
		(0.011)	
log(Mom's age)			277.504***
			(12.172)
Observations	50,000	50,000	50,000
R^2	0.030	0.013	0.011
Adjusted R^2	0.030	0.013	0.011
Residual Std. Error	557.854 (df = 49996)	562.717 (df = 49996)	563.354 (df = 49998)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: Effects on Birthweight - Models 9-10

	Baby's Birthweight in g	
	weight	
	(1)	(2)
Constant	2,229.000***	2,226.870***
	(233.372)	(233.282)
Mom's age	99.577***	99.704***
	(25.442)	(25.433)
Mom's age^2	−3.041***	−3.043***
	(0.899)	(0.898)
Mom's age^3	0.031***	0.031***
	(0.010)	(0.010)
High School	34.028***	33.849***
	(7.600)	(7.598)
Some College	52.110***	51.949***
	(8.413)	(8.411)
College	68.552***	68.471***
	(8.943)	(8.942)
Black	−215.476***	−213.728***
	(7.699)	(8.086)
Married	58.044***	58.004***
	(6.719)	(6.718)
Black×Smoke		−16.618
		(24.397)
Boy	115.931***	115.919***
	(4.883)	(4.883)
Smoke	−215.621***	−213.498***
	(7.792)	(8.221)
Observations	50,000	50,000
R^2	0.069	0.069
Adjusted R^2	0.069	0.069
Residual Std. Error	546.617 (df = 49989)	546.619 (df = 49988)

Note: *p<0.1; **p<0.05; ***p<0.01