Instructor’s Manual

**Answers to Odd-numbered Exercises**

1. Chapter 1

1-3. (a) The coefficient of Li represents the change in the percentage chance of making a putt when the length of the putt increases by one foot. In this case, the percentage chance of making the putt decreases by 4.1 for each foot longer the putt is.

(b) The equations are identical. To convert one to the other, note that i  = Pi – ei, which is true because ei = Pi - i (or more generally, ei = Yi – ).

(c) 42.6 percent, yes; 79.5 percent, no (too low); -18.9 percent, no (negative!).

(d) One problem is that the theoretical relationship between the length of the putt and the percentage of putts made is almost surely non-linear in the variables; we’ll discuss models appropriate to this problem in Chapter 7. A second problem is that the actual dependent variable is limited by zero and one but the regression estimate is not; we’ll discuss models appropriate to this problem in Chapter 13.

1-5. (a) βʏ is the change in S caused by a one-unit increase in Y, holding G constant and βG is the change in S caused by a one-unit increase in G, holding Y constant.

(b) +, -

(c) Yes. Richer states spend at least some of their extra money on education, but states with rapidly growing student populations find it difficult to increase spending at the same rate as the student population, causing spending per student to fall, especially if you hold the wealth of the state constant.

(d) Ŝi = -183 + 0.1422Yi - 59.26Gi. Note that 59.26 ∙ 10 = 5926 ∙ 0.10, so nothing in the equation has changed except the scale of the coefficient of G.

1-7. (a) β2 represents the impact on the wage of the *i*th worker of a one-year increase in the education of the *i*th worker, holding constant that worker’s experience and gender.

(b) β3 represents the impact on the wage of the *i*th worker of being male instead of female, holding constant that worker’s experience and education.

(c) There are two ways of defining such a dummy variable. You could define COLORi = 1 if the *i*th worker is a person of color and 0 otherwise, or you could define COLOR1 = 1 if the *i*th worker is not a person of color and 0 otherwise. (The actual name you use for the variable doesn’t have to be “COLOR.” You could choose any variable name as long as it didn’t conflict with the other variable names in the equation.)

(d) We’d favor adding a measure of the quality of the worker to this equation, and answer iv, the number of employee of the month awards won, is the best measure of quality in this group. As tempting as it might be to add the average wage in the field, it would be the same for each employee in the sample and thus wouldn’t provide any useful information.

1. Chapter 2

2-3. (a) The squares are “least” in the sense that they are being minimized.

(b) If R2  0, then RSS  TSS, and ESS  0. If R2 is calculated as ESS/TSS, then it cannot be negative. If R2 is calculated as 1 – RSS/TSS, however, then it can be negative if RSS > TSS, which can happen if  is a *worse* predictor of Y than  (possible only with a non-OLS estimator or if the constant term is omitted).

(c) positive.

(d) We prefer Model T because it has estimated signs that meet expectations and also because it includes an important variable that Model A omits. A higher R2 does not *automatically* mean that an equation is preferred.

2-5. (a) Even though the fit in Equation A is better, most researchers would prefer equation B because the signs of the estimated coefficients are as would be expected. In addition, X4 is a theoretically sound variable for a campus track, while X3 seems poorly specified because an especially hot *or* cold day would discourage fitness runners.

(b) The coefficient of an independent variable tells us the impact of a one-unit increase in that variable on the dependent variable holding constant the other explanatory variables in the equation. If we change the other variables in the equation, we’re holding different variables constant, and so the  has a different meaning.

2-7. (a) Yes. We’d expect bigger colleges to get more applicants, and we’d expect colleges that used the common application to attract more applicants. It might seem at first that the rank of a college ought to have a positive coefficient, but the variable is defined as 1 = best, so we’d expect a negative coefficient for RANK.

(b) The meaning of the coefficient of SIZE is that for every increase of one in the size of the student body, we’d expect a college to generate 2.15 more applications, holding RANK and COMMONAP constant. The meaning of the coefficient of RANK is that every one-rank improvement in a college’s USNews ranking should generate 32.1 more applications, holding SIZE and COMMONAP constant. These results do not allow us to conclude that a college’s ranking is 15 times more important than the size of that college because the units of the variables SIZE and RANK are quite different in magnitude. On a more philosophical level, it’s risky to draw any general conclusions at all from one regression estimated on a sample of 49 colleges.

(c) The meaning of the coefficient of COMMONAP is that a college that switches to using the common application can expect to generate 1222 more applications, holding constant RANK and SIZE. However, this result does not prove that a given college would increase applications by 1222 by switching to the common application. Why not? First, we don’t trust this result because there may well be an omitted relevant variable (or two) and because all but three of the colleges in the sample use the common application. Second, in general, econometric results are evidence that can be used to support an argument, but in and of themselves they don’t come close to “proving” anything.

(e) If you drop COMMONAP from the equation,  falls. This is evidence (but not proof) that COMMONAP belongs in the equation.

1. Chapter 3

3-3. (a) A male professor in this sample earns $817 more than a female professor, holding constant the other independent variables in the equation.

(b) Most students will expect a negative coefficient, so they will call this an unexpected sign. Most professors and administrators will expect a positive sign because of the growing competition among colleges for African-American professors, so they will call this an expected sign. A key point here is not to change expectations based solely on this result.

(c) R is not a dummy variable because it takes on more than two values. For each additional year in rank, the ith professor’s salary will rise by $406, holding constant the other independent variables in the equation.

(d) Yes. The coefficient is large and, as we’ll learn in Chapter 5, statistically significantly greater than zero. (In addition, it’s quite robust.)

(e) There’s no measure of the quality of the professor in the equation as it stands, so good suggestions might the number of articles published by the *i*th professor or the average teaching evaluation (on a standard scale) of the *i*th professor.

3-5. (a) A male student’s GRE subject score in Economics is likely to be 39.7 points higher than a female’s, holding constant their GPA and SATs.

(b) This result is evidence of, but not proof of, bias. If we were sure that we had the best possible specification (the topic of Chapter 6) and if this result turned out to be statistically significant (the topic of Chapter 5), and if we were able to reproduce this result in other samples, we’d be much closer to a “proof.” Even then, there still would be a possibility that some factor other than bias was the cause of these results.

(c) Possible variables include the number of upper division economics courses taken, the number of mathematics classes taken, and dummy variables measuring whether the student had taken econometrics or international economics (two fields frequently covered in the test). It’s vital that any suggested variable be cross-sectional by student.

(d) The equation would become   = 212.1 – 39.7Gi + 78.9GPAi + 0.203SATMi + 0.110SATVi..

3-7. (a) The best way to handle three discrete conditions is to specify two dummy variables. For example, one dummy variable could =1 if the iPod is new (and 0 otherwise) and the other dummy variable could =1 if the iPod is used but unblemished (and 0 otherwise). The omitted condition, that the iPod is used and scratched, would be represented by both dummy variables equaling zero.

(b) Positive, negative, positive

(c) In theory, the narrower the time spread of the observations, the better the sample, but three weeks probably is a short enough time period to ensure that the observations are from the same population. If the three weeks included a major shock to the iPod market, however, then the friend would be right, and the sample should be split into “before the shock” and “after the shock” subsamples.

(d) Yes, they match with the answer to part b.

(e)  is missing!

(f)  is .431.

1. Chapter 4

4-3. Pair “c” clearly violates Assumption VI, and pair “a” probably violates it for most samples.

4-5. (a) Yes, Yes. In particular, there’s no measure of prices in the equation.

(b) Yes

(c) Yes, Very unlikely

(d) No

(e) No

(f) No

(g) The nightclub should hire a dancer because the estimated coefficient is higher.

1. Chapter 5

5-3. (a) H0:β1  0, HA:β1 > 0

(b) H0:β1 ≥ 0, HA:β1 < 0; H0:β2 ≤ 0, HA:β2 > 0;

H0:β3 ≤ 0, HA:β3 > 0 (The hypothesis for β3 assumes that it is never too hot to go jogging.)

(c) H0:β1   0, HA:β1 > 0; H0:β2 ≤ 0, HA:β2 > 0;

H0:β3 ≥ 0, HA:β3 < 0; (The hypothesis for β3 assumes you’re not breaking the speed limit.)

(d) H0:βG = 0; HA:βG  0 (G for grunt.)

5-5. (a) t2 = (200 – 160)/25.0 = 1.6; tc = 2.052; therefore cannot reject H0. (Notice the violation of the principle that the null hypothesis contains that which we do not expect.)

(b) t3 = 2.37; tc = 2.756; therefore cannot reject the null hypothesis.

(c) t2 = 5.6; tc = 2.447; therefore reject H0 if it is formulated as in the exercise, but this poses a problem because the original hypothesized sign of the coefficient was negative. Thus, the alternative hypothesis ought to have been stated: HA: β2 < 0, andH0 cannot be rejected because the sign of t2 doesn’t agree with the sign in the alternative hypothesis.

5-7. (a) For both, H0: β ≤ 0 and HA: β > 0. For WIN, we cannot reject H0, even though the sign agrees with the sign implied by HA, because |+1.00| < 1.697, the 5 percent one-sided critical *t*-value for 30 degrees of freedom. For FREE, we can reject H0 at the 5 percent level of significance because |2.00|> 1.697 and because 2.00 has the sign implied by HA.

(b) H0: βWEEK ≥ 0 and HA: βWEEK < 0. We can reject H0 at the 1 percent level because |–4.00| > 2.457, the 1 percent, one-sided critical *t*-value for 30 degrees of freedom and because –4.00 has the sign implied by HA.

(c) H0: βDAY = 0 and HA: βDAY ≠ 0. We cannot reject H0 because |-1.00| < 2.042, the 5 percent two-sided critical *t*-value for 30 degrees of freedom.

(d) The coefficients of DAY and WIN are insignificantly different from zero. In addition, it’s hard to rule out the possibility that a variable that belongs in the equation might have been omitted.

(e) A potential omitted variable is more worrisome than an insignificant coefficient.

(f) We’d suggest adding a variable that measures the weather (like inches of rainfall that day) to the equation. Even given San Diego’s wonderful weather, there’s a good chance that rainy or cold weather could cut down on attendance at an outdoor event.

5-9. (a) All five tests are one-sided, so tc = 1.706 throughout.

GDPN: H0:β ≤ 0, HA: β > 0. Reject H0 because |+6.69| > 1.706 and 6.69 is positive as in HA.

CVN: H0:β ≥ 0, HA: β < 0. Reject H0 because |-2.66| > 1.706 and -2.66 is negative as in HA.

PP: H0:β ≤ 0, HA: β > 0. Do not reject H0 because |+1.19|< 1.706.

DPC: H0:β ≥ 0, HA: β < 0. Reject H0 because |-2.25| > 1.706 and -2.25 is negative as in HA.

IPC: H0:β ≥ 0, HA:β < 0. Do not reject H0 because |-1.59| < 1.706.

(b) Our confidence interval equation , and the 10 percent two-sided tc = 1.706 (the same as a one-sided 5 percent tc), so the confidence interval equals  1.706 ⋅ SE or

GDPN: 1.07 <  1.79

CVN: -0.98 < –0.22

PP: -3.13 < 17.75

DPC: -27.45 < –3.81

IPC: -23.59 <  0.83

(c) Yes. The important signs were as expected and statistically significant, and the overall fit was good.

(d) The sizes of the coefficients would change, but not their signs or significance.

1. Chapter 6

6-3. (a) Coefficient:   

Hypoth. Sign:   

t-value: 4.0 4.0 –2.0

t C  1.314 reject reject do not

(10% one-sided reject

with 27 d.f.)

The problem with the estimated coefficient of M is that it is significant in the unexpected direction, one indicator of a possible omitted variable.

(b) The estimated coefficient of M is unexpectedly negative, so we’re looking for a variable the omission of which would cause negative bias in the estimate of  We thus are looking for a variable that is negatively correlated with meat consumption with a positive expected coefficient *or* a variable that is positively correlated with meat consumption with a negative expected coefficient. For the six variables listed, the expected bias is:

|  |  |  |  |
| --- | --- | --- | --- |
| Possible Omitted Variable | Expected Sign of | Correlation with M | Direction of Bias |
| B |  | \* |  |
| F |  |  |  |
| W | \* |  |  |
| R | – | – |  |
| H | – |  | – |
| O | – | – |  |

\*indicates a weak expected sign or correlation.

As can be seen, the only one of these six suggested variables that is likely to have produced negative bias is H. This doesn’t “prove” that H belongs in the equation, but it does make it extremely unlikely that the negative bias was caused by the omission of any of the other five variables.

(c) The best suggested variables are aggregate time-series variables for the US the omission of which would cause negative bias. The expected bias equation is difficult to work with the first time around, so take the time to double check and make sure that your students suggest variables would create negative bias. Keep your fingers crossed that they don’t suggest a cross-sectional variable.

6-5. (a) Coefficient βPARENT βHSRANK

Hypothesized sign: – +

Calculated t-score: –11.26 4.22

tc = 1.679 (5% level), so: reject H0 reject H0

(b) There are no obvious signs of an omitted or irrelevant variable, but it seems probable that more than two variables determine financial aid grants in most colleges, so an omitted variable is very likely from a theoretical point of view.

(d) The estimated coefficient of MALE implies that a male financial aid applicant will receive $1570 less in grant aid than a female applicant, holding constant PARENT and HSRANK.

If we switch from MALE to FEMALE, the equation would be identical except that the estimated coefficient of FEMALE would be +1570 and the estimated constant term would adjust accordingly (to 8243).

(e) *Theory*: When asked, most colleges will state that they award financial aid without regard to gender, but liberal arts colleges attract more females than males, so it’s possible that a particular college might try to tilt its financial aid toward males. Given this possibility, and even taking the charge of bias into account, the theory behind MALE is weak. This provides evidence that MALE doesn’t belong in the equation.

*t-score*: The absolute value of the *t*-score is greater than the new critical t-value of 1.680, but the sign of the *t*-score is opposite that implied by HA, so we cannot reject the null hypothesis, providing evidence that MALE doesn’t belong in the equation.

: increases when MALE is added, providing evidence that the variable belongs in the equation.

*bias*: Neither estimated slope coefficient changes by anything close to a standard error when MALE is added to the equation, providing evidence that omitting MALE from the equation does not cause any bias.

Three of the four specification criteria favor Equation 6.21, so we prefer Equation 6.21 to Equation 6.20, particularly because theory is the most important of the criteria. However, the significant unexpected sign in Equation 6.21 cannot be ignored. It indicates that there very likely is an omitted variable in Equation 6.21. Since we were concerned about the possibility of an omitted variable on theoretical grounds already, this empirical evidence is very convincing. In essence, *neither* equation is the best equation! Most beginning econometricians will not be very happy with this answer, but it’s an important learning opportunity.

6-7. (a) *Theory:* If PERCENT is the best proxy available for the quality and reliability of the seller, then it has a strong theoretical basis until a better variable can be found.

*t-score:* The coefficient is in the expected direction, but it’s insignificant at the 5 percent level.

:  is not given, but it turns out that the addition of any variable with a t-score greater than one in absolute value will increase . 

*Bias:* none of the coefficients change by a standard error.

Thus, the four criteria are inconclusive. Theory and  support keeping PERCENT, but the other two criteria don’t. However, because PERCENT appears to be the best available measure of seller quality, and because the sign of the coefficient is in the expected direction, we’d tend to keep PERCENT.

(b) In theory, PERCENT seems like the best we can do, but it might be an unreliable measure if there are very few transactions.

(c) When you drop PERCENT from the equation,  falls from .434 to .431.

1. Chapter 7

7-3. (a) Linear in the coefficients but not the variables

(b) Linear in the coefficients but not the variables

(c) Linear in the coefficients but not the variables

(d) Nonlinear in both

(e) Nonlinear in both

7-5. (a) The Midwest (the fourth region of the country).

(b) Perfect multicollinearity! Including the omitted condition as a variable will cause the dummies to sum to a constant (1.0). This constant will be perfectly collinear with the constant term.

(c) Positive.

(d) Most correct = iii, least correct = i.

7-7. Since the equations are double-log, the elasticities are the coefficients themselves

|  |  |  |
| --- | --- | --- |
| Industry | Labor | Capital |
| Cotton | 0.92 | 0.12 |
| Sugar | 0.59 | 0.33 |

(b) The sum is an estimate of whether or not returns to scale are constant, increasing or decreasing. In this example, Cotton is experiencing increasing returns to scale while Sugar is experiencing decreasing returns to scale.

(c) This question contains a hidden difficulty in that the sample size purposely is not given. “D” students will give up, while “C” students will use an infinite sample size. “B” students will state the lowest sample size at which each of the coefficients would be significantly different from zero (listed below), and “A” students will look up the article in *Econometrica* and discover that there were 125 cotton producers and 26 sugar producers, leading to the tCs and hypothesis results listed below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Coefficient: |  |  |  |  |
| Hypoth. sign: |  |  |  |  |
| t-value: | 30.667 | 3.000 | 4.214 | 1.914 |
| Lowest d.f. at |  |  |  |  |
| which signif. (5%) | 1 | 2 | 2 | 7 |
| 5% tC given |  |  |  |  |
| actual d.f. | 1.645 | 1.645 | 1.714 | 1.714 |

All four coefficients are significantly different from zero in the expected direction.

1. Chapter 8

8-3. Perfect multicollinearity; each can be stated as an exact function of the other two. To solve the perfect multicollinearity problem, one of the three explanatory variables must be dropped.

8-5. (a) Don’t change your regression just because a student says you are going to have a problem. In particular, even if you do have multicollinearity, you may well end up doing nothing about it.

(b) There is a reasonable (.36) with all low *t*-scores (the highest is 0.84). Furthermore, the simple correlation coefficient between HR and RBI is 0.93. Also, the VIFs for HR and RBI are > 5.

(c) Since a sample of eight is extremely small, the first solution to try is to increase the sample size. In this particular case, a larger sample doesn’t rid the equation of damaging multicollinearity, so we’d favor dropping one of the redundant variables. There also are a number of potential omitted variables.

8-7. (a) 2.35, 2.50, 1.18.

(b) 1.01, 1.05, 1.04

(c) Since X1 and X2 are the only independent variables in the equation, VIF(X1) must equal VIF(X2) and hence VIF(X1) = 3.8.

(d) In a two-variable equation, r2 = R2. Thus R2 = (0.80)2, and VIF(X1) = VIF(X2) = 1/(1 – 0.64) = 2.78.

1. Chapter 9

9-3. The coefficient estimates for all three orders are the same:  The Durbin- Watson d results differ, however:

(a) DW  3.08

(b) DW  3.08

(c) DW  0.64

Note that any order change will be likely to change the DW except for the reverse order (for which DW will always be exactly the same).

9-5. (a) As we’ve mentioned, we prefer a one-sided Durbin-Watson d test, so with K  3 and N  40, the 5% critical values are dL  1.34 and dU  1.66. Since DW  0.85 is less than DL, we can reject the null hypothesis of no positive serial correlation.

(b) Coefficient:   

Hypoth. sign:   

t-value: 0.04 2.6 3.0

tC  2.423 do not reject reject

(1% one-sided reject

with 40 – closest to 36 in Table B-1 – d.f.)

(c) The estimated coefficient of P looks reasonable in terms of size and significance, but the one for L looks pathetically small. We would never expect such similar variables to have such dramatically different coefficients. Some students will want to drop L, pointing out that Laker fans love the team so much that they’ll come to games whether the team is playing well or not. We’d guess that continued losing seasons would show the true relevance of this variable, however.

(d) Pure serial correlation is certainly a possibility, but the fact that some fans “are most interested in games played late in the season” implies that an omitted variable with a temporal pattern exists. We’d want to include such a variable before concluding that pure serial correlation exists.

(e) We prefer dropping the first observation to including zeroes for L and P, but an even better alternative might be to use last season’s winning percentages as proxies for this season’s for opening day (or even a few games thereafter). While opening day might have always sold out in the past, there is no guarantee that it always will be sold out in the future.

9-7. (a) With a 5 percent, one-sided test and N = 19, K = 1, the critical values are dL = 1.18 and dU = 1.40. Since d = 0.48< 1.18, we can reject the null hypothesis of no positive serial correlation and conclude that there's evidence of positive serial correlation. (We ran a one-sided test because impure serial correlation caused by an incorrect functional form tends to be positive.)  Some students will complain that this is a cross-sectional data set and therefore "can't" have serial correlation, but order matters in this sample, so the Durbin-Watson test provides useful information.

(b) There are a number of reasons, but the most convincing to students is that a linear form can produce probabilities above one and below zero if you plug in extremely small and large distances (respectively).

(c) 1.40 > 1.22 > 1.18, so the answer is inconclusive.

(d) 9.33, but, as we’ll learn in Chapter 13, neither equation is perfect because estimated percentages are not limited by zero and one, even though in theory they should be.

1. Chapter 10

10-5. (a) Stock and Watson accurately describe the standard practice of many experienced econometricians. However, HC standard errors can be different from OLS standard errors in equations where the Breusch-Pagan and White tests do not indicate heteroskedasticity, and we prefer to use OLS SE’s in those situations.

(b) We think it’s crucial that beginning econometricians understand what heteroskedasticity is and how to combat it, so we continue to see benefits in covering heteroskedasticity in an elementary text.

10-7. (a) To test for serial correlation, first run:

|  |
| --- |
|  |
| (0.05) (0.16) |
| t  4.50 2.85 |

N  58 (monthly)  = .556 DW = 1.54

Since DW  1.54, the Durbin-Watson test is inconclusive at the 5% one-sided level. Lott and Ray, the source of these data, reach the same inconclusion but with slightly different numbers.

(b) As mentioned in the text, we do not recommend running GLS if the DW is in the inconclusive range. Our major reason is that a poor estimate of rho can introduce bias into an equation while pure serial correlation will not. This is especially true when the *t-*scores are not being used to decide whether to drop a variable, as is the case in this example.

(c) A mere doubling in the size of the dependent variable should not, in and of itself, cause you to be concerned about heteroskedasticity in a time series equation. If the dependent variable had gone up ten times, then heteroskedasticity (or nonstationarity, depending on the situation) would be a real concern.

(d) This answer depends on which variables you include as potential proportionality factors. If you use I and ln(1 + V), R2 = .0974 and NR2 = 5.65 < 5.99, the 5% chi square critical value with 2 DF. If you use t, R2 = .0454 and NR2 = 2.64 < 3.84, the 5% chi square critical value with 1 DF. We don’t recommend using all three simultaneously.

(e) Our recommendation is not to run GLS with an inconclusive DW.

(f) Our first instinct would be to use HC standard errors, but we’d do so only after investigating the possibility of nonstationarity, to be discussed in Chapter 12. Nonstationarity is a completely reasonable concern in a time-series study of the Brazilian black market for dollars.

1. Chapter 12

12-3. (a)   –243  5.2ADt  1.9ADt–1  3.1ADt–2  1.0ADt–3  3.3ADt–4

(b)   –38.86  2.98ADt  0.79SALESt–1

The lag structure in the ad hoc distributed lag equation makes no economic sense, because the estimated coefficients don’t follow the smoothly declining pattern that economic theory would suggest and that results from using a dynamic model.

12-5. We suggest that the farmers rethink either the form of their equation or their expectations. Their current equation is a dynamic model, so it posits that corn growth is a distributed lag function of rainfall, a not unreasonable idea. However, lambda is restricted to between zero and one, so the likelihood of observing a negative lambda is small, and in theory a negative lambda would be very difficult to explain for the impact of rainfall on corn.

12-7. (a) Such a split result is not unusual in correctly done applications of Granger causality when there is no real underlying causal relationship.

(b) Based on this research, it’s impossible to draw any general conclusions about the causal relationship between economic growth and democracy for two good reasons. First, the results are inconclusive. Second, reaching a conclusion about causality involves more than just the results of a Granger causality test, so even if the results for all 32 countries had provided evidence of Granger causality in the same direction, we would not feel justified in drawing a conclusion about the relationship between economic growth and democracy unless the theory was quite strong.

(c) An interesting next step in this research project would be to see if national characteristics shed any light on the results of the Granger causality test. To do this, we’d research the literature to find the attributes of a country that might impact the direction of the Granger causality, and then we’d estimate a model where the dependent variable would be the direction of the Granger causality and the independent variables would be these national attributes. Since the dependent variable in this case would be a dummy variable, we’d probably estimate the equation with a logit, and we’ll cover dummy dependent variable techniques in the next chapter.

1. Chapter 13

13-3. (a) Coefficient βUNIT  βALCO βYEAR βGREK

Hypothesized sign: + – – –

Calculated *t*-score: 0.84 –1.55 –8.25 –1.38

tc = 1.289, so: insig. insig. sig. sig.

Thus the results support their hypotheses only for YEAR and GREK.

(b) Defining YEAR this way constrains the coefficients of three classes to be fixed in their relationship to each other when there is no reason to expect that to be the case. For example, we’d expect seniors to be far more likely to live off campus than juniors or sophomores, and this definition wouldn’t allow that to happen.

A much better approach would have been to define two dummy variables, one equal to 1 for seniors (0 otherwise) and one equal to 1 for juniors (0 otherwise), which would make being a sophomore the omitted condition. We’d expect a positive coefficient for each variable, with the coefficient of senior being substantially larger than the coefficient of junior.

(c) The estimate of βALCO tells us that for each additional night (per week) that a student consumes alcohol, the log of the odds that that student will live on campus will decrease by 0.13, holding constant the other independent variables in the equation. If we divide 0.13 by four, thus turns out to be equivalent to saying that that for each additional night (per week) that a student consumes alcohol, probability of that student living on campus will decrease by 3.25 percentage points, holding constant the other independent variables in the equation. This is a little lower than we might have expected, but it certainly is plausible.

(d) We’d first fix the definition of YEAR as suggested in part b. After that, we’d add one of a number of potentially relevant variables, for instance the gender of the *i*th student or whether the *i*th student’s home was within 10 miles of campus.

13-5. (a) The trick here is getting the expected sign right, because it won’t be obvious to everyone that DISTANCE can be negative (if the patient lives farther from Cedars Sinai than they do from UCLA). Once you take this into account, it’s clear that the larger DISTANCE is, the less likely the *i*th patient is to choose Cedars Sinai, so the expected sign of the coefficient is negative, and we can reject the null (tc = 1.645).

(b) For every extra mile that it takes a patient to get to Cedars Sinai as compared to UCLA, the probability of that patient choosing Cedars Sinai falls by 9.5%, holding constant INCOME and OLD. [9.5% is the coefficient of distance (0.38) divided by 4.]

(c) Our guess is that most patients care about the relative distance to the two hospitals, not the absolute values of the individual distances to the hospitals.

(d) The coefficient of DISTANCE in the linear probability model is -0.072. We try to avoid estimating linear probability models when the dependent variable is a dummy variable because of the unlimited range of the dependent variable, so we have a strong preference for the logit in this and most other examples.

(e) The hypothesis behind this interaction term is that an elderly patient might be more likely than a younger patient to try to minimize the distance traveled to a hospital because of the limited mobility of elderly patients. Thus:

H0: β ≥ 0

HA: β < 0

Sure enough, the estimated coefficient of the interaction term is negative and produces a z-score of -3.03, which is greater than the critical value and is in the expected direction, so we can reject H0. We prefer the slope dummy logit, and all four of our specification criteria support that preference.

1. Chapter 14

14-3. (a) the first two equations are simultaneous, but the third equation is a recursive equation that feeds into the first two, so Y3 is not simultaneously determined.

Endogenous variables = Y1t, Y2t

Predetermined variables: Y3t, X1t, X1t-1, X2t-1, X3t, X4t, X4t-1

(b) All three equations are simultaneous. Note that Y is predetermined.

Endogenous variables = Zt, Xt, Ht

Predetermined variables: Yt, Pt-1, Bt, CSt, Dt

(c) The equations are recursive; solve for Y2 first and use it to get Y1.

14-5. (a)   –267  0.19Yt – 9.26rt–1

(0.01) (11.19)

t  15.87 –0.83

 .956 N  32 DW  0.47

(b)  –258.6 + 0.78Gt – 0.37NXt + 1.52Tt + 0.67COt–1 + 37.6rt–1

 .999 N  32

(c)   –261.5  0.19 – 9.55rt–1

(Standard errors obtained from this estimation are biased and should be disregarded.)

(d)   –261.5  0.19 – 9.55rt–1

(0.01) (11.2)

t  15.8 –0.85

 .956 N  32 DW  0.47

14-7. (a) QU: –, –, –, +, +, +

UR: +, +, +, +, +,

(b) Yes, since UR and QU are jointly determined in this system.

(c) This tells us that the UR equation is exactly identified but tells us nothing about the identification properties of the QU equation.

(d) The lack of significance makes us wonder if UR and QU are indeed simultaneously determined. We should be hesitant to jump to this conclusion, however, for three reasons. First, the theory indicates simultaneity; second, multicollinearity or other specification problems may be causing the insignificance; and third, the pooled data set makes it tricky to draw such a negative inference.

(e) Given the above reservations, we should be cautious. However, the results tend to support the theory that states interested in lowering their unemployment rates and lowering their budget deficits might consider lowering their unemployment benefits.

1. Chapter 15

15-3. For most students, the size of the residuals in this exercise will be surprisingly high.

15-5. Model A Model T

07 30.50 29.50

08 30.25 30.25

09 30.13 29.87

08 31.50 28.50

09 30.75 30.75

(c) Model A should exhibit smoother behavior because the negative coefficient in Model T will cause the forecast to fluctuate on an annual basis.

1. Chapter 16

16-3. In theory, one could design a random assignment experiment for which no additional explanatory variables would be necessary, but it’s virtually impossible to imagine a natural experiment not needing such variables. There are sure to be some differences between the “control” and the “treatment” groups, and we need to account for those differences.

16-5. (a) The estimated slope is positive, which certainly runs counter to our expectations:

 = –1.42 + 0.0457P

(0.014)

t = 3.28

N = 4  = .76

(b) While the fit and the size of the estimated coefficients differ from those in Part a, the sign of the estimated slope coefficient continues to be unexpected.

 = –0.22 + 0.0229P

(0.014)

t = 1.64

N = 4  = .36

(c) Some econometric software programs, including Stata, won’t be able to run this because of multicollinearity, but those programs that can produce an estimate do indeed show that the sign reverses.

(d) As expected, the fixed effects model is superior.