Answers to
Empirical Exercises for Chapter 6

This table contains the results from seven regressions that are referenced in these answers.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable						
	AHE	ln(*AHE*)	ln(*AHE*)	ln(*AHE*)	ln(*AHE*)	ln(*AHE*)	ln(*AHE*)
Age	0.314** (0.029)	0.021** (0.002)		0.159** (0.047)	0.162** (0.047)	0.151* (0.064)	0.155* (0.063)
Age^2				–.0023** (.0008)	–.0024** (.0008)	–.0023* (0.0011)	–0.0023* (0.0011)
ln(*Age*)			0.637** (0.061)				
Female×*Age*						0.025 (0.095)	
Female×Age^2						–.0005 (0.0015)	
Bachelor×*Age*							0.014 (0.095)
Bachelor×Age^2							–0.0002 (0.0016)
Female	–2.493** (0.162)	–0.180** (0.012)	–0.180** (0.012)	–0.179** (.012)	–0.218** (.016)	–0.531 (1.390)	–0.218** (0.016)
Bachelor	5.336** (0.171)	0.383** (.012)	0.383** (.012)	0.382** (.012)	0.346** (.016)	0.346** (0.016)	0.095 (1.400)
Female×*Bachelor*					0.085*** (0.023)	0.084*** (0.024)	0.086*** (0.023)
Intercept	3.300** (0.867)	1.79** (0.06)	0.268 (0.207)	–0.230 (0.696)	–0.255 (0.694)	–0.131 (0.943)	–0.139 (0.926)
***F*-statistic and *p*-values on joint hypotheses**							
(a) *F*-statistic on terms involving *Age*				57.88 (0.00)	59.01 (0.00)	29.58 (0.00)	29.74 (0.00)
(b) Interaction terms with *Age* and Age^2						0.44 (0.64)	0.32 (0.72)
SER	6.25	0.445	0.445	0.445	0.445	0.445	0.445
$\bar{R}^2$	.1819	.1824	.1827	.1834	.1850	.1849	.1848

Significant at the *5% and **1% significance level.

1. *Run a regression of average hourly earnings (AHE) on age (Age), gender (Female), and education (Bachelor). If Age increases from 25 to 26, how are earnings expected to change? If Age increases from 33 to 34, how are earnings expected to change?*

The regression results for this question are shown in column (1) of the table. If *Age* increases from 25 to 26, earnings are predicted to increase by $0.314 per hour. If *Age* increases from 33 to 34, earnings are predicted to increase by $0.314 per hour. These values are the same because the regression is a linear function relating *AHE* and *Age*.

2. *Run a regression of the logarithm average hourly earnings, ln(AHE), on Age, Female, and Bachelor. If Age increases from 25 to 26, how are earnings expected to change? If Age increases from 33 to 34, how are earnings expected to change?*

The regression results for this question are shown in column (2) of the table. If *Age* increases from 25 to 26, ln(*AHE*) is predicted to increase by 0.021. This means that earnings are predicted to increase by 2.1%. . If *Age* increases from 34 to 35, ln(*AHE*) is predicted to increase by 0.021. This means that earnings are predicted to increase by 2.1%. These values, in percentage terms, are the same because the regression is a linear function relating ln(*AHE*) and *Age*.

3. *Run a regression of the logarithm average hourly earnings, ln(AHE), on ln(Age), Female, and Bachelor. If Age increases from 25 to 26, how are earnings expected to change? If Age increases from 33 to 34, how are earnings expected to change?*

The regression results for this question are shown in column (3) of the table. If *Age* increases from 25 to 26, then ln(*Age*) has increased by ln(26) – ln(25) = .0392 (or 3.92%). The predicted increase in ln(*AHE*) is 0.637×(.0392) = .0250. This means that earnings are predicted to increase by 2.5%. . If *Age* increases from 34 to 35, then ln(*Age*) has increased by ln(35) – ln(34) = .0290 (or 2.90%). The predicted increase in ln(*AHE*) is 0.637×(.0290) = .0185. This means that earnings are predicted to increase by 1.85%.

4. *Run a regression of the logarithm average hourly earnings, ln(AHE), on Age, Age^2, Female, and Bachelor. If Age increases from 25 to 26, how are earnings expected to change? If Age increases from 33 to 34, how are earnings expected to change?*

When *Age* increases from 25 to 26, the predicted change in ln(*AHE*) is

$(0.159 \times 26 - .0023 \times 26^2) - (0.159 \times 25 - .0023 \times 25^2) = .0417$.

This means that earnings are predicted to increase by 4.17%.

When *Age* increases from 34 to 35, the predicted change in ln(*AHE*) is

$(0.159 \times 35 - .0023 \times 35^2) - (0.159 \times 34 - .0023 \times 34^2) = .0003$.

This means that earnings are predicted to increase by 0.03%.

5. *Do you prefer the regression in (3) to the regression in (2)? Explain.*

The regression differ in their choice of one of the regressors. They can be compared on the basis of the $\bar{R}^2$. The regression in (3) has a higher $\bar{R}^2$, so it is preferred.

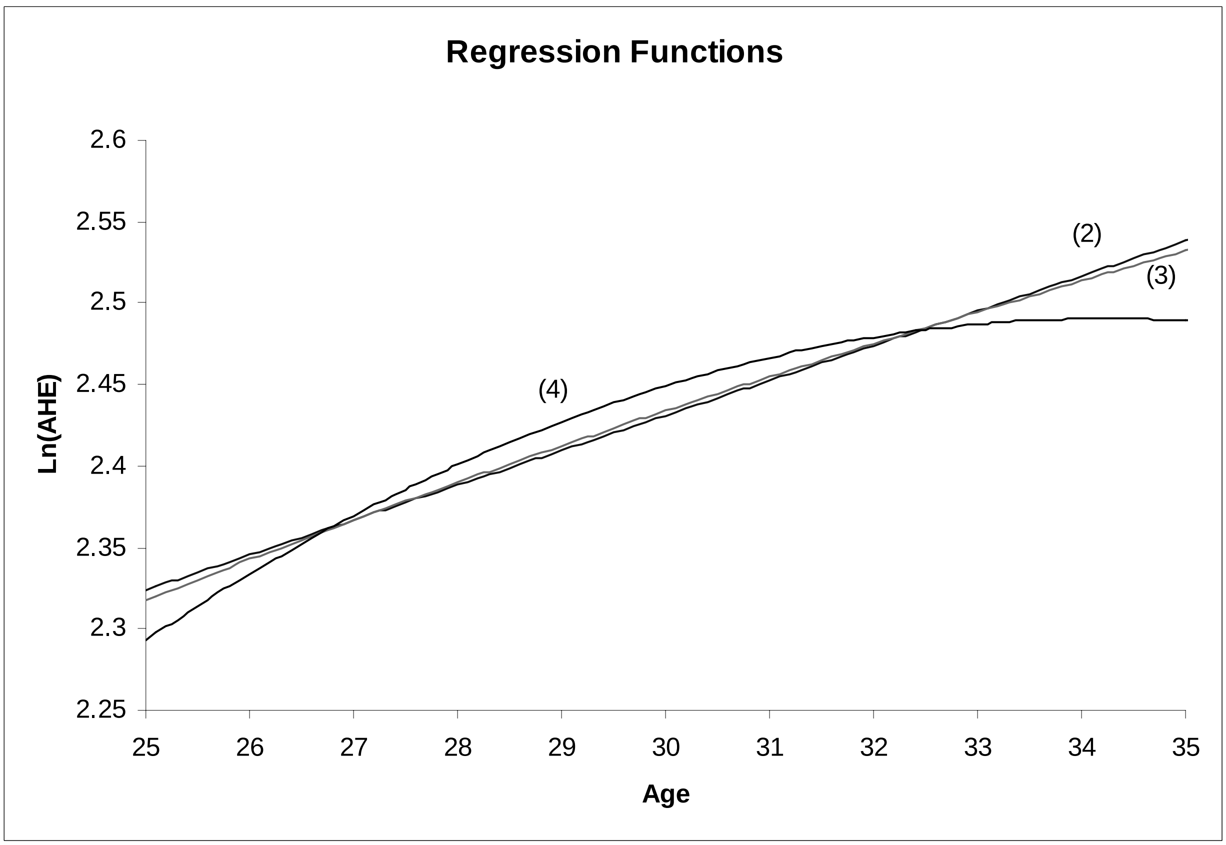
6. *Do you prefer the regression in (4) to the regression in (2)? Explain.*

The regressions in (4) adds the variable *Age*^{2} to regression (2). The coefficient on *Age*^{2} is statistically significant (t = –2.91), and this suggests that the addition of *Age*^{2} is important. Thus, (4) is preferred to (2).

7. *Do you prefer the regression in (4) to the regression in (3)? Explain.*

The regressions differ in their choice of one of the regressors. They can be compared on the basis of the $\bar{R}^2$. The regression in (4) has a higher $\bar{R}^2$, so it is preferred.

8. *Plot the regression relation between Age and ln(AHE) from (2), (3), and (4) for males with a high school diploma. Describe the similarities and differences between the estimated regression functions. Would your answer change if you plotted the regression function for females with college degrees?*



The regression functions using *Age* (2) and ln(*Age*) (3) are similar. The quadratic regression (4) is different. It shows a decreasing effect of *Age* on ln(*AHE*) as workers age.

The regression functions for a female with a high school diploma will look just like these, but they will be shifted by the amount of the coefficient on the binary regressor *Female*. The regression functions for workers with a bachelor's degree will also look just like these, but they would be shifted by the amount of the coefficient on the binary variable *Bachelor*.

9. *Run a regression of ln(AHE), on Age, Age^2, Female, Bachelor, and the interaction term Female×Bachelor. What does the coefficient on the interaction term measure? Alexis is a 30-year-old female with a bachelor's degree. What does the regression predict for her value of ln(AHE)? Jane is a 30-year-old female with a high school degree. What does the regression predict for her value of ln(AHE)? What is the predicted difference between Alexis's and Jane's earnings? Bob is a 30- year-old male with a bachelor's degree. What does the regression predict for his value of ln(AHE)? Jim is a 30-year-old male with a high school degree. What does the regression predict for his value of ln(AHE)? What is the predicted difference between Bob's and Jim's earnings?*

This regression is shown in column (5). The coefficient on the interaction term *Female*×*Bachelor* shows the "extra effect" of *Bachelor* on ln(*AHE*) for women relative the effect for men.

Predicted values of ln(*AHE*):

Alexis: $0.162\times30 - .0024\times30^2 - .218\times1 + .346\times1 + .085\times1 - 0.255 = 2.658$
Jane: $0.162\times30 - .0024\times30^2 - .218\times1 + .346\times0 + .085\times0 - 0.255 = 2.227$
Bob: $0.162\times30 - .0024\times30^2 - .218\times0 + .346\times1 + .085\times0 - 0.255 = 2.791$
Jim: $0.162\times30 - .0024\times30^2 - .218\times0 + .346\times0 + .085\times0 - 0.255 = 2.445$

Alexis – Jane = 2.658 – 2.227 = 0.431
Bob – Jim = 2.791 – 2.445 = 0.346

Notice the difference in the difference predicted effects is 0.431 – 0.346 = .085, which is the value of the coefficient on the interaction term.

10. *Is the effect of Age on earnings different for males than females? Specify and estimate a regression that you can use to answer this question.*

This regression is shown in (6), which includes two additional regressors that are interactions of *Female* and the age variables, *Age* and Age^2. The *F*-statistic testing the restriction that the coefficients on these interaction terms is zero is 0.44 with a *p*-value of 0.64. This implies that there is no statistically significant evidence that there is a different effect of *Age* on ln(*AHE*) for men and women.

11. *Is the effect of Age on earnings different for high school than college graduates? Specify and estimate a regression that you can use to answer this question.*

This regression is shown in (7), which includes two additional regressors that are interactions of *Bachelor* and the age variables, *Age* and Age^2. The *F*-statistic testing the restriction that the coefficients on these interaction terms is zero is 0.32 with a *p*-value of 0.72. This implies that there is no statistically significant evidence that there is a different effect of *Age* on ln(*AHE*) for high school and college graduates.

12. *After running all of these regressions (and any others that you want to run), summarize the effect of age on earning for young workers.*

The estimated regressions suggest that earnings increase as workers age from 25-35, the range of age studied in this sample. The effect of *Age* on the logarithm of earnings is summarized in the plot in question (8). Regression specifications using *Age* and ln(*Age*) suggest that the logarithm of earnings increases by approximately .12 (so that *AHE* increases by approximately 12%) as a worker's age increases from 25 to 32 years. The quadratic specification suggests a somewhat larger increase in *AHE* (approximately 16%) as a worker's age increases from 25-32. The regression specifications using *Age* or ln(*Age*) suggest that earnings increase another 5% (the predicted change in ln(*Age*) is .05) as a worker's age increases from 32 to 35 years. The quadratic specification implies little increase in earnings over this range of ages.

There is evidence that the quadratic term Age^2 belongs in the regression. This can be seen by comparing regression (2) and (4), and also by comparing the adjusted R^2 values from (3) and (4).

Gender and education were also significant predictors of earnings, but there was no significant evidence that the effect of *Age* on ln(*AHE*) was different for men and women or for high school and college graduates.