Age and Education Differences Among Couples

Ava Droznik

```
acs2021_couples$educ_numeric <- fct_recode(acs2021_couples$EDUC,
                                           "0" = "N/A or no schooling",
                                           "2" = "Nursery school to grade 4",
                                           "6.5" = "Grade 5, 6, 7, or 8",
                                           "9" = "Grade 9",
                                           "10" = "Grade 10",
                                           "11" = "Grade 11",
                                           "12" = "Grade 12",
                                           "13" = "1 year of college",
                                           "14" = "2 years of college",
                                           "15" = "3 years of college",
                                           "16" = "4 years of college",
                                           "17" = "5+ years of college")

acs2021_couples$educ_numeric <-
as.numeric(levels(acs2021_couples$educ_numeric))[acs2021_couples$educ_numeric
]

acs2021_couples$h_educ_numeric <- fct_recode(acs2021_couples$h_educ,
                                           "0" = "N/A or no schooling",
                                           "2" = "Nursery school to grade 4",
                                           "6.5" = "Grade 5, 6, 7, or 8",
                                           "9" = "Grade 9",
                                           "10" = "Grade 10",
                                           "11" = "Grade 11",
                                           "12" = "Grade 12",
                                           "13" = "1 year of college",
                                           "14" = "2 years of college",
                                           "15" = "3 years of college",
                                           "16" = "4 years of college",
                                           "17" = "5+ years of college")

acs2021_couples$h_educ_numeric <-
as.numeric(levels(acs2021_couples$h_educ_numeric))[acs2021_couples$h_educ_num
eric]

acs2021_couples$educ_diff <- acs2021_couples$educ_numeric -
acs2021_couples$h_educ_numeric

# Hypothesis:

# H0: β2 = β3 = 0 (the squared and cubic terms of education difference have
```

```r
no joint effect)

# HA: At least one of β2, β3 ≠ 0 (nonlinear relationship between age and
education differences)

# Model 1: Simple linear relationship between age and education differences

m1 <- lm(age_diff ~ educ_diff, data = acs2021_couples)

# Model 2: Add nonlinear terms for education difference

m2 <- lm(age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3),
data = acs2021_couples)

# Model 3: Add partner controls (ages and education levels)

acs2021_couples$avg_age <- (acs2021_couples$AGE + acs2021_couples$h_age) / 2

m3 <- lm(age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3) +
                    avg_age + I(avg_age^2),
         data = acs2021_couples)

# Robust coefficient test for Model 1

coeftest(m1, vcov = vcovHC)

##
## t test of coefficients:
##
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept) -0.7556547  0.0072612 -104.0678   <2e-16 ***
## educ_diff    0.0072887  0.0041757    1.7455   0.0809 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Wald test for nonlinearity (do polynomial terms jointly matter?)

waldtest(m1, m2, vcov = vcovHC)

## Wald test
##
## Model 1: age_diff ~ educ_diff
## Model 2: age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3)
##   Res.Df Df     F    Pr(>F)
## 1 776062
## 2 776060  2 368.8 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Wald test for additional age/education controls

waldtest(m2, m3, vcov = vcovHC)
```

```
## Wald test
##
## Model 1: age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3)
## Model 2: age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3) + avg_age
+
##     I(avg_age^2)
##   Res.Df Df      F    Pr(>F)
## 1 776060
## 2 776058  2 151.16 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
models <- list(
  "M 1" = lm(age_diff ~ educ_diff, data = acs2021_couples),
  "M 2" = lm(age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3),
             data = acs2021_couples),
  "M 3" = lm(age_diff ~ educ_diff + I(educ_diff^2) + I(educ_diff^3) +
                        avg_age + I(avg_age^2),
             data = acs2021_couples)
)

modelsummary(models, stars = TRUE, output = "default")
```

	M 1	M 2	M 3
(Intercept)	-0.756***	-0.650***	-0.087
	(0.007)	(0.008)	(0.070)
educ_diff	0.007**	-0.038***	-0.037***
	(0.003)	(0.004)	(0.004)
I(educ_diff^2)		-0.014***	-0.014***
		(0.000)	(0.000)
I(educ_diff^3)		0.000***	0.000***
		(0.000)	(0.000)
avg_age			-0.030***
			(0.003)
I(avg_age^2)			0.000***
			(0.000)
Num.Obs.	776064	776064	776064
R2	0.000	0.003	0.003
R2 Adj.	0.000	0.003	0.003

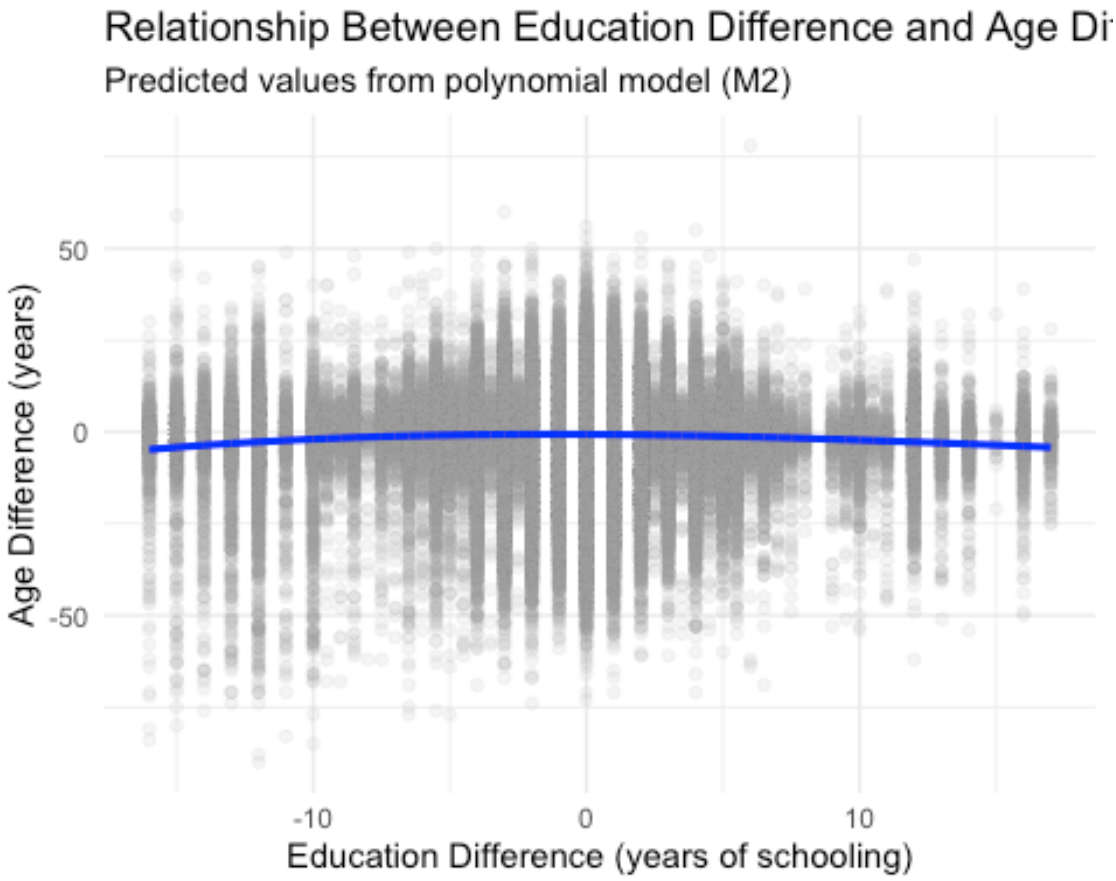
	M 1	M 2	M 3
AIC	5085935.4	5083663.0	5083402.2
BIC	5085970.1	5083720.8	5083483.1
Log.Lik.	-2542964.709	-2541826.519	-2541694.079
F	7.212	762.314	510.528
RMSE	6.41	6.40	6.40

- p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

```
# Predicted values from Model 2

plot_data <- acs2021_couples %>%
mutate(pred_age_diff = predict(m2))

ggplot(plot_data, aes(x = educ_diff, y = age_diff)) +
geom_point(alpha = 0.1, color = "gray60") +
geom_line(aes(y = pred_age_diff), color = "blue", linewidth = 1.1) +
theme_minimal() +
labs(
title = "Relationship Between Education Difference and Age Difference",
subtitle = "Predicted values from polynomial model (M2)",
x = "Education Difference (years of schooling)",
y = "Age Difference (years)"
)
```


Relationship Between Education Difference and Age Di
Predicted values from polynomial model (M2)
Age Difference (years)
50
0
-50
-10
0
10
Education Difference (years of schooling)