Answers to Odd-Numbered Exercises

Chapter One: An Overview of Regression Analysis

1-3.  (a) Positive, (b) negative, (c) positive, (d) negative, (e) ambiguous, (f) negative.

1-5.  (a) The coefficients in the new equation are not the same as those estimated in the previous equation because the sample is different. When the sample changes, so too can the estimated coefficients. In particular, the constant term can change substantially between samples; in our research for this exercise, for instance, we found one sample that had a negative intercept (and a very steep slope).

(b) Equation 1.21 has the steeper slope (6.38 > 4.30) while Equation 1.24 has the greater intercept (125.1 > 103.4). They intersect at 9.23 inches above 5 feet (162.3 pounds).

(c) Equation 1.24 misguesses by more than ten pounds on exactly the same three observations that Equation 1.21 does, but the sum of the squared residuals is greater for Equation 1.24 than for Equation 1.21. This is not a surprise, because the coefficients of Equation 1.21 were calculated using these data.

(d) If it were our last day on the job, we’d probably use an equation that we’d calculate from both equations by taking the mean, or by taking an average weighted by sample size, of the two.

1-7.  (a) The estimated slope coefficient of 5.77 represents the change in the size of a house (in square feet) given a one thousand dollar increase in the price of the house. The estimated intercept of 72.2 is the value of SIZE when PRICE equals zero; it includes the constant value of any omitted variables, any measurement errors, and/or an incorrect functional form.

(b) No. All we have shown is that a statistical relationship exists between the price of a house and its size.

(c) The new slope coefficient would be 0.00577 (5.77/1000), but nothing else would change.

1-9.  (a) 101.40 is the estimated constant term, and it is an estimate of the bond price when the Federal funds rate equals zero, but since zero is outside the normal range of observation, this interpretation doesn’t mean much (as is often the case). −4.78 is an estimate of the slope coefficient, and it tells us by how much bond prices will change when the federal funds rate increases by one unit (one percentage point). The sign of the estimated slope coefficient is as expected; as interest rates rise, the price of existing bonds is expected to fall. We typically do not develop hypotheses involving constant terms.

(b) See equation 1.17; the equation could have been equivalently stated as $Y_i = 101.40 - 4.78X_i + e_i$

(c) An error term is unobservable and shouldn’t be included in an estimated equation from which we actually calculate a $\hat{Y}$. If the question is reworded to ask why a residual isn’t included, then the answer becomes the same as for part (b) above.
(d) If the Federal funds rate increases by 1 percentage point, bond prices will fall by $4.78. Possible criticisms are: (1) The appropriate interest rate to use when explaining the value of a long-term asset is the long-term interest rate, not the short-term interest rate. (2) It seems reasonable that there are more than just one relevant explanatory variable, and (3) Given how responsive capital markets are to interest rate changes, it seems likely that a monthly (or even more frequent) data set would provide more observations (and therefore a better fit) for the same number of years.

1-11. (a) The error term is the theoretical, unobservable, difference between the true (population) regression line and the observed point. The residual is the measured difference between the observed point and the estimated regression line.

<table>
<thead>
<tr>
<th>$Y_i$</th>
<th>2</th>
<th>6</th>
<th>3</th>
<th>8</th>
<th>5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_i$</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$e_i$</td>
<td>0.20</td>
<td>0.24</td>
<td>-0.12</td>
<td>0.92</td>
<td>0.56</td>
<td>-1.76</td>
</tr>
<tr>
<td>$\epsilon_i$</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
<td>0.50</td>
<td>0.50</td>
<td>-2.00</td>
</tr>
</tbody>
</table>

1-13. (a) Negative, because a high unemployment rate would imply that jobs are scarce; as a result, employers would not need to advertise much, if at all, to attract applicants for a job opening. Yes.

(b) While 97% of the variation of HWI around its mean can be explained by UR in this sample, it seems possible that other variables might matter as well, especially over a longer time period. When reviewing suggested additional independent variables, the key at this point is to make sure that every variable (like the cost of help-wanted advertising or the efficiency of networking alternatives to advertising) is time-series, not cross-sectional.

(c) Yes. The EViews estimates are $\hat{\beta}_0 = 364.0404$ and $\hat{\beta}_1 = -46.39477$.

### Chapter Two: Ordinary Least Squares

2-3. (a) Because of Equation 2.5, $\hat{\beta}_0$ and $\hat{\beta}_1$ tend to compensate for each other in a two-variable model. Thus if $\hat{\beta}_1$ is too high, then $\hat{\beta}_0$ is likely to be too low.

(b) There are only two possibilities:
   
   i) $\hat{\beta}_0 = \hat{\alpha}_0 + \hat{\alpha}_2 X_2$
   
   ii) $\hat{\alpha}_2 = 0$ (and therefore $\hat{\beta}_0 = \hat{\alpha}_0$)

2-5. (a) The squares are “least” in the sense that they are being minimized.

(b) If $R^2 = 0$, then $RSS = TSS$, and $ESS = 0$. If $R^2$ is calculated as $ESS/TSS$, then it cannot be negative. If $R^2$ is calculated as $1 - RSS/TSS$, however, then it can be negative if $RSS > TSS$, which can happen if $\hat{Y}$ is a worse predictor of $Y$ than $\bar{Y}$ (possible only with a non-OLS estimator or if the constant term is omitted).

(c) We prefer Model T because it has estimated signs that meet expectations and also because it includes an important variable (assuming that interest rates are nominal) that Model A omits. A higher $R^2$ does not automatically mean that an equation is preferred.
2-7.  (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, $X_4$ is a theoretically sound variable for a campus track, while $X_3$ 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 $\hat{\beta}$ has a different meaning.

2-9.  As the students will learn in Chapters 6 and 7, there’s a lot more to specifying an equation than maximizing $R^2$.

2-11.  (a) It might seem that the higher the percentage body fat, the higher the weight, holding constant height, but muscle weighs more than fat, so it’s possible that a lean, highly muscled man could weigh more than a less well-conditioned man of the same height.

(b) We prefer Equation 1.24 because we don’t think F belongs in the equation on theoretical grounds. The meaning of the coefficient of X changes in that F now is held constant.

(c) The fact that $R^2$ drops when the percentage body fat is introduced to the equation strengthens our preference for 1.24.

(d) This is subtle, but since 0.28 times 12.0 equals 3.36, we have reason to believe that the impact of bodyfat on weight (holding constant height) is very small indeed. That is, moving from average bodyfat to no bodyfat would lower your weight by only 3.36 pounds.


Chapter Three: Learning to Use Regression Analysis

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) $\hat{S}_i = 36721 + 817(1) + 426(0) + 406(3) + 3539(1) = \hat{S}_i = $42,295. This is an approximation because the actual research included a number of other independent variables. Luckily, $42,295 is fairly close to what the entire equation cranks out, because most third-year assistant professors have zeroes for the other variables in the equation. (By the way, this estimate is ten years old, so it’s a bit low.)
3-5.  (a) Positive.
(b) Obviously, the best equation includes the actual traffic data (which, it turns out, are available).
Since the traffic dummy variable is correlated with the actual traffic variable and since the
new equation has expected signs and reasonable coefficients, it seems slightly better than
Equation 3.5.
(c) No! The theoretical underpinnings of the model are much more important. Of course, the
higher $R^2$ is certainly a plus.

3-7. Most students at this stage can exactly duplicate the Woody’s results.

3-9. (a) Negative, positive, none
(b) Holding all other included explanatory variables constant, a car with an automatic
transmission gets 2.76 miles less per gallon than a model with a manual transmission, and a
car with a diesel engine gets 3.28 miles more per gallon than one without a diesel engine.
(c) Lovell added the EPA variable because he wanted to test the accuracy of EPA estimates. If
these estimates were perfectly accurate, then the EPA variable would explain all the variation
in miles per gallon.

Chapter Four: The Classical Model

4-3. (a) An additional pound of fertilizer per acre will cause corn yield (bushels/acre) to increase by
0.10 bushel/acre, holding rainfall constant. An additional inch of rain will increase corn yield
(bushels/acre) by 5.33 bushels/acre holding fertilizer/acre constant.
(b) No. (This is a typical student mistake.) First, since it’s hard to imagine zero inches of rain
falling in an entire year, this intercept has no real-world meaning. In addition, recall that the
OLS estimate of the intercept includes the non-zero mean of the error term in order to validate
Classical Assumption II (as explained in the text), so even if rainfall were zero, it wouldn’t
make sense to attempt to draw inferences from the estimate of the $\beta_0$ term unless it was
known that the mean of the (unobservable) error term was zero.
(c) No; this could be an unbiased estimate. 0.10 is the estimated coefficient for this sample, but
the mean of the coefficients obtained for the population could still equal the true $\beta_r$.
(d) Not necessarily; 5.33 could still be close to or even equal to the true value. An estimated
coefficient produced by an estimator that is not BLUE could still be accurate. If the estimator
is biased, its bias could be small and its variance smaller still.

4-5. (a) Classical Assumption II.
(b) Classical Assumption VI.
(c) R: A one-unit increase in yesterday’s R will result in a 0.1% increase in today’s Dow Jones
average, holding constant the other independent variables in the equation.
M: The Dow Jones will fall by 0.017% on Mondays, holding constant the other variables in
the equation.
(d) Technically, C is not a dummy variable because it can take on three different values. Saunders assumed (at least implicitly) that all levels of cloud cover between 0% and 20% have the same impact on the Dow and also that all levels of cloud cover between 21% and 99% have the same impact on the Dow. In addition, by using the same variable to represent both sunny and cloudy days, he constrained the coefficient of sun and cloud to be equal.

(e) In our opinion, this particular equation does little to support Saunders’ conclusion. The poor fit and the constrained specification combine to outweigh the significant coefficients of R_{t-1} and M.

4-7. (a) The estimated coefficient of C shows that (for this sample) a one percent increase in the nonwhite labor force in the ith city adds 0.002 percentage points to the overall labor force participation rate in that city, holding constant all the other independent variables in the equation. The estimated coefficient of the dummy variable, D, shows that if a city is in the South, the labor force participation rate will be 0.80 percentage points lower than in other cities, holding constant the other explanatory variables in the equation.

(b) Perfect collinearity is virtually impossible in a cross-section like this one because no variable is a perfect linear function of another; some are closely related, but none is a perfect linear function.

(c) This does not imply that one of the estimates is biased. The estimates were taken from two different samples and are quite likely to differ. In addition, the true value may have changed between decades.

(d) Disagree. Beginners often confuse the constant term with the mean of the dependent variable. While the estimated constant term shows the value of the dependent variable in the unlikely case that all of the explanatory variables equal zero, it also includes the mean of the observations of the error term as mentioned in question 3, part (b) above.

4-9. (a) This possibly could violate Assumption III, but it’s likely that the firm is so small that no simultaneity is involved. Well cover simultaneous equations in Chapter 14.

(b) Holding constant the other independent variables, the store will sell 134.4 more frozen yogurts per fortnight if it places an ad. If we ignore long-run effects, this means that the owner should place the ad as long as the cost of the ad is less that the increase in profits brought about by selling 134.4 more frozen yogurts.

(c) The result doesn’t disprove the owner’s expectation. School is not in session during the prime yogurt-eating summer months, so the variable might be picking up the summer time increases demand for frozen yogurt from non-students.

(d) Answers will vary wildly, so perhaps it’s best just to make sure that all suggested variables are time-series for two-week periods. For students who have read Chapters 1–4 only, the best answer would be any variable that measures the existence of, prices of, or advertising of local competition. Students who have read Chapter 6 might reasonably be expected to try to find a variable whose expected omitted-variable bias on the coefficient of C is negative. Examples include the number of rainy days in the period or the number of college students returning home for vacation in the period.
Chapter Five: Hypothesis Testing

5-5. For $\beta_N$: Reject $H_0: \beta \geq 0$ if $[-4.42] > t_c$ and $-4.42$ is negative.

For $\beta_p$: Reject $H_0: \beta \leq 0$ if $[4.88] > t_c$ and $4.88$ is positive.

For $\beta_i$: Reject $H_0: \beta \leq 0$ if $[2.37] > t_c$ and $2.37$ is positive.

(a) $t_c = 1.943$; reject the null hypothesis for all three coefficients.
(b) $t_c = 1.311$; reject $H_0$ for all three coefficients.
(c) $t_c = 6.965$; cannot reject the null hypothesis for any of the three coefficients.

5-7. This is a concern for part (a) but not for parts (b) and (c). In part (a), 160 probably is the coefficient we expect; after all, if our expectation was something else, why did we specify 160 in the null? In parts (b) and (c), however, it seems unlikely that we’d expect zero.

5-9. (a) For all three, $H_0: \beta \leq 0$, $H_0: \beta > 0$, and the critical 5% one-sided $t$-value for 24 degrees of freedom is 1.711. For LOT, we can reject $H_0$ because $[+7.0] > 1.711$ and $+7.0$ is positive. For BED, we cannot reject $H_0$ because $[-1.0] < 1.711$ even though $-1.0$ is positive. For BEACH, we can reject $H_0$ because $[-10.0] > 1.711$ and $-10.0$ is positive.

(b) $H_0: \beta \geq 0$, $H_0: \beta < 0$, and the critical 10% one-sided $t$-value for 24 degrees of freedom is 1.318, so we reject $H_0$ because $[-2.0] > 1.318$ and $-2.0$ is negative.

(c) $H_0: \beta = 0$, $H_0: \beta \neq 0$, and the critical 5% two-sided $t$-value for 24 degrees of freedom is 2.064, so we cannot reject $H_0$ because $[-1.0] < 2.064$. Note that we don’t check the sign because the test is two-sided and both signs are in the alternative hypothesis.

(d) The main problems are that the coefficients of BED and FIRE are insignificantly different from zero.

(e) Given that we weren’t sure what sign to expect for the coefficient of FIRE, the insignificant coefficient for BED is the most worrisome.

(f) Unless the students have read Chapter 6, this will be a difficult question for them to answer. It’s possible that the dataset is unrepresentative, or that there’s an omitted variable causing bias in the estimated coefficient of BED. Having said that, the most likely answer is that BED is an irrelevant variable if LOT also is in the equation. Beach houses on large lots tend to have more bedrooms than beach houses on small lots, so BED might be irrelevant if LOT is included.

5-11. (a) For the $t$-tests:

<table>
<thead>
<tr>
<th>Coefficient:</th>
<th>$\beta_p$</th>
<th>$\beta_M$</th>
<th>$\beta_S$</th>
<th>$\beta_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypoth. Sign:</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$t$-value:</td>
<td>5.8</td>
<td>6.3</td>
<td>1.0</td>
<td>-3.3</td>
</tr>
<tr>
<td>$t_c = 1.671$</td>
<td>reject</td>
<td>reject</td>
<td>do not</td>
<td>reject</td>
</tr>
<tr>
<td>(5% one-sided with 60 d.f., as close to 73 as Table B–1 goes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(b) No. We still agree with the authors’ original expectations despite the contrary result.

(c) Keynes’ point is well taken; empirical results will indeed allow an econometrician to discover a theoretical mistake now and then. Unfortunately, far too many beginning researchers use this loophole to change expectations to get “right” signs without enough thinking or analysis.

(d) Holding all other included explanatory variables constant, an increase in winning percentage of 150 points will increase revenues by $7,965,000 ($53.1 times 150 times 1000) and thus it would be profitable for this team to hire a $4,000,000 free agent who can raise its winning percentage to 500 from 350.

5-13. (a) All the expected signs for the coefficients are positive, so: $H_0: \beta \leq 0, \ H_A: \beta > 0$. (Some students will see the Wall Street Journal as a competitor and will hypothesize a negative coefficient for J, but the authors intended the Journal’s circulation as a measure of what the Post’s circulation would have been without Watergate.)

(b) Coefficient: $\beta_j, \beta_s, \beta_W$
Hypoth. Sign: + + +
t-value: 14.27 6.07 1.31
$\tau_c = 1.717$ reject reject do not reject (5% one-sided with 22 d.f.)

(c) Assuming that the specification is correct, Watergate had a positive but statistically insignificant effect on the Post’s circulation.

Chapter Six: Specification: Choosing the Independent Variables

6-3. (a) Coefficient: $\beta_c, \beta_E, \beta_M$
Hypoth. Sign: + + +
t-value: 4.0 4.0 −2.0
$\tau_c = 1.314$ reject reject do not reject (10% one-sided with 27 d.f.)

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

(b) The 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 $\beta_M$. We thus need 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.
So:

<table>
<thead>
<tr>
<th>Possible Omitted Variable</th>
<th>Expected Sign of $\beta$</th>
<th>Correlation with M</th>
<th>Direction of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>F</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>W</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>R</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>H</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>O</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

* indicates a weak expected sign or correlation.

(c) The only one of the above variables that seems likely to have caused any of the observed negative bias in $\hat{\beta}_M$ is H, and that variable is a good proxy for the obvious omission of the general level of quality of cardiac health care, so we’d choose to add it. We wouldn’t overly penalize a student who chose to add another variable as long as the alternative was an annual aggregate variable. (Students who add disaggregate variables would probably benefit from some extra review sessions.)

6-7. Some students will come to the conclusion that sequential specification searches are perfectly reasonable in business applications, and they need to be reminded of the biased estimated coefficients generated by such searches.

6-9. (a) In a supply equation, the coefficient of price will have a positive expected sign because the higher the price, holding all else constant, the more suppliers would be willing to produce.

(b) The price of inputs (such as labor, seeds, transportation, machinery, and fertilizer), the price of a “production substitute” (a crop or product that could be produced instead of the crop or product being modeled), and exogenous factors (like local growing conditions, local strikes that don’t have an impact on the price, etc.) are just a few examples of important variables in a supply-side equation.

(c) Lag those independent variables that influence the production decision on a delayed basis. In particular, lag them by the length of time it takes for that particular event to have an impact on production. For example, if growers must make production decisions a year before the crop is harvested, then price should be lagged one year, etc. If a product can be stored at a fairly low cost, then such a lag might not be appropriate because producers could choose to wait until prices rose before going to market.

6-11. (a) Coefficient: $\beta_w$ $\beta_t$ $\beta_C$ $\beta_L$

| Hypoth. sign: | + | + | + | + |
| t-value: 4.0 | 3.0 | 2.0 | 1.0 |
| $t_c = 1.697$ | reject | reject | reject | do not reject |
| (5% one sided with 30 d.f.) | | | | |
(b) Almost any equation could potentially have an omitted variable, and this one is no exception. In addition, L might be irrelevant. Finally, the coefficient of C seems far too large, suggesting at least one omitted variable. C appears to be acting as a proxy for other luxury options or the general quality of the car. (Some students will claim that they shouldn’t be expected to know that 5.8 is too large. Even students who don’t know what cruise control is, however, should be expected to know that $5,800 is too large a portion of a car’s price to be a reasonable amount to add to that price if the car has cruise control, whatever it is!)

(c) We prefer T. Using the four criteria:

Theory: Engine size seems a small component of car price, especially for sedans.

t-score: L’s coefficient is insignificant at the 5% level.

$R^2$: The overall fit adjusted for degrees of freedom does not change when L is dropped.

Bias: The other coefficients do not change significantly when L is dropped.

(d) The coefficient of L had a t-score of exactly 1.0, so dropping L will have no impact on $R^2$. If the t-score had been less than 1.0, dropping L would have increased $R^2$, and if the t-score had been greater than 1.0, dropping L would have decreased $R^2$.

6-13. This question is similar to Exercise 6-8 but since the answers to odd-numbered exercises are not in the back of the text, 13 is more suitable for use on a problem set or exam.

(a) $X_1$ = either dummy variable

$X_2$ = either dummy variable

$X_3$ = Parents’ educational background

$X_4$ = Iowa Test score

(b) We have two variables for which we expect positive coefficients (Iowa score and Parents’ education) and two positive estimated coefficients ($\hat{\beta}_3$ and $\hat{\beta}_4$), so we’d certainly expect $X_3$ and $X_4$ to be those two variables. In choosing between them, it’s fair to expect a more significant coefficient for Iowa than for Parents. Next, we have two variables for which we expect a zero coefficient (the dummies) and two estimated coefficients ($\hat{\beta}_1$ and $\hat{\beta}_2$) that are not significantly different from zero, so we’d certainly expect $X_1$ and $X_2$ to be the dummies. There is no evidence to allow us to distinguish which dummy is $X_1$ and which is $X_2$. (Students who justify this answer by expecting negative signs for coefficients of the two dummies are ignoring the presence of the Iowa test score variable in the equation that holds constant the test-taking skills of the student.)

(c) Coefficient: $\beta_D$, $\beta_D$, $\beta_{PE}$, $\beta_{IT}$

Hypoth. sign: 0 0 + +

t-value: $t = -1.0$ $t = 0.25$ $t = 2.0$ $t = 12.0$

(5% two-sided with 19 d.f.)

$|t_c| = 2.093$ do not reject do not reject

(5% one-sided with 19 d.f.)

$t_c = 1.729$ reject reject

(d) As you can see, we used a one-sided test for those coefficients for which we had a specific prior expectation but a two-sided test around zero for those coefficients for which we did not.
6-15.  
(a) (i) The coefficient of CV is –0.19 with a SE (\(\hat{\beta}\)) of 0.23 and a t-score of –0.86. The \(R^2\) is .773, and the rest of the equation is extremely similar to Equation 5.14 except that the coefficient of CVN falls to –0.48 with a t-score of –1.86.

(ii) The coefficient of N is 0.00054 with a SE (\(\hat{\beta}\)) of 0.063 and a t-score of 0.0086. The \(R^2\) is .766, and the rest of the equation is identical (for all intents and purposes) to Equation 5.10.

(b) Theory: P is a price ratio, and while it’s possible that a price ratio would be a function of the size of a market or a country, it’s not at all obvious that either variable would add anything since CVN is already in the equation.

t-score: Both t-scores are insignificant.

\(R^2\): \(R^2\) falls when either variable is added.

bias: None of the coefficients change at all when N is added, so it clearly is irrelevant. The omission of CV does change the coefficient of CVN somewhat, making it likely that CV is redundant since CVN is in the equation.

(c) Since CVN = f([CV/N], it would make little theoretical sense to include all three variables in an equation, even though technically you don’t violate Classical Assumption VI by doing so.

(d) It’s good econometric practice to report all estimated equations in a research report, especially those that were undertaken for specification choice or sensitivity analysis.

Chapter Seven: Specification: Choosing a Functional Form

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.  

(b) A linear form, a double-log form, a polynomial form (with no \(\beta_1U_t\) term), and a semi-log form all could have been used (with varying success).

(c) There is evidence that macropolicies and external shocks were shifting the short-run Phillips curve upward and to the right during this period, causing a plotting of the annual inflation/unemployment rates to appear to be positively sloped. While there is strong theoretical support for a vertical long-run Phillips curve, students who let the data convince them that a short-run Phillips curve was positively sloped during this time period are making the same mistake that many Keynesians did for over a decade.
7-7. (a) This question contains a hidden difficulty in that the sample size is purposely 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 t_Cs and hypothesis results listed below.

<table>
<thead>
<tr>
<th>Coefficient:</th>
<th>$\beta_{1C}$</th>
<th>$\beta_{2C}$</th>
<th>$\beta_{1S}$</th>
<th>$\beta_{2S}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypoth. sign:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>t-value:</td>
<td>30.667</td>
<td>3.000</td>
<td>4.214</td>
<td>1.914</td>
</tr>
<tr>
<td>Lowest d.f. at which signif. (5%)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>5% t_C given actual d.f.</td>
<td>1.645</td>
<td>1.645</td>
<td>1.714</td>
<td>1.714</td>
</tr>
</tbody>
</table>

(So all four coefficients are significantly different from zero in the expected direction.)

(b) A double-log function seems quite appropriate.

(c) Since the equations are double-log, the elasticities are the coefficients themselves:

<table>
<thead>
<tr>
<th>Industry</th>
<th>Labor</th>
<th>Capital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton</td>
<td>0.92</td>
<td>0.12</td>
</tr>
<tr>
<td>Sugar</td>
<td>0.59</td>
<td>0.33</td>
</tr>
</tbody>
</table>

(d) The sum indicates 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.

7-9. (a) $\alpha$ and $\beta$ both have positive hypothesized signs because as either L or K increases holding the other constant, output should increase.

(b) $H_{0}: \beta \leq 0$; $H_{A}: \beta > 0$; for both.

(c) Reject $H_{0}$ if $|t\text{-value}| > 1.708$ and t-value is positive.

(d) L: $t = 2.02$; K: $t = 5.86$, so reject $H_{0}$ for both.

(e) The relative prices of the two inputs need to be known.

7-11. (a) polynomial (second-degree, with a negative expected coefficient for age and a positive expected coefficient for age squared)

(b) double-log (We would not quibble with those who chose a linear form to avoid the constant elasticity properties of a double-log.)

(c) semilog (lnX)

(d) linear (All intercept dummies have a linear functional relationship with the dependent variable by definition.)

(e) inverse (Most students will remember from the text that a U-shaped polynomial typically is used to model a cost-curve and will want to apply it here. The problem is that the telephone industry appears to be an industry in which costs continually decrease as size increases, making an inverse our choice.)
7-13. (a) Coefficient: \( \beta_{LQ} \quad \beta_A \quad \beta_V \)

Hypoth. sign: + – –
t-value: 4.0 –2.0 0.13
t_C = 1.725 reject reject do not
(5% one-sided reject reject do not with 20 d.f.)

(b) \( Q \) constant, \( A \) and \( V \) non-constant.

(c) No. The coefficient of \( V \) is quite insignificant, and the equation (simplified from an unpublished article) is flawed to boot. Note, however, that the violence may be causing the absentee rate to rise, so that the significant coefficient for \( A \) does indicate some support for the charge.

(d) In our opinion, this is a classic case of “spurious correlation” because actual total output appears on both sides of the equation, causing almost all of the fit by definition. If we could make one change, we’d drop \( LQ \) from the equation, but we worry that little will be left when we do.

7-15. (a) Coefficient: \( \beta_s \quad \beta_G \quad \beta_D \quad \beta_{SD} \quad \beta_{GD} \)

Hypoth. sign: + + + + ? +
t-value: 4.5 0.4 2.9 –5.0 2.3
t_C = 1.645 reject do not reject do not reject
(5% one-sided reject do not reject reject with infinite d.f.)

(b) As Primeaux puts it (on page 622 of his article), “A duopoly firm of small size spends more than a monopoly firm of the same size. However, as scale increases, eventually, the duopoly firm spends less.”

(c) Again, from page 622, “There is no difference between monopoly and duopoly firms at zero rates of growth in sales. However, as growth takes place, the duopoly firms engage in more sales promotion activity.”

Chapter Eight: Multicollinearity

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) Using the mean method:
\[ X_a = 16 \text{HEIGHT}_i + \text{WEIGHT}_i \]
Using the regression method:
\[ X_a = 6.4 \text{HEIGHT}_i + \text{WEIGHT}_i \]

(b) Using the mean method:
\[ X_a = 5I_i + P_i \]
Using the regression method:
\[ X_a = 2.7I_i + P_i \]
(c) Using the mean method:
\[ X_n = 2.01Y_i + YD_i \]
Using the regression method:
\[ X_i = 3.63Y_i + YD_i \]

8-7.  (a) No; no explanatory variable is a perfect function of another.
(b) Yes; income in any quarter will be strongly correlated with income in previous quarters.
(c) If all the variables were specified in terms of first differences, it’s likely that much of the multicollinearity would be avoided.

8-9.  (a) Coefficient: \( \beta_{PC} \quad \beta_{PQ} \quad \beta_Y \quad \beta_C \quad \beta_N \)
Hypoth. sign: + - + + +
t-value: 0.801 -1.199 0.514 -1.491 1.937
tc = 1.725 at the 5% level, so only \( \hat{\beta}_N \) is significantly different from zero in the expected direction.
(b) The obviously low t-scores could be caused by irrelevant variables, by omitted variables biasing the estimated coefficients toward zero, or by severe imperfect multicollinearity.
(c) The high simple correlation coefficient between \( Y \) and \( C \) indicates that the two are virtually identical (redundant), which makes sense theoretically. The r between the two price variables is not as high, but mild multicollinearity can still be shown to exist.
(d) \( Y \) and \( C \) both serve as measures of the aggregate buying power of the economy, so they are redundant, and one should be dropped. It doesn’t matter statistically which one is dropped, but \( Y \) seems analytically more valid than \( C \), so we’d drop \( C \). Dropping one of the price variables would be a mistake, since they have opposite expected signs. While forming a relative price variable is an option, the low level of multicollinearity, the reasonable coefficients, and the possibility that \( C \) is also multicollinear with prices (so dropping it will improve things) all argue for making just one change.

8-11. (a) Coefficient: \( \beta_M \quad \beta_B \quad \beta_A \quad \beta_S \)
Hypoth. sign: + + + +
t-value: 5.0 1.0 -1.0 2.5
tc = 1.645 reject do not reject do not reject
(5% one-sided reject reject reject with infinite d.f.)
(b) The insignificant t-scores of the coefficients of A and B could have been caused by omitted variables, irrelevance, or multicollinearity (a good choice, since that’s the topic of this chapter). In particular, since most students graduate at about the same age, the collinearity between A and B must be fairly spectacular (Stanford gave us no clues).
(c) It’s probably a good idea, since the improvement in GPA caused by extra maturity may eventually be offset by a worsening in GPA due to separation from an academic environment.
(d) We believe in making just one change at a time to best be able to view the impact of the change on the estimated regression. Thus, our first choice would be to drop either A or B (we’d prefer to drop A, but on theoretical grounds, not as a result of the unexpected sign). Switching to a polynomial before dropping one of the redundant variables will only make things worse, in our opinion.

8-13. (a) Coefficient: \( \beta_{\text{EMP}} \), \( \beta_{\text{UNITS}} \), \( \beta_{\text{LANG}} \), \( \beta_{\text{EXP}} \)

<table>
<thead>
<tr>
<th>Hypoth. sign:</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-value:</td>
<td>-0.98</td>
<td>2.39</td>
<td>2.08</td>
<td>4.97</td>
</tr>
<tr>
<td>( t_c = 1.725 )</td>
<td>do not reject</td>
<td>reject</td>
<td>reject</td>
<td>reject</td>
</tr>
</tbody>
</table>

(5% one-sided with 20 d.f.)

(b) The functional form is semilog left or semilog (ln Y). Semilog left is an appropriate functional form for an equation with salary as the dependent variable because salaries often increase in percentage terms when an independent variable (like experience) increases by one unit.

(c) There’s a chance that an omitted variable is pulling down the coefficient of EMP. However, it’s more likely that EMP and EXP are redundant (since they in essence measure the same thing) and are causing multicollinearity.

(d) These results confirm our opinion that EMP and EXP are redundant.

(e) If we knew that this particular school district doesn’t give credit for teaching experience elsewhere, then it would make sense to drop EXP. Without that specific knowledge, however, we’d drop EMP because EXP includes EMP.

(f) Theory: The two variables are redundant

- t-test: The coefficient of EMP is indeed insignificantly different from zero.
- \( \bar{R}^2 \): \( \bar{R}^2 \) does indeed increase when EMP is dropped from the equation.
- Bias: Students will have to work backwards to calculate the SE(\( \hat{\beta} \))s, but if they do, they’ll find that the coefficient of EXP does indeed change by a standard error. This is exactly what we’d expect to happen when we drop a redundant variable from an equation; the coefficient of the remaining redundant variable will adjust to pick up the effect of both variables.

8-15. The ever-innovative Rao and Miller used this example (developed by Professor Maurice G. Kendall) to show that the inspection of simple correlation coefficients is not an adequate test for multicollinearity.

(a) Since R and L are obviously correlated but R and (L – R) are not, many beginning students will want to drop either R or L from Model A, but this would leave out the difference between leg lengths that is the inherent causal variable.

(b) As Rao and Miller point out (on page 48 of their text), the implicit estimates of the coefficients are identical “because the conditions imposed on the residuals for estimation in either case are implicitly the same.” To calculate the coefficients of one model from the other, multiply out the \( \beta_2 \) term of Model B, reconfigure to correspond to Model A, and solve for the coefficients of Model B: \( \beta_0 = \alpha_0 \), \( \beta_1 = (\alpha_1 - \alpha_2) \), and \( \beta_2 = \alpha_2 \).
(c) Since the coefficient estimates are identical for every sample, their distributions must also be identical, meaning that the two models are identically vulnerable to multicollinearity.
(d) If you drop L from Model A, then the linkage between the Models cited in the answers above is lost.

Chapter Nine: Serial Correlation

9-3. The coefficient estimates for all three orders are the same: $\hat{HS}_t = -28187 + 16.86P_t$. 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).

9-7. (a) $d_L = 1.49$; $DW = 0.81 < 1.49$, so we’d reject the null hypothesis of no positive serial correlation.
(b) This is not necessarily a sign of pure serial correlation. It’s reasonable to think that residuals from the same country would have more in common than would residuals from other countries (that is, the model could be consistently underestimating for France and overestimating for Canada, producing six positive residuals followed by six negative residuals). As a result, the DW for such pooled datasets will at times give indications of serial correlation when it does not indeed exist. The appropriate measure is the Durbin-Watson d for each country taken individually, since the order of the countries will influence the overall DW statistic, and that order is arbitrary.
(c) If the serial correlation is impure, then a variable needs to be added to the equation to help distinguish better between the countries. If the serial correlation is judged to be pure, however, then generalized least squares might be applied one country at a time. It is possible to specify different first-order serial correlation coefficients for each country and then estimate one pooled regression equation.

9-9. (a) An outlier in the residuals can occur even if no outlier exists in the dataset if all the Xs are very low (or very high) simultaneously, producing an unusually low or high $\hat{Y}$. In such a situation, $\hat{Y}$ would be dramatically lower (or higher) than Y.
(b) When an extreme outlier exists in the residuals, the Durbin-Watson test will not necessarily produce an accurate measure of the existence of serial correlation because the outlier will give the appearance of severe negative serial correlation. That is, there will be a large $(e_t - e_{t-1})$ of one sign followed by a large $(e_t - e_{t-1})$ of the opposite sign, so the two large squared terms will move the DW dramatically toward four. In such a circumstance, some researchers will drop the outlier from the DW calculation (but not from the data set). A one-time dummy equal to one in the observation with the outlier residual will solve this problem by in essence setting the residual equal to zero; this is almost (but not quite) the same as dropping the observation.

9-11 (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 $d_L = 1.34$ and $d_U = 1.66$. Since $DW = 0.85$ is less than $D_L$, we can reject the null hypothesis of no positive serial correlation.
(b) Coefficient: \( \beta_L \quad \beta_P \quad \beta_W \)

<table>
<thead>
<tr>
<th>Hypoth. sign</th>
<th>+</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-value:</td>
<td>0.04</td>
<td>2.6</td>
<td>3.0</td>
</tr>
<tr>
<td>( t_c \approx 2.423 )</td>
<td>do not reject</td>
<td>reject</td>
<td>reject</td>
</tr>
<tr>
<td>(1% one-sided with 40 – closest to 36 in Table B-1 – d.f.)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(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. Many students will want to drop \( L \), pointing out that the Lakers “almost always play well,” so fans may not pay much attention to exactly how the Lakers are doing at any given point. We’d guess that a long losing streak 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.

9-13. (a) This is a cross-sectional dataset and we normally wouldn’t expect autocorrelation, but we’ll test anyway since that’s what the question calls for. DL for a 5% one-sided, \( K = 3 \), test is approximately 1.61, substantially higher than the DW of 0.50. (Sample sizes in Table B-4 only go up to 100, but the critical values at those sample sizes turn out to be reasonable estimates of those at 200.) As a result, we can reject the null hypothesis of no positive serial correlation, which in this case seems like evidence of impure serial correlation caused by an omitted variable or an incorrect functional form.

(b) Coefficient: \( \beta_G \quad \beta_D \quad \beta_F \)

<table>
<thead>
<tr>
<th>Hypoth. sign</th>
<th>+</th>
<th>+</th>
<th>–?</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-value:</td>
<td>3.5</td>
<td>7.0</td>
<td>–2.5</td>
</tr>
<tr>
<td>( t_c = 1.645 )</td>
<td>reject</td>
<td>reject</td>
<td>reject</td>
</tr>
<tr>
<td>(5% one-sided with infinite d.f.)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We certainly have impure serial correlation. In addition, some students will conclude that \( F \) has a coefficient that is significant in the unexpected direction. (As it turns out, the negative coefficient could have been anticipated because the dependent variable is in percentage terms but \( F \) is in aggregate terms. We’d guess that the more food a pig eats, the bigger it is, meaning that its chances of growing at a high rate are low, thus the negative sign.)

(c) The coefficient of \( D \) is significant in the expected direction, but given the problems with this equation, we’d be hesitant to conclude much of anything just yet.

(d) In this case, the accidental ordering was a lucky stroke (not a mistake), because it allowed us to realize that younger pigs will gain weight at a higher rate than their older counterparts. If the data are ordered by age, positive residuals will be clustered at one end of the data set, while negative ones will be clustered at the other end, giving the appearance of serial correlation.
9-15. (a) For the first equation (standard errors in parentheses):

\[
\ln SDL_t = -3.9 + 0.17\ln USD_t + 1.12\ln SY_t - 0.037\ln SP_t
\]

\[
(0.071) (0.062) (0.030)
\]

\[
t = 2.47 \quad 18.24 \quad -1.22
\]

\[N = 25 \text{ (annual 1960–1984)} \quad R^2 = .980 \quad DW = 0.70\]

Coefficient: \(\beta_1\) \(\beta_2\) \(\beta_3\)

Hypothesized sign: \(+\) \(+\) \(+\)
\[t_c = 1.721, \text{ so: reject reject do not reject}\]

For the second equation:

\[
\ln SDL_t = -3.29 + 0.143\ln USD_t + 1.062\ln SY_t
\]

\[
(0.067) (0.035)
\]

\[
t = 2.15 \quad 30.50
\]

\[N = 25 \text{ (annual 1960–1984)} \quad R^2 = 0.980 \quad DW = 0.59\]

Coefficient: \(\beta_1\) \(\beta_2\)

Hypothesized sign: \(+\) \(+\)
\[t_c = 1.717, \text{ so: reject reject}\]

(b) The specification criteria now imply that SP is an irrelevant variable: \(R^2\) increases slightly when SP is added, but SP’s coefficient is not significantly different from zero, and the other coefficient estimates do not change more than one standard error when SP is added.

(c) For both equations, DW is far below the critical value for a 5% one-sided test, so we can reject the null hypothesis of no positive serial correlation. (For the first equation, 0.70 < 1.12, and for the second equation 0.59 < 1.21.)

(d) Once again, such a small improvement in the DW statistic is not evidence that the serial correlation is impure.

(e) Since the SDL equations imply that SP is irrelevant, it makes sense to run GLS on the second equation:

\[
\ln SDL_t = -2.12 + 0.094\ln USD_t + 0.914\ln SY_t
\]

\[
(0.058) (0.069)
\]

\[
t = 1.64 \quad 13.15
\]

\[N = 24 \text{ (annual 1960–1984)} \quad R^2 = .994 \quad \hat{\rho} = 0.61\]
Chapter Ten: Heteroskedasticity

10-5. (a) \( \overline{CO} = 1273.2 + 0.72I_i \quad R^2 = .97 \)

\[
\begin{align*}
(0.044) \\
t &= 16.21
\end{align*}
\]

where: \( CO \) = average consumption
\( I \) = average income.

(b)

\[
\ln(e_i) = 29.54 - 2.34 \ln(I_i) \quad \overline{R^2} = .39
\]

\[
(0.94) \\
t &= -2.49
\]

\( t_c = 3.499 \) at the 1% level (two-tailed) so we cannot reject the null hypothesis of homoskedasticity.

(c) Some students will feel that when there is only one independent variable, weighted least squares (with \( X \) as the proportionality factor) will produce the same answer as OLS because you could always “multiply through” the WLS equation by \( Z \) to get the OLS result. Since the dependent variables are different, such a hypothesis can be proven to be incorrect. We can illustrate this with the data from this example by using \( I_i \) as the proportionality factor, obtaining:

\[
\frac{CO_i}{I_i} = \beta_0/I_i + \beta_1 + u_i
\]

(where \( u_i \) is a homoskedastic error term)

weighted least squares produces:

\[
\frac{\overline{CO_i}}{I_i} = 2400.1/I_i + 0.40 \quad \overline{R^2} = .94
\]

\[
(0.14) \\
t &= 2.77
\]

Thus even with only one independent variable, weighted least squares does indeed make a difference. Note that the relative magnitudes of the estimated coefficients have now “switched” due to the WLS regression; 2400.1 in the WLS regression is comparable to 1273.2, not 0.72, in the OLS regression.

(d) It does not suggest anything should necessarily be done about it. Interestingly, there is another potential cause of heteroskedasticity in this model. That is, the variables are means of ranges of differing widths and different sample sizes. Thus it would seem reasonable to think that each of these means might come from a distribution with a different variance, and some sort of heteroskedasticity would be fairly likely.

10-7. \( R^2 = .122, N = 33, \) so \( NR^2 = 4.026 < 21.7 = \) the critical 1% Chi-square value with 9 d. f; so we cannot reject the null hypothesis of homoskedasticity. Thus both tests show evidence of heteroskedasticity.
10-9. (a) Multicollinearity and heteroskedasticity (but not positive serial correlation) appear to exist. We’d tackle the multicollinearity first. Since the heteroskedasticity could be impure, you should get the best specification you can before worrying about correcting for heteroskedasticity.

(b) For all intents and purposes, the two equations are identical. Given that, and given the reasonably strong t-score of STU, we’d stick with Equation 10.29. Note that the ratio of the FAC/STU coefficients is far more than 10/1 in Equation 10.29. This means that Equation 10.29 overemphasizes the importance of faculty, compared to Equation 10.30. (On second thought, what’s wrong with overemphasizing the importance of faculty?)

(c) Both methods show evidence of heteroskedasticity. For instance, if TOT = Z, the Park test \( t = 4.85 > 2.67 \), the critical t-value for a two-sided, 1 percent test with 57 degrees of freedom (interpolating).

(d) We don’t provide a t-score for the coefficient of \((1/TOT)\) because that coefficient is an estimate of the intercept of the original equation. (Note, however, that we do provide a t-score for the “constant” of the WLS equation.) This is a good example of “coefficient switching.” The pathetic fit is mainly due to the fact that the dependent variable in the WLS equation \((VOL/TOT)\) is significantly more difficult to predict than is VOL, where size alone can explain most of the variation. In addition, it turns out that not one of the WLS variables in this equation has any theoretical validity at all, making a bad situation even worse.

(e) There are many possible answers to this question. Using EViews, the easiest might be to reformulate the equation, using SAT and STU/FAC (the student/faculty ratio) as proxies for quality:

\[
\frac{VOL}{TOT_i} = 0.067 + 0.00011SAT_i - 0.0045STU_i/FAC_i \\
(0.00007) \quad (0.0013)
\]

\[ t = 1.59 \quad -3.44 \]

\[ R^2 = .19 \quad N = 60 \quad DW = 2.11 \]

10-11. (a) Coefficient: \( \beta_W \) \( \beta_U \) \( \beta_{hp} \)

Hypoth. sign: + – –?

t-value: 10.0 -1.0 -3.0

t_\text{c} = 1.66 reject do not reject

(5% one-sided reject with 90 d.f., — interpolating)

(b) We disagree, mainly because econometrics cannot “prove” that discrimination is the cause of any differences in employment. Less importantly, we disagree that the nondiscriminatory expected value of the coefficient of W is 1.00; for starters, a constant term and other variables are in the equation.

(c) Heteroskedasticity seems reasonably unlikely, despite the cross-sectional nature of the dataset, because the dependent variable is stated in per capita terms.

(d) The two-sided 1% critical t-value is approximately 2.64 (interpolating), so we cannot reject the null hypothesis of homoskedasticity.

(e) The theory behind P or lnP seems quite weak (despite its significance). Our preference would be to change P to a non-aggregate measure, for instance the percentage of the population that is black in the ith city, or some measure of the unemployment benefits available in the ith city.
10-13. (a) Coefficient:

<table>
<thead>
<tr>
<th>Hypoth. sign:</th>
<th>( \beta_p )</th>
<th>( \beta_I )</th>
<th>( \beta_Q )</th>
<th>( \beta_A )</th>
<th>( \beta_S )</th>
<th>( \beta_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-value:</td>
<td>-0.97</td>
<td>6.43</td>
<td>3.62</td>
<td>1.93</td>
<td>1.6</td>
<td>-2.85</td>
</tr>
<tr>
<td>( t_c = 1.684 )</td>
<td>do not reject</td>
<td>reject</td>
<td>reject</td>
<td>reject</td>
<td>do not</td>
<td>do not</td>
</tr>
</tbody>
</table>

(5% one-sided with 40 d.f., closest to 43)

The last two variables cause some difficulties for most students when hypothesizing signs. Our opinion is that having more suburban newspapers should hurt metropolitan newspaper circulation but that the number of television stations is a measure more of the size of a city than of the competition a newspaper faces. By the way, we see Q as a proxy for quality and A as an endogenous variable (note that the authors did indeed estimate the equation with 2SLS, a technique beyond the scope of this chapter).

(b) Heteroskedasticity seems extremely likely, since larger cities will have larger newspaper circulation, leading to larger error term variances. Using a two-sided 1% test, we can reject the null hypothesis of homoskedasticity since 3.13 > 2.66, the critical t-value with 60 degrees of freedom (closest to 48).

(c) Heteroskedasticity, multicollinearity, and omitted variables all seem likely.

(d) While it’s tempting to divide the equation through by population (or reformulate the equation by making the dependent variable per capita circulation), this would dramatically lessen the equation’s usefulness. Instead, we would attempt to improve the specification. Reasonable answers would include attempting to reduce some of the multicollinearity (redundancy) among the independent variables, trying to find a better measure of quality than the number of editorial personnel or substituting the number of major metropolitan newspaper competitors for S and T.

10-15. (a) To test for serial correlation, first run:

\[
\hat{S}_t = 0.73 + 0.22I_t + 0.46\ln(1 + V_t)
\]

\[
(0.05) \quad (0.16)
\]

\[
t = 4.50 \quad 2.85
\]

N = 58 (monthly) \[ \bar{R}^2 \cdot 5.56 \quad DW = 1.54 \]

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. This means that there is a chance that we transcribed the dataset incorrectly. If so, comparability with Lott and Ray is reduced, but the quality of the exercise for students is undiminished.

(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) A Park test with T as a proportionality factor produces a t-score of 0.45, providing no evidence whatsoever of heteroskedasticity. A White test produces an $R^2$ of .20, for an $NR^2$ of 11.6, which indicates heteroskedasticity at the 5% level but not at the 1% level we typically suggest for heteroskedasticity tests.

(e) A Park test with T as a proportionality factor produces a t-score of 0.99, once again providing no evidence of heteroskedasticity.

(f) Our preference would be to reformulate the equation (perhaps after testing for nonstationarity). We do not favor blindly using WLS in a time-series equation.

Chapter Twelve: Time-Series Models

12-3. (a) The double-log functional form doesn’t change the fact this is a dynamic model and that Y and M are almost surely related by a distributed lag.

(b) In their relationship to M, both Y and R have the same distributed lag pattern over time, since the lambda of 0.60 applies to both (in logs).

(c) Many students will run a Durbin-Watson $d$ test on this equation and conclude that there is no positive serial correlation because $DW = 1.80$, which is greater than 1.65, the 5% one-sided critical $d_a$. The best students won’t be so sure, however, because the Durbin-Watson $d$ is biased toward 2 in the presence of a lagged dependent variable and the question doesn’t provide enough information to run an LMSC. One student in a hundred will do more research, find out about Durbin’s $h$ test (which we no longer include in the text), and use the $h$ test to conclude that we cannot reject the null hypothesis of no serial correlation because $h = (1 - 0.5 \cdot 1.80)\sqrt{26/(1 - 26 \cdot 0.10^2)} = 0.59$, which is less that 1.96, the 5% critical $z$-value.

12-5. (a) \(SALES = -243 + 5.2AD_t + 1.9AD_{t-1} + 3.1AD_{t-2} + 1.0AD_{t-3} + 3.3AD_{t-4}\)

(b) \(SALES = -38.86 + 2.98AD_t + 0.79SALES_{t-1}\)

12-7. (a) Second-order serial correlation means that the current observation of the error term is a function of the previous two observations of the error term.

(b) \[e_t = a_0 + a_1X_t + a_2Y_{t-1} + a_3e_{t-1} + a_4e_{t-2} + u_t\]

There would be 2 degrees of freedom in the test because there are two restrictions in the null hypothesis ($a_3 = a_4 = 0$).

(c) \(\hat{e}_t = -11.8 - 0.22A_t + 0.04S_{t-1} - 0.06e_{t-1} - 0.25e_{t-2}\)

\[R^2 = 0.066 \quad N = 23 \text{ (1978–2000)}\]

\[LM = N \cdot R^2 = 23 \cdot 0.066 = 1.52 < 5.99, \text{ the 5% critical Chi-square value with 2 d.f., so we cannot reject the null hypothesis of no second-order serial correlation.}\]

12-9. We suggest that the farmers rethink either the form of their equation or their expectations. Their current equation posits that corn growth is a distributed lag function of rainfall, a not unreasonable idea. Unfortunately, lambda is restricted to between zero and one, so the likelihood of observing a negative lambda is small.

12-11. (a) \(2.01 > -3.22\), so we cannot reject the null hypothesis of nonstationarity for Y.

(b) \(-0.72 > -3.22\), so we cannot reject the null hypothesis of nonstationarity for r.

(c) \(4.16 > -3.22\), so we cannot reject the null hypothesis of nonstationarity for CO.

(d) \(0.04 > -3.22\), so we cannot reject the null hypothesis of nonstationarity for I.

All four variables appear to be nonstationary, at least at the 2.5% level. This is a bit surprising, because interest rate variables often are stationary, and it’s not a coincidence that the interest rate is the only variable to have a negative $t$-score.
Chapter Thirteen: Dummy Dependent Variable Techniques

13-3. (a) This equation is a linear probability model, and as such it can encounter at least two problems in addition to any “normal” specification problems it might have: $R^2$ is not an accurate measure of the overall fit, and $\hat{Y}_i$ is not bounded by 0 and 1.

(b) Some might question the importance of PV in the equation, and others might suggest adding variables to measure recent changes in profits or book value, but our inclination would be to switch to a logit before analyzing the specification much further.

(c) The best way to win this argument is to note that the linear probability model produces nonsensical forecasts (greater than 1.0 and less than 0.0).

13-5. It won’t take long for students to realize that with a logit, $\hat{D}_i$ is never greater than 1 or less than 0.

13-7. (a) It’s reasonable to conclude that if $D_i = 2$, then the logit computer program will balk at taking the log of a negative number (–2). As mentioned in the text, however, logit programs iterate using Equation 13.7 (or a version thereof), so it’s possible that a software package could produce some sort of estimates. (We’ve never seen one, however.)

(b) With a linear probability model, the answer is unambiguous. The estimated coefficients of WN and ME will double in size. If we divide each coefficient by two, the theoretical meanings will be unchanged from the original model.

13-9. In all three models, we find evidence that A is an irrelevant variable. (Note that in the datafile $D$ = “J”)

(a) $\hat{D}_i = -0.22 - 0.38M_i - 0.001A_i + 0.09S_i$
    $\quad (0.16) \quad (0.007) \quad (0.04)$
    $t = -2.43 \quad -0.14 \quad 2.42$
    $R^2 = .29 \quad N = 30 \quad R^2_p = .81$

(b) $\ln[D_i/(1-D_i)] = -5.27 - 2.61M_i - 0.01A_i + 0.67S_i$
    $\quad (1.20) \quad (0.04) \quad (0.32)$
    $t = -2.17 \quad -0.25 \quad 2.10$
    $\bar{R}^2 = .76$

(c) $\hat{Z}_i = F^{-1}(P_i) = -3.05 - 1.45M_i - 0.006A_i + 0.39S_i$
    $\quad (0.63) \quad (0.02) \quad (0.18)$
    $t = -2.31 \quad -0.26 \quad 2.20$
    $\bar{R}^2 = .76 \quad \text{iterations} = 5 \quad LR = 13.85$

13-11. (a) Three

(b) The three dependent variables are $\ln(P_u/P_c)$, $\ln(P_j/P_c)$, and $\ln(P_a/P_c)$, which are the log of the ratio of the probability of attending the choice in question divided by the probability of attending your college.
Chapter Fourteen: Simultaneous Equations

14-3. (a) If ε₂ decreases, Y₂ and then Y₁ decrease.
(b) If ε₀ increases, Q₀ increases, and then Qₛ increases (equilibrium condition) and Pᵣ increases.
(Remember that the variables are simultaneously determined, so it doesn’t matter which one is on the left-hand side.)
(c) If ε increases, CO increases, and then Y increases and YD increases.

14-5. All these cases can be shown to involve a positive correlation between the εs and the Ys.

14-7. (a) A: Predetermined = 2 < 3 = # of slope coefficients, so underidentified.
     B: Predetermined = 2 = 2 = # of slope coefficients, so exactly identified.
(b) Note that X₂ is endogenous to this system, so:
     Y₁: Predetermined = 3 < 4 = # of slope coefficients, so underidentified.
     Y₂: Predetermined = 3 > 1 = # of slope coefficients, so overidentified.
     X₂: Predetermined = 3 = 3 = # of slope coefficients, so exactly identified.
(c) Note that you can consider the change in Y to be endogenous to the system with a non-
     stochastic equation in which it equals Y t – Y t–1 . Given this, there are six predetermined
     variables, Y t–1 , E t, D t, M t, R t–1  and G t, so the identification properties of the four stochastic
     equations can be determined by using the order condition (which is necessary but not
     sufficient):
     CO t: Predetermined = 6 > 1 = # of slope coefficients, so overidentified.
     I t: Predetermined = 6 > 4 = # of slope coefficients, so overidentified.
     R t: Predetermined = 6 > 3 = # of slope coefficients, so overidentified.
     Y t: Predetermined = 6 > 3 = # of slope coefficients, so overidentified.

14-9. (a) \( \hat{I}_t = -270 + 0.19Y_t - 8.26r_{t-1} \)
     \[ (0.01) \quad (11.62) \]
     \( t = \quad 13.93 \quad -0.71 \)
     \( R^2 = .933 \quad N = 28DW = 0.38 \)
(b) \( \hat{Y}_t = 235.17 + 0.77G_t - 0.51NX_t + 1.51T_t + 0.67CO_{t-1} + 36.4r_{t-1} \)
     \[ (0.23) \quad (0.17) \quad (0.16) \quad (0.10) \quad (9.32) \]
     \( t = \quad 3.32 \quad -2.94 \quad 9.43 \quad 6.81 \quad 3.91 \)
     \( R^2 = .999 \quad N = 28 \)
(c) \( \hat{I}_t = -26.5 + 0.19 \hat{Y}_t - 8.48r_{t-1} \)
     (Standard errors obtained from this estimation are biased and should be disregarded.)
(d) \( \hat{I}_t = -265.5 + 0.19 \hat{Y}_t - 8.48r_{t-1} \)
     \[ (0.01) \quad (11.6) \]
     \( t = \quad 13.9 \quad -0.73 \)
     \( R^2 = .933 \quad N = 28 \quad DW = 0.38 \)
14-11. Most reasonable models of the labor market are simultaneous and therefore potentially involve simultaneity bias and should be estimated with 2SLS.

14-13. (a) OLS estimation will still encounter simultaneity bias because price and quantity are simultaneously determined. Not all endogenous variables will appear on the left-hand side of a structural equation.

(b) The direction of the bias depends on the correlation between the error term and the right-hand-side endogenous variable. If the correlation between the error term and price is positive, as it most likely is, then the simultaneity bias will also be positive.

(c) Three: stage one: \( P = f (YD, W) \)
stage two: \( Q_d = f (\hat{P}, YD) \) and \( Q_s = f (\hat{P}, W) \)

(d) OLS: \[
\hat{Q}_D = 57.3 - 0.86P + 1.03YD \\
\hat{Q}_S = 167.5 + 3.95P - 1.42W
\]

2SLS: \[
\hat{Q}_D = 95.1 - 6.11 \hat{P} + 2.71YD \\
\hat{Q}_S = 480.2 + 13.5 \hat{P} - 5.50W
\]

14-15. (a) All three variables are nonstationary. In Exercises 12-10 and 12-11, we showed that both \( CO_t \) and \( Y_t \) are nonstationary. If \( CO_t \) is nonstationary, then so too must be \( CO_{t-1} \). Since \( YD_t \) and \( Y_t \) are highly correlated, it’s reasonable to think that if one is nonstationary then so too is the other. As a test of this, a Dickey-Fuller test on \( YD_t \) produces a \( t \)-score of \( |-0.02| < 3.22 \), the critical value, so we cannot reject the null hypothesis. This is further evidence that \( YD_t \) is nonstationary.

(b) If we run a Dickey-Fuller test on the residuals, we get \( |2.8| > 2.048 \), a 2.5% 1-sided 28 df critical \( t \) value (which we use as an approximation—see the text). Thus we can reject the null hypothesis of nonstationarity and conclude that the residuals are stationary. This implies that Equation 14.31 is reasonably cointegrated.

(c) A lag model is more likely to be cointegrated dynamic model distributed because the lagged values of the dependent variable that appear on the right hand side of the equation should have the same degree of nonstationarity as the dependent variable.

(d) We agree with the majority of applied econometricians who think that the concept of cointegration is unrelated to the estimation technique. As a result, we do not hesitate to recommend the use of the Dickey-Fuller test when testing 2SLS residuals for cointegration. There are those who disagree, however, by pointing out that nonstationarity in a truly simultaneous system implies that a test for cointegration should go beyond testing the residuals of just one of the equations in the system.

Chapter Fifteen: Forecasting

15-3. Answers to this question will depend on the forecasted \( X \) values chosen.

15-7. If the answers to Exercise 15-6 were calculated correctly, then calculating “backwards” will indeed reproduce the original series.

15-9. (a) \( e_{99} = Y_{99} - \hat{Y}_{99} = 27 - 27.5 = - 0.5 \)

(b) \( Y_{100} = 0 + 1(27) - 0.5(- 0.5) = 27.25 \)
\( Y_{100} = 0 + 1(27.25) = 27.25 \)
\( Y_{102} = 0 + 1(27.25) = 27.25 \)
Chapter Sixteen: Statistical Principles

16-3. The mean is 16.89 and the standard deviation is 6.43. Thus the 1999 P/E ratio was more than two
standard deviations above the mean.

16-5. Standardized scores: 1.9, 0.0, and –0.8, raw score: 90.

16-7. \[ Z = \frac{(70 - 66)}{2.5} = 1.60. \] \[ P[Z > 1.60] = 0.0548. \]

16-9. People who have visited France for pleasure more than once during the past two years are more
likely to have had good experiences than are people who visited France just once and then never
returned and/or people making their first visit to France.

16-11. The standard deviation is the square root of the variance, or 0.686, and the 95\% two-sided \( t \) with
34 degrees of freedom is approximately 2.03, so a 95\% confidence interval is 6.19 +
2.03 \( \left( \frac{0.686}{\sqrt{35}} \right) \) or 6.19 + 0.24. There is a 95\% probability that a confidence interval
constructed in this fashion will include the true value of the mean prediction of the population, so:
(a) No. This says nothing about how accurate or inaccurate the forecasters are.
(b) No. If anything, we might estimate that approximately 95\% of the forecasts are in an interval
equal to our estimate of the mean plus or minus 2 standard deviations of the individual
observations:

\[ 6.19 + 1.96 \left( \frac{0.686}{\sqrt{35}} \right) \quad \text{or} \quad 6.19 + 1.34. \]

16-13. The calculated \( t \)-value = \( \frac{(13.333 - 11.2)}{\sqrt{5.499}} \) = 1.344. Since a two-sided 5\% critical \( t \)-
value is 2.201, we cannot reject the null hypothesis.

16-15. All three statements are wrong:
(a) This statement confuses \( P[A \text{ if } B] \) with \( P[B \text{ if } A] \).
(b) Statistical significance is not the same as practical importance. We cannot tell the estimated
size of the effect from the P value.
(c) Not being able to reject the null hypothesis is not the same as proving the null hypothesis to
be true.

The correct interpretation is that the test shows that if the cold vaccine is ineffective, there is a
0.08 probability of observing a difference this large (or larger) between the fractions of each
group that caught colds.
Chapter One: An Overview of Regression Analysis

1-1. What is Econometrics?
Here’s the perfect opportunity for an inspirational talk about how econometrics is used today. You could mention anything from the high percentage of academic articles and government studies that involve econometrics in one way or another to how knowledge of econometrics will help students do well in business school and (hopefully) business.

1-2. What is Regression Analysis?
Many instructors feel that students learn more if an example is introduced early on and then the “nitty gritty” details of this section are applied to the same example as the class progresses. One possibility for such an example is a simple model of something that all the students are familiar with, for instance a model of the number of students (N) that sign up for a class as a function of how easy a grader the professor of the class is (measured by the professor’s average GPA) and the number of majors in the department (NM):

\[ N_i = f(GPA_i, NM_i) + \epsilon_i = \beta_0 + \beta_1 GPA_i + \beta_2 NM_i + \epsilon_i \]

or a model of the number of tickets sold for a football game (T) as a function of the rank (1 = first place) of the home team in the league standings (P) and the number of miles away the visiting team’s campus is (M):

\[ T_i = f(P_i, M_i) + \epsilon_i = \beta_0 + \beta_1 P_i + \beta_2 M_i + \epsilon_i \]

By the way, you’re doing well if only a couple students expect a positive sign for the coefficient of P!

1-3. The Estimated Regression Equation
Then in this section you could continue the example by providing an estimated equation, complete with numbers and “hats.” A real key here is to focus on the difference between the error term and the residual.

1-4. A Simple Example of Regression Analysis
This weight/height example is worth taking the time to go into in detail; we refer back to it a number of times later in the text. One way of improving the example is to ask your male students to take a moment on the first day of class to write down their heights and weights for you. When you cover this example on the second or third day, you could announce the student-estimated coefficients and compare them to the textbook’s, thus bringing the example to life and also planting the seeds of the idea that estimated regression coefficients come from a sampling distribution of beta-hats.

1-5. Using Regression to Explain Housing Prices
The estimates in this section represent the simplest possible model that can be estimated from the data presented in Section 11.7, the Housing Price Interactive Exercise. Since that section discusses the data and the theory behind hedonic models more than Section 1.5 does, you might consider looking at Section 11.7 before you present Section 1.5 in class.
1-6. Summary and Exercises

For a problem set on this chapter, we’d recommend Exercises 1-4, 1-11, and 1-12.

**Chapter Two: Ordinary Least Squares**

2-1. Estimating Single-Independent-Variable Models with OLS

Students often get bogged down in Section 2.1.2, and we urge you to require the students to work through an estimation example (like Section 2.1.3 or Exercise 2-4) on their own. No matter how hard we try, we still encounter students who have difficulty understanding what regression estimation actually is until they’ve done it for themselves.

If you want an easy estimation problem for which the coefficient estimates are not in the text, use Ando and Modigliani’s data in Exercise 10-5. The coefficient estimates for this dataset are:

\[
\hat{CO}_i = 1273.2 + 0.72I_i \quad R^2 = .97
\]

\[
(0.044) \quad t = 16.21
\]

where: CO = average consumption

I = average income.

2-2. Estimating Multivariate Regression Models with OLS

The key here is understanding the meaning of multivariate regression coefficients. In our experience, there are two levels of such understanding. At the first level, the student can recall that all the other independent variables in the equation are held constant only if asked directly. At the second level, the student understands the fact that the other regressors are held constant well enough to use and interpret the estimated coefficients correctly in an example. We think it’s vital to cover this concept in enough detail and from enough different points of view that all students can reach the second level by the end of the chapter.

2-3. Evaluating the Quality of a Regression Equation

This short section lists a number of items that should be considered when evaluating the quality of a regression equation. The most important concept to communicate is that fit is just one of many attributes of an equation that should be considered.

2-4. Describing the Overall Fit of the Estimated Model

One of the problems with introducing R^2 and then immediately replacing it with R^2 is that some students ask why we bothered mentioning it at all. Our answer is that it’s important to know R^2 in its own right and that understanding R^2 makes the advantages of R^2 more obvious.

Unfortunately, if you do a good enough job of presenting the benefits of R^2 then some students will become curve-fitters for the rest of the semester. It’s worthwhile to suggest to these “true believers” that they reread (aloud with a friend, if they dare) the imaginary conversation on page 58 that points out the problems with relying on R^2 as the primary measure of the quality of a regression.
2-5. An Example of the Misuse of $R^2$

We do not return to this water-demand example, and most students can understand the point of the example without it being discussed, so it can be skipped in class if necessary. Note, however, that this example and the interactive regression learning example at the back of this instructor’s manual are based on the same research.

2-6. Summary and Exercises

For a problem set on this chapter, we’d recommend Exercises 2-4, 2-7, 2-8 and 2-11.

Chapter Three: Learning to Use Regression Analysis

We’ve heard from some users of the text that they cover Chapter 3 after Chapter 7 or (less frequently) after Chapter 10, and we have no objections to this practice. To help these instructors, we’ve made slight revisions in order to allow Chapter 3 to be skipped without much loss of continuity. The two topics in Chapter 3 that are needed in future chapters are some coverage of dummy variables (pages 69–70) and some exposure to the Woody’s dataset.

3-1. Steps in Applied Regression Analysis

This section is best taught around an example. Indeed, if you’re going to cover the Woody’s example in detail and you’re short on time, Sections 3.1 and 3.2 can be presented simultaneously in class. If you have the time, however, it’d be better to present the six steps with a distinct example (from your own research, for example, perhaps a bit simplified), before going on to the Woody’s example.

3-2. Using Regression Analysis to Pick Restaurant Locations

This Woody’s restaurant location example is one of the two most important examples in the book. We return to it in virtually every chapter, so the time you spend going over the example in class will pay off in enhanced student understanding of a number of different topics. In this chapter, the example puts the six steps in applied regression analysis to work, acts as an illustration of estimation for those who are still stuck on Section 2.1.2, and is a first introduction to what the computer output of a packaged regression program looks like. A good project assignment is to ask the students to estimate this regression themselves on your computer system. If you grade this assignment pass/fail and allow students to work together, it will help students with computer phobias (or who still don’t quite know what a regression is) to work with other students and make some progress in an unthreatening environment.

3-3. Summary and Exercises

For a problem set on this chapter, we’d recommend Exercises 3-3, 3-5, 3-7, and 3-10.

Chapter Four: The Classical Model

4-1. The Classical Assumptions

The students won’t want to hear you say it, but the classical assumptions simply must be known cold by econometrics students. There is little value to a week spent studying serial correlation or heteroskedasticity if the underlying assumption that’s being violated isn’t understood. As a result, any trick or threat you can think of to get the students to learn the classical assumptions is probably fair to the students in the long run. It’s sad but true that the most effective way we know to get even the weakest students to learn the assumptions is to tell them that the assumptions will be on the next midterm (and then follow through).
4-2. The Sampling Distribution of $\hat{\beta}$

A very effective demonstration technique to bring this concept home quite dramatically is to give each student his or her own dataset drawn from a larger population and to ask each student to calculate a $\beta$-hat from that sample. If you then collect the calculated $\hat{\beta}s$ and graph them, it produces a fairly normal-looking distribution of estimates. Such an in-class example requires you to develop a different small subsample for each student and also requires each student to run a small regression, but we find that the results are worth the effort.

One way to cut down on the instructor workload of the previous suggestion is to ask each student to run a weight-height regression on a dataset (males only) that they collect on their own. To keep the assignment manageable, of course, the sample size should be fairly low and some sort of computer facilities should be available. Such an exercise can serve as a student’s first introduction to the computer if the Woody’s example of the previous chapter was not assigned to be estimated individually.

4-3. The Gauss-Markov Theorem and the Properties of OLS Estimators

Needless to say, this short section is quite important. While we don’t recommend that you require it, note that a proof of the Gauss-Markov Theorem is in the answer to Exercise 4-8 in Appendix A.

4-4. Standard Econometric Notation

Some students who are unfamiliar with the expectation function will have trouble with Section 4.4. These students often have an easier time if they think of an expected value as a mean. Some instructors may want to save time by presenting Sections 4.1 and 4.4 at the same time.

4-5. Summary and Exercises

For a problem set on this chapter, we’d recommend Exercises 4-4, 4-5, and 4-9.

Chapter Five: Hypothesis Testing

For some groups of students, you may want to cover Chapter 16 before starting Chapter 5. For the majority, however, we find that presenting basic statistical concepts as students need them is much more effective for student learning than is a complete review of statistical inference, but we recognize that not everyone agrees on this issue.

5-1. What is Hypothesis Testing?

To give students exposure to more than one example of hypothesis testing, it’s often worth the effort to use a different example in class than we used in the text. For instance, consider a supply equation for cotton:

$$C_t = \beta_0 + \beta_1 A_H + \beta_2 P_{C_t-1} + \beta_3 P_{S_t-1} + \beta_4 M_t + \epsilon_t$$

where:

- $C_t$ = thousands of bales of cotton produced in year $t$
- $A_H$ = thousands of acres of cotton harvested in year $t$
- $P_{C_t-1}$ = the real price of a thousand pounds of cotton in year $t - 1$
- $P_{S_t-1}$ = the real price of a thousand bushels of soybeans in year $t - 1$
- $M_t$ = price index for farm machinery in year $t$
This is a typical student effort at such an equation. Acres harvested is not tautologically related to production (as acres planted might be if technology was constant during the sample), but it is likely to be jointly determined with production. Still, a positive sign would be expected. The price of cotton is lagged one time period because of the length of time it takes to switch crops; a positive sign is expected. Soybeans can be planted on the same kinds of land on which cotton is planted, and so a negative sign is expected. Finally, the price of cotton machinery (negative expected sign) was unavailable, so the student instead chose the average price of all farm machinery, leading to an ambiguous expected sign (since cotton is fairly labor intensive, an increase in price in farm machinery might shift farmers into cotton production).

Based on this theory, the appropriate null and alternative hypotheses would be:

\[
\begin{align*}
\beta_1: & \quad H_0: \beta_1 = 0, \quad H_A: \beta_1 > 0 \\
\beta_2: & \quad H_0: \beta_2 = 0, \quad H_A: \beta_2 > 0 \\
\beta_3: & \quad H_0: \beta_3 = 0, \quad H_A: \beta_3 < 0 \\
\beta_4: & \quad H_0: \beta_4 = 0, \quad H_A: \beta_4 \neq 0
\end{align*}
\]

5-2. The \( t \)-Test

Continuing on with the cotton example, the appropriate border value for all four null hypotheses is zero, so the formula for the \( t \)-statistic to use is Equation 5.3:

\[ t_k = \frac{\hat{\beta}_k}{\text{SE}(\hat{\beta}_k)} \]

The equation will be estimated for annual U.S. data from 1947 through 1972, so there will be 21 degrees of freedom. Given this, and given a 5 percent level of significance, the appropriate decision rules and critical \( t \)-values (from Statistical Table B.1) are:

- \( \beta_1 \): Reject \( H_0 \) if \( |t_1| > 1.721 \) and if \( t_1 > 0 \)
- \( \beta_2 \): Reject \( H_0 \) if \( |t_2| > 1.721 \) and if \( t_2 > 0 \)
- \( \beta_3 \): Reject \( H_0 \) if \( |t_3| < 1.721 \) and if \( t_3 < 0 \)
- \( \beta_4 \): Reject \( H_0 \) if \( |t_4| > 2.080 \)

5-3. Examples of \( t \)-Tests

Continuing on with the cotton example, the estimated equation is:

\[
\hat{C}_t = 624 + 0.33AH_t + 1929PC_{t-1} - 154PS_{t-1} + 12.4M_t
\]

\[
\begin{align*}
(0.10) & & (1078) & & (82) & & (7.4) \\
t & = & 3.16 & & 1.79 & & -1.89 & & 1.67
\end{align*}
\]

(these \( t \)-scores are computer-calculated and differ from those calculated by hand because of rounding.)

\( t \)-tests of the first three coefficients are all one-sided tests, and in all three cases we can reject the null hypothesis. This is because the calculated \( t \)-score is greater in absolute value than the critical \( t \)-value and has the same sign as the alternative hypothesis for all three.

A \( t \)-test on the coefficient of \( M \) is appropriately a two-sided test, however. We cannot reject the null hypothesis of no effect for this variable because the calculated \( t \)-score, 1.67, is not greater than the critical \( t \)-value of 2.080.
5-4. Limitations of the $t$-Test

This section sets the stage for our discussion in Chapter 6 of the problems associated with using the $t$-test as a method for selecting independent variables. While the section is self-explanatory, it also can be applied to the cotton production equation if desired. For example, the high $t$-score for AH does not imply theoretical validity or importance. Indeed, it is one of the two worst variables in the equation. By the way, we’ve not yet covered multicollinearity, but there is a fairly high simple correlation coefficient in this dataset. The $r$ between $AH_i$ and $PC_{i-1}$ is 0.792.

5-5. Summary and Exercises

For a problem set on this chapter, we’d recommend Exercises 5-2, 5-5, 5-6a, 5-9, and 5-13.

5-6. Appendix: The $F$-Test

Sections 5.6.1 and 5.6.2 are fairly simple, since they involve little more than the comparison of an $F$-score (printed out by the computer) with a critical $F$-value in the back of the book. We find that it enhances student understanding of statistical hypothesis testing and of the $t$-test to distinguish between tests of hypotheses about one coefficient and tests of multiple hypotheses.

The textbook does not include any exercises for the appendices, so this instructor’s manual will include a problem that can be used as an in-class teaching example (or as an exercise on a problem set or as a question on an exam) for each appendix in the text.

For Appendix 5.5, a possible example/exercise/exam question would be:

Suppose you want to test the hypothesis that the first 20 restaurants in the Woody’s example are different in some way from the rest of the sample. Given the following two regression results, use the $F$-test to test the null hypothesis that the slope coefficients are the same in the two samples (also called a Chow test):

That is, given:

$$A: \hat{Y}_i = 89,977 - 9693N_i + 0.431P_i + 1.67I_i$$

$$t = -3.10 \quad 3.52 \quad 2.18$$

$$N = 20 \text{ (observations 1-20)} \quad R^2 = .554$$

$$RSS_A = 3.890E + 09 \text{ (RSS for this sample)}$$

$$B: \hat{Y}_i = 119,761 - 9460N_i + 0.324P_i + 0.68I_i$$

$$t = -3.39 \quad 3.56 \quad 0.86$$

$$N = 13 \text{ (observations 21-33)} \quad R^2 = .596$$

$$RSS_B = 1.68E + 09 \text{ (RSS for this sample)}$$

Test the hypotheses:

$$H_0: \beta_{NA} = \beta_{NB}; \beta_{PA} = \beta_{PB}; \beta_{IA} = \beta_{IB}$$

$$H_A: H_0 \text{ is not true.}$$

at the 5 percent level of significance.
**Answer:** The appropriate $F$-statistic to use is:

$$F = \frac{(\text{RSS}_T - \text{RSS}_A - \text{RSS}_B)/(K+1)}/\left(\frac{\text{RSS}_A + \text{RSS}_B}{(N+M-2(K+1))}\right)$$

where: $\text{RSS}_T =$ the residual sum of squares for the entire sample (6.13E+09 as in Table 3.2 on page 81 of the text)

$N =$ observations in sample A (20)

$M =$ observations in sample B (13)

$K =$ number of explanatory variables (3)

and $F$ has $K+1$ numerator and $N+M-2(K+1)$ denominator degrees of freedom. The decision rule to use is to reject the null hypothesis if $F$ exceeds the critical $F$-value of 2.76 found in Statistical Table B.2 for 4 and 25 degrees of freedom:

$$\text{Reject } H_0 \text{ if } F > 2.76$$

As it turns out, substitution into the $F$ formula produces:

$$F = \frac{(0.56E+09)/(4)}{(5.57E+09)/(25)} = (0.14E+09)/(0.223E+09) = 0.628$$

Since $F$ is less than 2.76, we cannot reject the null hypothesis. That is, the coefficients are not significantly different at the 5 percent level of significance.

**Chapter Six: Specification: Choosing the Independent Variables**

6-1. Omitted Variables

The demand for chicken is one of the two most important examples in the text and is referred to in almost all the remaining chapters. As a result, we recommend that you take the time to go over the example while discussing omitted and irrelevant variables in class.

Equation 6.7 is sometimes difficult for students to understand from just the examples in the book, so another example (or exam question) might be helpful:

In the Woody’s equation, what would be the expected bias on the coefficient of the competition variable if we were to drop the population variable from the equation?

Well, Equation 6.7 becomes:

$$\text{bias} = \beta_p \cdot f(r_{\text{P},p})$$

The expected sign for the population coefficient is positive, and you’d expect the simple correlation coefficient between population and competition to be positive (because the more densely populated an area is, the more restaurants you’d expect to be built . . . indeed, $r$ turns out to be +0.726), so:

$$\text{bias} = \beta_p \cdot f(r_{\text{P},p}) = + + + +$$

So we’d expect dropping P from the equation to cause positive bias in $\hat{\beta}_n$, and that’s what happens:

$$\hat{Y}_i = 84,439 - 148N_i + 2.32I_i$$

(1778) (0.66)

$$t = -0.84 \quad 3.50$$

$N = 33 \quad \bar{R}^2 = .258$

Note that the coefficient of N has increased from −9075 (in Equation 3.5) to −1487, which is the direction it would move if it had been positively biased by the omission of P.
6-2. Irrelevant Variables and

6-3. An Illustration of the Misuse of Specification Criteria

To continue the example above, if we had started with the Woody’s equation with only N and I as above, we might have concluded (mistakenly) that N is irrelevant, because without N, the equation looks like:

\[
\hat{Y}_i = 77,543 + 2.34I_i \quad \hat{R}^2 = .265
\]

\[
(0.66) \\
 t = 3.54
\]

That is, if we apply the four criteria for an irrelevant variable:

- **Theory** - Perhaps Woody’s patrons are so dedicated that they eat there no matter what else is available?
- **t-test** - The t-score is −0.84, insignificantly different from zero, indicating irrelevance.
- **\( R^2 \)** - \( R^2 \) rises when N is dropped, indicating that it is irrelevant.
- **Bias** - When N is dropped, the coefficient of I doesn’t change at all, an indication that N is irrelevant (or that P and N are uncorrelated).

Thus we’d mistakenly come to the conclusion that N was irrelevant when in fact the problem was being caused by the omitted variable P!

6-4. Specification Searches


6-5. An Example of Choosing Independent Variables

This example turns out to be an excellent way for students to learn the principles outlined in the previous sections. Most students’ first impulse is to include every variable and then drop the ones that have insignificant coefficients. As mentioned previously, this approach is almost sure to cause bias. By the way, this is a great place to ask students to collect their own data sets and compare estimated coefficients with each other and with the book. Finally, this dataset was collected at a college that avoids the use of multiple choice questions on tests.

6-6. Summary and Exercises

For a problem set on this chapter, we’d recommend Exercises 6-2, 6-3, 6-11, and 6-13.

6-7. Appendix: Additional Specification Criteria

Our choices of RESET, AIC and Schwarz for this section were not based on an opinion that they are the best or even the most-used additional specification criteria. Instead, we looked for criteria that are easy to use and that represent different approaches to specification choice. (For instance, AIC and SC are criteria pure and simple, but RESET can more appropriately be called a test.) Suggestions of criteria to add for the next edition would be appreciated.
At any rate, for Appendix 6.7, a possible exam Question/exercise/example would be:

For practice with our additional specification criteria, return to the Woody’s dataset, drop I from the equation, and attempt to detect this specification error using:

(a) Ramsey’s RESET.
(b) Akaike’s Information Criterion
(c) The Schwarz Criterion

Answer:

(a) Using the Y^4 version of RESET, we obtain: \[ F = \frac{[(6.13328 \times 10^9 - 4.92506 \times 10^9)/3]/[(4.92506 \times 10^9)/26]}{[(4.92506 \times 10^9)/26]} = 19.14 > 2.99 = F_{0.05} \] the closest in the table to a 5% critical value with 3 and 26 degrees of freedom. We thus reject the null hypothesis that the coefficients of the added terms are jointly equal zero and conclude that there is a specification error according to RESET.

(b) For \( Y = f(N, P, I) + \varepsilon \) \[ AIC = 22.12 \]
   For \( Y = f(N, P) + \varepsilon \) \[ AIC = 22.24 \]

   Thus the original specification (with I) has a lower AIC and is preferable according to that criterion.

(c) For \( Y = f(N, P, I) + \varepsilon \) \[ SC = 22.30 \]
   For \( Y = f(N, P) + \varepsilon \) \[ SC = 22.37 \]

   Thus the original specification (with I) has a lower SC and is preferable according to that criterion.

Chapter Seven: Specification: Choosing a Functional Form

7-1. The Use and Interpretation of the Constant Term

Some instructors may disagree with the “no exceptions” point of view we take in the text, and we think students at this level would probably benefit from hearing both sides of the debate. In particular, counter-examples, while not convincing to us, may make for an interesting class session.

7-2. Alternative Functional Forms

The translog function mentioned in footnote 6 combines three different functional forms, the double-log, the polynomial quadratic, and an interaction term, to come up with an equation uniquely suited to estimating various kinds of cost functions:

\[ C = \beta_0 + \beta_1 Q + \beta_2 Q^2 + \beta_3 B + \beta_4 B^2 + \beta_5 QB + \varepsilon \]

where: \( C \) = the log of total cost of production  
\( Q \) = the log of the level of output  
\( B \) = the log of the number of offices or plants

The translog cost function allows the production of the item to exhibit increasing or decreasing costs as production is increased, and it even lets different economies of scale exist over different ranges of output. While the translog function in general is too complex to warrant detailed consideration in class, it serves as an excellent example of the innovative ways in which various functional forms can be combined to meet the requirements of complex economic theories. For more, see Laurits R. Christensen and and William H. Greene, “Economies of Scale in U.S. Electrical Power Generation,” Journal of Political Economy, August 1976, pp. 655–676.
7-3. Lagged Independent Variables

A good first example of a lagged independent variable is that of a price variable in a supply equation for an agricultural good. For example, Section 5.1 of this Instructor’s Manual estimated the supply of cotton (C) as a function a number of variables, including the price of cotton lagged one year:

\[ \hat{C}_t = 624 + 0.33AH_t + 1929PC_{t-1} - 154PS_{t-1} + 12.4M_t \]

\[ (0.10) \quad (1078) \quad (82) \quad (7.4) \]

\[ t = 3.16 \quad 1.79 \quad -1.89 \quad 1.67 \]

7-4. Using Dummy Variables

The two key concepts here are that:
1. you must have one fewer dummy variable than you have conditions and
2. that if you have more than one dummy variable describing a situation, the coefficient of a particular dummy represents the impact of that condition on the dependent variable compared to the omitted condition (holding all non-related variables in the equation constant).

7-5. Slope Dummy Variables

One interesting extension of the use of slope dummies is piecewise regression (also called segmented regression), where the $D = 1$ if $X$ is greater than a specified value $X^*$ and $D = 0$ if not. A piecewise regression allows the slope of a linear relationship to change, as in the old kinked demand curve of oligopoly theory. If the equilibrium price is considered $X^*$, then a kinked demand curve for an individual oligopoly would predict a different slope at prices greater than $X^*$ than at lower prices because competitors will match price decreases (changing industry supply) but not price increases.

7-6. Summary and Exercises

For a problem set on this chapter, we’d recommend exercises 7.4, 7.7 (which purposely omits the sample size), 7.15, and 7.16.

**Chapter Eight: Multicollinearity**

8-1. Perfect vs. Imperfect Multicollinearity

Perfect multicollinearity is so unusual that it often can be covered in class in less than five minutes. A typical example of perfect multicollinearity that results from a student mistake is a variable that does not change at all during the sample period. In such a case, the variable is perfectly multicollinear with the constant term. Similarly, other students include dummy variables for all four seasons in an equation that also includes the constant term. They encounter perfect multicollinearity because the sum of the four dummy variables will always equal one.

8-2. The Consequences of Multicollinearity

Both examples are used again later in the text. The student consumption function is used to illustrate remedies for multicollinearity in Section 8.4, while the demand for gasoline example is used as an example of heteroskedasticity in Section 10.5.

8-3. The Detection of Multicollinearity

A refresher in the simple correlation coefficient may be in order here; at a minimum, make sure that the students have read Section 2.4.2.
8-4. Remedies for Multicollinearity
The last paragraph of Section 8.4.4 warns against combining cross-sectional and time series data sets to avoid multicollinearity. The basic reason we do not advocate such “pooled” datasets has to do with the dynamic or time-wise interpretation of the coefficients. Time series coefficients show the impact on the dependent variable of a one-unit change in the independent variable for a specified unit of time, such as a quarter or year (holding the other Xs constant). Cross-sectional coefficients measure the same impact over an indefinite time period and therefore can be interpreted as measuring long-run effects (where “long-run” is not precisely quantifiable in time-series units). Thus mixing the two makes the interpretation of the coefficients quite difficult in situations where the coefficients are likely to change depending on whether they’re long-run or short-run. For more on this, see Edwin Kuh and John R. Meyer, “How Extraneous are Extraneous Estimates?” Review of Economics and Statistics, Nov. 1957, pp. 380–393.

Also note that such pooled data sets have potential inherent heteroskedasticity (if the variances of the error terms of the different pooled samples are different, as would seem likely). Finally, pooling makes it difficult to test or adjust for serial correlation without doing so for each time series before pooling. This is because the first observation of one time series will follow the last observation of the previous time series. In such a case, there usually is no meaning to a model which posits that one observation of the error term is a function of the previous one.

8-5. Choosing the Proper Remedy
The fish/Pope study in Section 8.5.2 is a good one for class use because students seem to enjoy it and because alternative specifications can not only be discussed but also estimated, since the data are given.

8-6. Summary and Exercises
For a problem set on this chapter, we’d recommend exercises 8-6, 8-9, 8-11, and 8-12.

8-7. Appendix: The SAT Interactive Regression Learning Exercise
The fact that we’ve devoted almost 40 pages of this text to interactive regression learning exercises lets you know how potentially valuable we feel they can be, and we urge you to take advantage of them. One way to use the SAT interactive exercise of this section is to discuss it in class (after assigning the reading up through but not including Section 8.7.2). We try to let the class discuss the issues and then vote on which specification to start with, which regression run to move to, etc. We find it’s best to let the students make their choices and then give them feedback after their decisions, but others choose to give feedback as they go along. Such an in-class exercise will promote active learning and also prepare the students to do the interactive exercise in Chapter 11 on their own.

Chapter Nine: Serial Correlation

9-1. Pure vs. Impure Serial Correlation
Instructors who choose to spend only a short amount of time on this section will suffer no loss of continuity as long as students understand that serial correlation can be “caused” by a specification error.
9-2. The Consequences of Serial Correlation

This topic is one of the hardest in elementary econometrics for beginning students to understand, and instructors are urged to pay special attention to it.

9-3. The Durbin-Watson $d$ Test

One way to liven this section up is to ask one student to randomly choose a number between zero and four (the DW), then ask another student to pick a sample size and a number of explanatory variables and finally ask the rest of the students to decide what these numbers imply about serial correlation.

9-4. Remedies for Serial Correlation

Section 9.4.2 is certainly worth covering even if you encourage the use of GLS.

9-5. Summary and Exercises

For a problem set on this chapter, we’d recommend exercises 9-2, 9-9, 9-11, and 9-13.

Chapter Ten: Heteroskedasticity

10-1. Pure vs. Impure Heteroskedasticity

10-2. The Consequences of Heteroskedasticity

See our comments on Sections 9.1 and 9.2.

10-3. Testing for Heteroskedasticity

We know that our decision to focus on the Park test will cause some difficulty for instructors who typically use another test, and our inclusion of the White test is meant to offset the problem. In addition, the rest of the text is worded flexibly enough to accommodate whichever test the instructor decides to use.

While we understand the theoretical problems with the Park test, we find that its conclusions are typically the same as more complex tests. More importantly, beginning students seem to intuitively understand the Park test more readily and find it easier to calculate than other tests.

10-4. Remedies for Heteroskedasticity

The coefficient-switching that goes on when WLS is used (described at the top of page 365) sometimes is difficult for students to understand right off the bat. In such a situation, you might consider pulling the answer to Exercise 10-5 out of the first section of this instructor’s manual and using the exercise (especially part c) as an in-class example. That is, students who have trouble understanding Equation 10.16 sometimes see things more clearly when there is only one independent variable and when coefficient estimates are included (so that the similarity of magnitudes can be seen).

10-5. A More Complete Example

This reprise of the demand for gasoline example from Section 8.2.2 is an extremely useful in-class exercise, but it can be left to the students’ own reading without any loss of continuity.

10-6. Summary and Exercises

For a problem set on this chapter, we’d recommend exercises 10-2, 10-4, 10-5, 10-13 and 10-7 (only if you want students to have experience with the White test as well as with the Park test).
Chapter Eleven: A Regression User’s Handbook

11-1. A Regression User’s Checklist and Guide
The checklist is mainly a reference for student use and doesn’t deserve much class time. The guide, however, is a useful summary of the previous five chapters and serves as a good review for an examination or as a good preparation for doing the interactive exercise. We often find that a class devoted to reviewing the items in Table 11.2 synthesizes student knowledge that otherwise was fragmented by chapter and topic.

11-2. Running Your Own Regression Project
You may want to assign this section well before you get to Chapter 11, as it lays out the steps necessary for a student to complete his or her own independent project.

Once the students have completed this section, they should be ready to run their own regressions, following the six steps of applied regression first outlined in Chapter 3.

It’s usually worth your time to require that students get their topics approved by you before they begin any estimations. Topics should not be approved unless the student has found an adequate data set (including all the independent variables in the first regression). This approval requirement will help students avoid projects with inadequate data (like a study of the black market value of the Vietnamese piaster during the Vietnam war) or which are fairly difficult without more advanced techniques (like an aggregate investment function).

11-3. Economic Data
The internet is becoming the best source of data for students. A good place is start is the textbook’s website: <www.aw-bc.com/studenmund>. The website contains links to good data sources, and we’ll continually update those links as the sources (or their addresses) change.

11-4. The Ethical Econometrician
Before assigning a project or requiring that an interactive regression learning exercise be done independently, it’s a good idea to review the problems associated with sequential specification searches. Otherwise, the students may draw the wrong conclusion from the lists of many different specifications back to back.

11-5. Practical Advice for Applied Econometricians
We’re delighted that Peter Kennedy coauthored this section and allowed us to include in it some material from his Guide to Econometric. It should be extremely useful for all applied econometricians.
11-7. Appendix: The Housing Price Interactive Exercise

We urge you to require this interactive exercise as part (or all) of a problem set. Once a student has finished this assignment, he or she should be ready to tackle an individually chosen research project without too much help from you. As mentioned in the Hints in Appendix A, our favorite model on theoretical grounds is:

\[ P = f(S, N, A, A^2, Y, CA) + \varepsilon \]

The results of the estimation of this particular specification are:

\[ \hat{P} = 182 + 0.09S - 29.9N - 1.86A + 0.0156A^2 + 0.005Y + 0.89CA \]

\[ (0.01) \ (5.0) \ (0.94) \ (0.0097) \ (0.001) \ (11.7) \]

\[ t = 7.14 \ -5.98 -1.98 \ 1.60 \ 3.58 \ 0.08 \]

\[ R^2 = 0.90 \ \ N = 43 \ \ DW = 1.76 \]

We admit that CA’s coefficient is absurdly insignificant, but we still think the variable belongs in the equation.

Chapter Twelve: Time-Series Models

12-1. Dynamic Models

While most students can identify Equation 12.3 as a dynamic model (or distributed lag equation) when discussing the chapter, they will completely miss the boat when analyzing a similar equation outside the context of the chapter. We find that only repeated practice will help solve this problem.

12-2. Serial Correlation and Dynamic Models

As we hint at in this section, some econometricians disagree with the majority view that serial correlation is more likely in dynamic models. Their position is that omitted variables introduce bias whenever they are correlated with included variables and that lagged errors therefore introduce bias here but not in other cases.

12-3. Granger Causality

The purpose of this section is to present the idea, not to cover it fully. You will lose little by skipping this section.

12-4. Spurious Correlation and Nonstationarity

This section attempts to walk the narrow line between being understandable by a college sophomore and doing justice to the material in a fairly short space. We’d love to get your feedback on this section, particularly on the usefulness of the boxes material on page 441.

12-5. Summary and Exercises

For a problem set on this chapter, we’d recommend exercises 12-2, 12-5, 12-6, 12-9, and 12-11.
Chapter Thirteen: Dummy Dependent Variable Techniques

13-1. The Linear Probability Model
We use the women’s labor force example throughout the chapter and even have an exercise on it at the end of the chapter, so it’s probably worth mentioning in class.

13-2. The Binomial Logit Model
This section is the heart of the entire chapter, and it’s worth quite a bit of time. Students at first will balk at equations like 13.7, but if you work through enough examples, they’ll come to realize that a logit isn’t much harder than OLS. We think that going through the example in Section 13.2.3 in class will help bring this realization about.
Other good in-class examples are: 1. how students chose whether or not to attend your college or university and 2. what determines whether or not a student will graduate within four (or five) years. Finally, some students will be excited to learn that logits are used by marketing research companies to plan advertising and by political campaign media consultants to help predict and impact election votes.

13-3. Other Dummy Dependent Variable Techniques
Little will be lost if you skip this section, but we find that it’s well worth the time.

13-4. Summary and Exercises
For a problem set on this chapter, we’d recommend exercises 13-2, 13-5, 13-6, and 13-9.

Chapter Fourteen: Simultaneous Equations

14-1. Structural and Reduced-Form Equations
Our notation does not change to the double-subscript “β_{12}” system typically used in the discussion and analysis of simultaneous systems. We feel that switching notation for such a short chapter would probably do more harm than good.

14-2. The Bias of Ordinary Least Squares (OLS)
If Section 12.2.1 and/or Section 14.6 have or will be assigned, this is the perfect time to make the point that all three sections basically are talking about the same problem. That is, the violation of Classical Assumption III means that the observations of the error term are no longer independent of the explanatory variables, and OLS attributes variation in Y caused by ε to one of the Xs, causing bias.

14-3. Two-Stage Least Squares (2 SLS)
Most students will learn 2SLS better and a lot more quickly if they’re asked to estimate an equation with 2SLS “by hand” (that is, estimating the reduced-forms themselves with OLS and then using the estimated Ys in place of the actual Ys on the right-hand side of an OLS estimate of the second stage). After completing this project, most students will feel fairly comfortable with 2SLS. In addition, if the project also includes required OLS estimates and “packaged” 2SLS estimates of the structural equations, then the students will also be aware of the differences between 2SLS and OLS β estimates and between 2SLS-packaged SE (\(\hat{\beta}\)) and 2SLS “by hand” SE (\(\hat{\beta}\)). Exercise 14-9 provides just such a project, and the data required for the estimation are in Table 14.1.
14-4. The Identification Problem

Here’s an in-class example you can use to give the students experience with using the order condition:

Use the order condition to determine the likely identification properties of the equations in the following system:

1. \( Y_1 = f(Y_2, X_1, X_2, \varepsilon_1) \)
2. \( Y_2 = f(Y_3, Y_4, X_3, \varepsilon_2) \)
3. \( Y_3 = f(X_4, X_5, \varepsilon_3) \)
4. \( Y_4 = Y_1 + Y_3 \)

where the \( \varepsilon \)s are classical error terms.

(answer: Equation three is recursive to the system and doesn’t provide us with any information that will help identify the other equations. In addition, equation 4 is a nonstochastic identity and need not be identified. Thus there are two endogenous variables, \( Y_1 \) and \( Y_2 \) and four predetermined variables, \( X_1, X_2, X_3, \) and \( Y_3 \) (\( X_4 \) and \( X_5 \) show up only in Equation 3). As a result, the identification properties of only two equations need to be determined, and both of them have one fewer slope coefficient (3) than there are predetermined variables in the system (4), so they are overidentified.)

14-5. Summary and Exercises

For a problem set on this chapter, we’d recommend exercises 14-2, 14-7, 14-8, 14-9, and 14-13.

14-6. Appendix: Errors in the Variables

The key here is for students to understand that only errors in the measurement of independent variables cause bias.

Chapter Fifteen: Forecasting

15-1. What is Forecasting?

Most students will know what forecasting is by the time they get to this section, but the examples are nonetheless worth mentioning in class because we return to them later in the chapter and because the data sets for all of them are available in the text.

15-2. More Complex Forecasting Problems

At some point the students should be told that forecasts often are made by an individual with a superb “feel” for a market (or the economy as a whole) and then the econometric models are adjusted to come out with the answers that the “guru” expects. Whether in simulation analysis or forecasting, the use of such “add factors” and “mull factors” will catch most students by surprise and will open their eyes to the fact that econometric forecasting is not accurate enough by itself to place a lot of faith in.

15-3. ARIMA Models

We’ve shortened this coverage to do little more than introduce the topic to the reader.

15-4. Summary and Exercises

For a problem set on this chapter, we’d recommend exercises 15-2, 15-4a, and 15-8.
Chapter Sixteen: Statistical Principles

This chapter was written by Gary Smith of Pomona College. Gary is the author of our favorite statistics text, *Statistical Reasoning* (McGraw-Hill Publishers, 1998, ISBN 0-07-059276-4), and we owe a real debt of gratitude to Gary for taking on this task.

We’ve tried to set up the chapter so that it can be assigned in its entirety (in courses with a minimal stats requirement) or assigned on a section-specific basis (for readers who need brushing up on particular topics).

16-1. Probability Distributions
   Placing heavy emphasis on the section on standardized variables, Section 16.1.4, will pay off later in enhanced student understanding of the principle behind the normal distribution and the $t$-test.

16-2. Sampling
   This short section is so well-written and so universally applicable that we anticipate assigning it to all of our students when we cover data collection in Chapter 11.

16-3. Estimation

16-4. Hypothesis Tests
   We purposely created some overlap between Chapter 5 and these two sections so that students who are having trouble with Chapter 5 can get a bit more background and detail.

16-5. Summary and Exercises
   For a problem set on this chapter, we’d recommend exercises 16-5, 16-7, 16-9, 16-11, 16-13, and 16-15.
Sample Examinations

Each of the following sample exams follows the same format. Question one consists of four identifications from the chapters and one short application of the material. Question two asks the student “did you read the book?” about one of the most important topics covered. Question three asks the student to apply the materials of the chapters in a manner that parallels work already done, and question four attempts to see if the student can apply skills learned in the chapters to a regression result that he or she has not seen before. Questions from different sample exams can thus be combined to form tests that cover more chapters or that are harder or easier than these are. Answers to the questions (typically just textbook page references) follow each sample exam.

By the way, if your students seem especially nervous before a particular exam, you might relax them a bit with the following actual quote from a mayonnaise jar: “Keep cool but don’t freeze.”

Sample Exam for Chapters 1–3

1. Briefly identify the following in words or equations as appropriate:
   (a) Degrees of freedom
   (b) Estimated regression equation
   (c) The Six Steps in Applied Regression Analysis
   (d) Ordinary Least Squares
   (e) The meaning of $\beta_1$ in:

   \[ Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \]

2. Two of the most important econometric concepts to date have been the stochastic error term and the residual. Carefully distinguish between these two concepts, being sure to:
   (a) define both terms
   (b) state how they are similar,
   (c) state how they are different,
   (d) give an example of an equation that contains a stochastic error term, and
   (e) give an example of an equation that contains a residual.

3. A source of constant discussion among applied econometricians is the degree to which measures of the overall fit of an estimated equation also measure the quality of that regression. To date, we have introduced something like four different measures of overall fit, but the two most used are $R^2$ and $\bar{R}^2$.
   (a) Carefully distinguish between $R^2$ and $\bar{R}^2$.
   (b) Of the two, which do you recommend typically using? Why?
   (c) What drawbacks are there to the use of the measure you chose (as your answer to part (b) above) as the primary determinant of the overall quality of a regression?
4. Consider the following estimated equation:

\[ \hat{C}_t = 18.5 - 0.07P_t + 0.93YD_t - 0.74D_{1t} - 1.3D_{2t} - 1.3D_{3t} \]

where:
- \( C_t \) = Per-capita pounds of pork consumed in the U.S. in quarter \( t \).
- \( P_t \) = The price of a hundred pounds of pork (in dollars) in quarter \( t \).
- \( YD_t \) = Per capita disposable income (in dollars) in quarter \( t \).
- \( D_{1t} \) = Dummy equal to 1 in the first quarter (Jan.–Mar.) of the year and 0 otherwise.
- \( D_{2t} \) = Dummy equal to 1 in the second quarter of the year and 0 otherwise.
- \( D_{3t} \) = Dummy equal to 1 in the third quarter of the year and 0 otherwise.

(a) What is the meaning of the estimated coefficient of \( YD \)?
(b) Specify expected signs for each of the coefficients. Explain your reasoning.
(c) Suppose we changed the definition of \( D_{3t} \) so that it was equal to 1 in the fourth quarter and zero otherwise and re-estimated the equation with all the other variables unchanged. Which of the estimated coefficients would change? Would your answer to part (b) above change? Explain your answer.

**Answers:**

1. (a) See page 54.
(b) See pages 16–18.
(c) See page 67.
(d) See Section 2.1.
(e) See page 41.

2. See Sections 1.2.3 and 1.3.

3. See Sections 2.4.1 and 2.4.3.

4. We’d expect pork consumption to be the highest in the fourth quarter due to holidays, so the expected signs of the dummy coefficients are negative. For part (c), the coefficients of \( P \) and \( YD \) shouldn’t change, but the others should, because the omitted condition is now the third quarter and not the fourth quarter. Thus the expected signs of the coefficients of \( D_1 \) and \( D_2 \) are no longer negative. (Some of the best students will note that the estimate of the coefficient of \( D_2 \) will be almost exactly zero.)

**Sample Exam for Chapters 4–5**

*Note:* To do this test, students will need to consult a \( t \)-table like Statistical Table B-1 in the text. Since the table on page 611 is accompanied by a description of how to use it, we have repeated Table B-1 without the description on the inside back cover of this text. This will allow students to use a \( t \)-table (as long as no notes are written on it) without seeing the additional material if you so desire.

1. Briefly identify the following in words or equations as appropriate:
   (a) \( P \)-value
   (b) Sampling distribution of \( \hat{\beta} \)
   (c) Type I Error
   (d) Level of significance
   (e) Given \( \hat{\beta} = 4.0 \) and SE (\( \hat{\beta} \)) = 2.0, test the null hypothesis that \( \beta = 0 \) (versus the alternative hypothesis that it does not equal zero) with 13 degrees of freedom at the 1 percent level of significance.
2. In Chapter Four, we discussed seven explicit assumptions about the properties of a regression equation. These assumptions (or, more accurately, the first six of them) are usually referred to as the Classical Assumptions.
   (a) Carefully and accurately state the seven assumptions.
   (b) State what each assumption actually means in real-world terms (be extremely brief).

3. One of the shortest but most important sections in the book is that on the Gauss-Markov Theorem.
   (a) What is the Gauss-Markov Theorem?
   (b) Carefully explain what the precise properties specified by the Gauss-Markov theorem mean and why they are desirable for an equation.
   (c) What could cause the Gauss-Markov Theorem to no longer hold? What should we do in such a situation?

4. Consider the following regression equation for the United States (standard errors in parentheses):
   \[ \hat{P}_t = 4.00 - 0.010PRP_t + 0.030PRB_t + 0.20YD_t \]
   \[ (0.005) (0.020) (0.04) \]
   \[ R^2 = .98 \quad n = 29 \]
   where: \( P_t \) = per capita pounds of pork consumed in time period \( t \)
   \( PRP_t \) = the price of pork in time period \( t \)
   \( PRB_t \) = the price of beef in time period \( t \)
   \( YD_t \) = per capita disposable income in time period \( t \)
   (a) Hypothesize signs and specify the appropriate null and alternative hypotheses for the coefficients of each of these variables.
   (b) State your decision rules and then test your hypotheses on the above results using the \( t \)-test at a 5% level of significance.
   (c) If you could add one variable to the regression, what variable would you add? Why?

Answers:
1. (a) See pages 128–130.
   (b) See pages 97–98.
   (c) See Section 5.1.2.
   (d) See Section 5.2.3.
   (e) \( t = 2.0 \), and the critical two-sided, 1 percent, 13 degree of freedom critical \( t \)-value is 3.012, so we can not reject the null hypothesis.

2. See Section 4.1.

3. See Section 4.3.
4. Coefficient

<table>
<thead>
<tr>
<th>Hypothesized sign</th>
<th>$\beta_{PRP}$</th>
<th>$\beta_{PRH}$</th>
<th>$\beta_{YD}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculated $t$-score for a 5 percent one-sided test with 25 d.f. is 1.708, so</td>
<td>reject</td>
<td>do not reject</td>
<td>reject</td>
</tr>
</tbody>
</table>

For part (c), the two most important criteria are whether or not the suggested variable is a time series variable for the U.S. and whether or not that variable can be measured. “Tastes,” for example, are important but hard to measure.

**Sample Exam for Chapter 6–7**

1. Briefly identify the following in words or equations as appropriate:
   (a) The major consequence of including an irrelevant variable in a regression equation.
   (b) Four valid criteria for determining whether a given variable belongs in an equation.
   (c) The problem with sequential specification searches.
   (d) The elasticity of $Y$ with respect to $X$ in:

\[ \ln Y = \beta_0 + \beta_1 \ln X + \varepsilon \]

(e) The sign of the bias on the coefficient of age caused by omitting experience in an equation explaining the salaries of various workers.

2. Most of Chapter Six is concerned with specification bias.
   (a) What exactly is specification bias?
   (b) What is known to cause specification bias?
   (c) Are unbiased estimates always better than biased estimates? Why or why not?
   (d) What’s the best way to avoid specification bias?

3. There are at least two different possible approaches to the problem of building a model of the costs of production of electric power.
   I: Model I hypothesizes that per unit costs ($C$) as a function of the number of kilowatt-hours produced ($Q$) continually and smoothly falls as production is increased, but it falls at a decreasing rate.
   II: Model II hypothesizes that per unit costs ($C$) decrease fairly steadily as production ($Q$) increases, but costs decrease at a much faster rate for hydroelectric plants than for other kinds of facilities.

Given this information,
   (a) What functional form would you recommend for estimating Model I? Be sure to write out a specific equation.
   (b) What functional form would you recommend for estimating Model II? Be sure to write out a specific equation.
   (c) Would $R^2$ be a reasonable way to compare the overall fits of the two equations? Why or why not?

*Note: In the following question, you may want to change the absolute size of the coefficients, depending on the size of your school, but remember to change the size of the standard errors proportionally.*
4. One your way to the cashier’s office (to pay yet another dorm damage bill), you overhear U. R. Accepted (the Dean of Admissions) having a violent argument with I. M. Smart (the Director of the Computer Center) about an equation that Smart built to understand the number of applications that the school received from high school seniors. In need of an outside opinion, they turn to you to help them evaluate the following regression results (standard errors in parentheses):

\[
\hat{N}_t = 150 + 180A_t + 1.50\ln T_t + 30.0P_t
\]

\[
(90) \quad (1.50) \quad (60.0)
\]

\[R^2 = .50 \quad N = 22 \quad \text{(annual)}\]

where: \(N_t\) = the number of high school seniors who apply for admission in year \(t\).

\(A_t\) = the number of people on the admission staff who visit high schools full time spreading information about the school in year \(t\).

\(T_t\) = dollars of tuition in year \(t\).

\(P_t\) = the percent of the faculty in year \(t\) that had PhDs in year \(t\).

How would you respond if they asked you to:

(a) Discuss the expected signs of the coefficients.

(b) Compare these expectations with the estimated coefficients by using the appropriate tests.

(c) Evaluate the possible econometric problems that could have caused any observed differences between the estimated coefficients and what you expected.

(d) Determine whether the semi-log function for \(T\) makes theoretical sense.

(e) Make any suggestions you feel are appropriate for another run of the equation.

**Answers:**

1. (a) See Section 6.2.1.

(b) See Section 6.2.3.

(c) See Section 6.4.2.

(d) Constant = \(\beta_1\).

(e) positive, since the expected sign of the coefficient of experience is positive and since the simple correlation coefficient between age and experience is positive. See Section 6.1.3.

2. See Section 6.1. Note in particular that biased estimates can at times be closer to the true \(\beta\) than unbiased ones and that the solution to omitted variable bias is not to simply include every justifiable variable.

3. (a) A number of forms are possible, but a reciprocal form would be perhaps the most appropriate:

\[C_t = \beta_0 + \beta_1Q_t + \epsilon_t\]

(b) Such an hypothesis calls for the use of a slope dummy defined (for instance) as \(D = 1\) if the plant is hydroelectric and 0 otherwise. The resulting equation would be:

\[C_t = \beta_0 + \beta_1Q_t + \beta_2D_t + \beta_3D_tQ_t + \epsilon_t\]

(c) \(R^2\) is perfectly appropriate for comparing the overall fits of the two equations because the number of independent variables changes but the functional form of the dependent variable does not.
4. (a)/(b) Coefficient $\beta_1$ $\beta_2$ $\beta_3$

<table>
<thead>
<tr>
<th>Expected Sign</th>
<th>+</th>
<th>$-$</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-score</td>
<td>+2.0</td>
<td>+1.0</td>
<td>+0.5</td>
</tr>
<tr>
<td>Decision</td>
<td>reject</td>
<td>do not</td>
<td>do not</td>
</tr>
</tbody>
</table>

(given the 5 percent level and a resulting critical t-score of 1.734.)

(c) $\hat{\beta}_2$ has an insignificant unexpected sign, and $\hat{\beta}_3$ is not significant, so an irrelevant variable and an omitted variable(s) are both possible.

(d) Many would expect that the impact of an extra dollar of tuition (on applications) decreases as tuition increases. If this is so, the semi-log functional form makes sense; if not, then we should use the linear form as a default.

(e) Percent PhDs appears to be intended as a proxy for the way in which the quality of the school has changed over the years. If we could find a better proxy for quality, then it would make sense to substitute it for $X_r$ and that would be a possible answer. If the unexpected sign for $\ln T$ is being caused by an omitted variable, that omitted variable must be causing positive bias. Thus for a suggested omitted variable to be a good suggestion, it would have to be likely to have caused positive bias. Therefore, it must be either positively correlated with tuition and have a positive expected coefficient (like the number of high-school seniors in the country or in the region) or else have a negative correlation with tuition and have a negative expected coefficient (like the number of inadequate facilities, etc.).

Sample Exam for Chapters 8–10

Note: This exam uses heteroskedasticity as the topic of the “major” question, question number two, so the “IDs” in question number one avoid heteroskedasticity as a topic. Question number two can easily be reworded to cover serial correlation or multicollinearity, but then question number one should also be changed to cover heteroskedasticity.

1. Briefly identify the following in words or equations as appropriate:
   (a) Impure serial correlation
   (b) Dominant variable
   (c) Variance Inflation Factor
   (d) Generalized Least Squares
   (e) Given a calculated Durbin-Watson d statistic of 2.58, a $d_L$ of 1.21 and a $d_U$ of 1.55, what would you conclude?

2. Carefully outline (be brief!) a description of the problem typically referred to as pure heteroskedasticity.
   (a) What is it?
   (b) What are its consequences?
   (c) How do you diagnose it?
   (d) What do you do to get rid of it?
3. A model of the number of cars sold in the United States from 1980 through 2004 produced the following results (standard errors in parentheses):

\[
\hat{C}_t = 3738 - 48.0P_t + 10.0Y_t + 6.0A_t - 360.0R_t
\]

(12.0) (2.0) (2.0) (120.0)

\[R^2 = 0.85 \quad DW = 1.86 \quad N = 25\text{ (annual)}\]

where: \(C_t\) = thousands of cars sold in year \(t\)

\(P_t\) = price index for domestic cars in year \(t\)

\(Y_t\) = disposable income (billions of dollars) in year \(t\)

\(A_t\) = billions of dollars of auto industry advertising expenditures in year \(t\)

\(R_t\) = the interest rate in year \(t\)

(a) Hypothesize the expected signs of the coefficients and test the appropriate null hypotheses at the 1 percent level.

(b) What econometric problems appear to be present in this equation? Why?

(c) Suppose you were now told that the simple correlation coefficients between \(P\), \(A\), and \(Y\) were all between 0.88 and 0.94 and that a Park test with \(Y\) as \(Z\) produced a \(t\)-score of 0.50. Would your answer to part (b) above change? Why or why not? How would it change?

(d) What suggestions would you have for another run of this regression?

4. In a study of the long-run and short-run demand for money, Chow estimated the following demand equation (standard errors in parentheses) for the U.S. from 1947:1 through 1965:4:

\[
\hat{M}_t = 0.14 + 1.05Y^*_t - 0.01Y_t - 0.75R_t
\]

(0.15) (0.10) (0.05)

\[R^2 = 0.996 \quad DW = 0.88 \quad N = 76\text{ (quarterly)}\]

where: \(M_t\) = the log of the money stock in quarter \(t\)

\(Y^*_t\) = the log of permanent income (a moving average of previous quarters’ current income) in quarter \(t\)

\(Y_t\) = the log of current income in quarter \(t\)

\(r_t\) = the log of the rate of interest in quarter \(t\)

(a) Hypothesize signs and test the appropriate null hypotheses at the 5 percent level of significance.

(b) What econometric problems seem likely to be in this equation?

(c) In particular, are there any problems related to the coefficient of \(Y\)? If so, are these problems more likely to have been caused by multicollinearity, serial correlation, or heteroskedasticity?

(d) What suggestions would you have for another estimation of this equation? Why?
Answers:

1. (a) See Section 9.1.2.
(b) See Section 8.1.1.
(c) See Section 8.3.2.
(d) See Section 9.4.1.
(e) The answer depends on whether you encourage your students to use one-sided or two-sided tests. For a one-sided test, the correct answer is that we can reject the null hypothesis of positive serial correlation. For a two-sided test, the correct answer is that the test is inconclusive.

2. The best answer would be only slightly more detailed than the summary on pages 376–377.

3. (a) Coefficient $\beta_y$ $\beta_Y$ $\beta_A$ $\beta_R$
   
   | Expected Sign | $-\$ | $+$ | $+$ | $-$ |
   | Calculated t-score | $-4.0$ | $5.0$ | $3.0$ | $-3.0$ |
   | Decision | reject | reject | reject | reject |

   (given a 1 percent critical $t$-score of 2.528) In addition, the DW = 1.86 is higher than the $d_u$ of 1.77, so there is no evidence of positive serial correlation.

(b) There may be theoretically sound variables that have been omitted, for instance some measure of competition from foreign-made cars, but the results themselves give no indication of any problems.

(c) These results are clear indications of multicollinearity but not of heteroskedasticity.

(d) Unless a theoretically sound variable can be added measuring competition, the regression need not be changed at all.

4. (a) Coefficient $\hat{\beta}_Y$ $\beta_Y$ $\beta_R$
   
   | Expected Sign | $+$ | $+$ | $-$ |
   | Calculated $t$-score | $7.0$ | $-0.1$ | $-15.0$ |
   | Decision | reject | do not reject | reject |

   (given a 5 percent critical $t$-score of about 1.67) In addition, the DW = 0.88 is less than the $d_L$ of 1.54, so we can reject the null hypothesis of no positive serial correlation.

(b) The poor significance of $\hat{\beta}_Y$ could have been caused by multicollinearity, an omitted variable, or by an irrelevant variable ($Y$), and in addition, as mentioned above, we have serial correlation.

(c) Of the three, only multicollinearity between $Y^*$ and $Y$ could have caused the problem, since serial correlation and heteroskedasticity would cause the $t$-scores to appear higher than they actually should be.

(d) This is not an easy question, but we would not suggest dropping $Y$, nor would we consider GLS right now. It turns out that making the equation a distributed lag by adding $M_{t-1}$ to the right-hand side switches the insignificance to $Y^*$ (and raises the DW, naturally) giving another indication of multicollinearity. The best action right now would seem to be to review the theory in search of an omitted variable and failing to find one, do nothing. This problem was adapted from a source, Edward Tower, *Econometrics Exams, Puzzles and Problems* (Durham, N.C.: Eno River Press, 1985) p. 167, that is just full of questions of a similar (or of a more theoretical) nature.
Sample Exam for Chapters 12–15

1. Briefly identify the following in words or equations as appropriate:
   (a) Problems with ad hoc distributed lags
   (b) Unconditional forecasting
   (c) Moving-average process
   (d) How to test for serial correlation in a dynamic model.
   (e) The identification properties (use the order condition only) of each of the equations in the following system:
      i. \( Y_t = f(Y_{t-1}, X_1, X_2) \)
      ii. \( Y_{t-1} = f(Y_t, Y_{t-1}) \)
      iii. \( Y_{t-2} = f(X_{t-3}, X_{t-4}) \)

2. Virtually all of Chapter 14 is spent discussing the violation of the assumption that the error term is independent of the explanatory variables.
   (a) Under what circumstances might that assumption be violated?
   (b) What would the violation of that assumption be likely to cause?
   (c) What general technique is used to rid the equation of this problem? Specifically, how does it work?

3. Suppose you’ve been hired by your school’s admissions department to help them decide whether to change admissions procedures. You are given the files of all of the students in the last graduating class (including those students who didn’t graduate) and told to build a model to explain why some admitted students graduate and others don’t.
   (a) Specify the functional form you would use in building such a model and carefully explain why that form is appropriate.
   (b) Specify the independent variables you would include in the equation and briefly explain how they apply to the dependent variable in question.
   (c) Carefully explain the meaning of the coefficient of your first independent variable.

4. You have been hired to forecast GDP (\( Y \)) for the Caribbean island of Tabasco. Tabasco has domestic food (\( F \)) and shelter (\( S \)) industries, a tourist (\( T \)) industry, and an export (\( X \)) industry. All tourists come from the United States, while the exports are split between Mexico and the U.S. Investment is virtually zero, and government expenditures (\( G \)) can be considered to be exogenously determined. Imports (\( I \)) are a function of GDP. Thus the structural equations for a simultaneous model of the Tabascan economy would look something like:
   \( Y = F + S + T + X + G - I \)
   \( F = f_F(Y, ?) \)
   \( S = f_S(Y, ?) \)
   \( T = f_T(USGNP, ?) \)
   \( X = f_X(USGNP, MEXICOGNP, ?) \)
   \( I = f_I(Y, ?) \)
   \( G = \text{exogenous} \)
   (a) Develop a theory for Tabasco’s economy. Then choose which predetermined variables you would like to add to the simultaneous system and specify to which of the five stochastic structural equations (see question marks) you would like to add them. Explain your reasoning.
   (b) Comment on the identification properties of each of the five stochastic equations in the system you outlined in your answer to part a above.
(c) How should the coefficients of the system be estimated?
(d) What technique would you use to forecast Tabasco’s GNP? Why?

Answers:
1. 
   (a) See Section 12.1.1.
   (b) See Section 15.2.1.
   (c) See Section 15.3.
   (d) See Section 12.2.2.
   (e) There are three predetermined variables in the system (X1, X2, and Y3). This means that i is exactly identified because there are also three slope coefficients to be estimated and that ii is overidentified, since there are only two slope coefficients to be estimated.

2. See Sections 14.2 and 14.3. This answer will also depend on whether or not Chapter 12 and/or Section 14.6 have been assigned.

3. The appropriate functional form is the logit because of the problems with the linear probability model outlined in Section 13.1.2. The key to choosing independent variables is the type of variable suggested; some students will misunderstand the disaggregate nature of the variables required by such a study and will suggest variables that are constant for all observations in the dataset. Each coefficient tells the impact of a one-unit change in the independent variable in question (holding constant all the other independent variables in the equation) on the log of the odds that the person graduated.

4. OK, OK, we know this will be hard to grade, since each answer will be different depending on the exact variables and equations added, but this question tends to work well. The student will be forced to apply the identification, estimation, and forecasting techniques to a system of their own choosing in much the same way they will have to in their work later on.

   The key to the questions on estimation and forecasting have to do with the size and importance of the feedback loops (as compared to exogenous factors in determining GDP). In this case, there is a good chance that the most accurate forecast of Tabasco’s GDP would be based on a “simplified reduced-form” equation that included only USGNP and MEXICOGNP as explanatory variables.
An Additional Interactive Regression Learning Exercise

In this final section, we present a “demand for water” interactive learning regression exercise. This interactive exercise is intended to be used in class to give students experience with specification issues like omitted variables, irrelevant variables, multicollinearity, and, possibly, serial correlation. Thus the best time to use the exercise is sometime between Chapters Eight and Ten. A hoped-for byproduct is to make the students more comfortable with interactive exercises in general.

One way to use this exercise in class is to introduce the topic, define all the available variables, encourage class discussion of expected signs and significance, and then have the class pick their guess at a “best” specification. After you put the estimated result on the board (or hand out copies of it), the class should then test hypotheses, discuss probable econometric difficulties, and choose whether to change the specification. While only 26 regression runs are provided here, instructors who wish to have more available are encouraged to use the dataset and estimate them. Indeed, one way of handling this (or any other) interactive exercise is to bring a PC to class and estimate the regressions after they’ve been specified.

Obviously, it’s vital to give feedback along the way, particularly since otherwise some students might get the idea that we can estimate as many runs as we want. Even given this, we try hard to make the students choose a course of action before we comment on it.

Building A Model of the Demand for Water in Los Angeles

This model is an extension of the short example presented in Section 2.5. The goal is to build a model of the demand for water in Los Angeles County as a function of population, rainfall and other variables. The dependent variable is:

\[ W_t = \text{Consumption of water in Los Angeles county in year } t \text{ (hundreds of millions of gallons).} \]

A few comments about the data are in order. Population data for Los Angeles county were unavailable for the entire sample when the data were collected, so population in Glendale was used as a proxy. Second, conservation efforts took place only during the last two years in the sample, so if conservation efforts have an effect but with a lag, we’d be forced to use a “one-time dummy.” Thus there are reasons to be concerned that population and conservation are not accurately measured in the sample. The means, standard deviations, and simple correlation coefficients of the variables, plus the raw data themselves, are contained in tables on the next two pages.
The available independent variables are:

\[ \text{POP}_t = \text{Population in Glendale, California in year } t \text{ (thousands of people).} \]

\[ T_t = \text{Average annual temperature at the Los Angeles civic center in year } t \text{ (degrees F).} \]

\[ R_t = \text{Inches of rainfall at the Los Angeles civic center in year } t. \]

\[ \text{PR}_t = \text{Average price of a gallon of water in Los Angeles in year } t \text{ (dollars).} \]

\[ \text{CO}_t = \text{A dummy variable to measure the conservation efforts undertaken in the last two years in the sample, when there was a severe drought in Northern California (the source of much of L.A.'s water) even though no such drought took place in Southern California.} \]

\[ Y_t = \text{Total personal income for Los Angeles county in year } t \text{ (in billions of dollars).} \]

**Means and Variances for the Demand for Water Data**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>71.01</td>
<td>7.97</td>
</tr>
<tr>
<td>POP</td>
<td>126.83</td>
<td>9.78</td>
</tr>
<tr>
<td>T</td>
<td>65.15</td>
<td>1.24</td>
</tr>
<tr>
<td>R</td>
<td>13.75</td>
<td>6.20</td>
</tr>
<tr>
<td>PR</td>
<td>0.33</td>
<td>0.13</td>
</tr>
<tr>
<td>CO</td>
<td>0.08</td>
<td>0.27</td>
</tr>
<tr>
<td>Y</td>
<td>26.91</td>
<td>13.55</td>
</tr>
</tbody>
</table>

**Simple Correlation Coefficients for the Data**

| Y, PR  | 0.954 | CO, PR  | 0.764 |
| Y, T   | 0.108 | CO, T   | 0.096 |
| Y, CO  | 0.639 | CO, R   | -0.036|
| Y, POP | 0.828 | CO, POP | 0.294 |
| Y, R   | 0.009 | T, POP  | 0.158 |
| PR, POP| 0.630 | T, PR   | 0.079 |
| PR, R  | -0.024| R, T    | -0.063|
|        |       | R, POP  | 0.124 |
An Additional Interactive Regression Learning Exercise  57

**Data for the Demand for Water Interactive Exercise**

<table>
<thead>
<tr>
<th>YEAR</th>
<th>W</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53.385</td>
<td>108.7</td>
<td>63.6</td>
<td>26.2</td>
<td>0.23</td>
<td>9.5345</td>
</tr>
<tr>
<td>2</td>
<td>60.311</td>
<td>110.7</td>
<td>64.6</td>
<td>9.46</td>
<td>0.22</td>
<td>10.783</td>
</tr>
<tr>
<td>3</td>
<td>60.064</td>
<td>113.3</td>
<td>64.9</td>
<td>11.99</td>
<td>0.22</td>
<td>11.991</td>
</tr>
<tr>
<td>4</td>
<td>58.939</td>
<td>114.7</td>
<td>63.7</td>
<td>11.94</td>
<td>0.23</td>
<td>13.019</td>
</tr>
<tr>
<td>5</td>
<td>62.330</td>
<td>115.7</td>
<td>64.8</td>
<td>16.00</td>
<td>0.22</td>
<td>14.288</td>
</tr>
<tr>
<td>6</td>
<td>62.361</td>
<td>117.1</td>
<td>65.7</td>
<td>9.54</td>
<td>0.23</td>
<td>15.428</td>
</tr>
<tr>
<td>7</td>
<td>65.409</td>
<td>118.2</td>
<td>67.1</td>
<td>21.13</td>
<td>0.25</td>
<td>15.876</td>
</tr>
<tr>
<td>8</td>
<td>71.914</td>
<td>118.3</td>
<td>67.3</td>
<td>5.58</td>
<td>0.25</td>
<td>17.327</td>
</tr>
<tr>
<td>9</td>
<td>70.417</td>
<td>119.4</td>
<td>65.7</td>
<td>8.18</td>
<td>0.26</td>
<td>18.002</td>
</tr>
<tr>
<td>10</td>
<td>73.549</td>
<td>121.8</td>
<td>65.3</td>
<td>4.83</td>
<td>0.26</td>
<td>18.888</td>
</tr>
<tr>
<td>11</td>
<td>69.093</td>
<td>123.1</td>
<td>62.3</td>
<td>18.79</td>
<td>0.27</td>
<td>20.342</td>
</tr>
<tr>
<td>12</td>
<td>67.479</td>
<td>126.7</td>
<td>65.1</td>
<td>8.38</td>
<td>0.26</td>
<td>21.581</td>
</tr>
<tr>
<td>13</td>
<td>73.277</td>
<td>129.8</td>
<td>63.2</td>
<td>7.93</td>
<td>0.26</td>
<td>22.979</td>
</tr>
<tr>
<td>14</td>
<td>68.947</td>
<td>132.4</td>
<td>64.3</td>
<td>13.69</td>
<td>0.26</td>
<td>24.228</td>
</tr>
<tr>
<td>15</td>
<td>73.437</td>
<td>135.7</td>
<td>66.0</td>
<td>20.44</td>
<td>0.26</td>
<td>26.099</td>
</tr>
<tr>
<td>16</td>
<td>72.340</td>
<td>137.0</td>
<td>66.5</td>
<td>22.00</td>
<td>0.28</td>
<td>27.924</td>
</tr>
<tr>
<td>17</td>
<td>76.044</td>
<td>137.2</td>
<td>66.2</td>
<td>16.58</td>
<td>0.30</td>
<td>29.908</td>
</tr>
<tr>
<td>18</td>
<td>75.155</td>
<td>137.9</td>
<td>65.4</td>
<td>27.47</td>
<td>0.34</td>
<td>32.407</td>
</tr>
<tr>
<td>19</td>
<td>81.664</td>
<td>132.6</td>
<td>66.1</td>
<td>7.77</td>
<td>0.37</td>
<td>34.098</td>
</tr>
<tr>
<td>20</td>
<td>80.544</td>
<td>132.7</td>
<td>65.1</td>
<td>12.32</td>
<td>0.39</td>
<td>34.314</td>
</tr>
<tr>
<td>21</td>
<td>81.418</td>
<td>133.1</td>
<td>66.5</td>
<td>6.54</td>
<td>0.43</td>
<td>37.116</td>
</tr>
<tr>
<td>22</td>
<td>76.707</td>
<td>132.7</td>
<td>65.0</td>
<td>17.45</td>
<td>0.44</td>
<td>39.191</td>
</tr>
<tr>
<td>23</td>
<td>80.615</td>
<td>136.5</td>
<td>65.3</td>
<td>16.69</td>
<td>0.47</td>
<td>43.784</td>
</tr>
<tr>
<td>24</td>
<td>80.062</td>
<td>138.0</td>
<td>63.0</td>
<td>10.70</td>
<td>0.51</td>
<td>47.822</td>
</tr>
<tr>
<td>25</td>
<td>81.467</td>
<td>139.2</td>
<td>65.2</td>
<td>11.01</td>
<td>0.60</td>
<td>53.334</td>
</tr>
<tr>
<td>26</td>
<td>69.315</td>
<td>133.2</td>
<td>65.9</td>
<td>14.97</td>
<td>0.75</td>
<td>59.267</td>
</tr>
</tbody>
</table>

**How to Find the Desired Regression Run**

Note that the possible variables include a theoretically strong variable (R), a probable irrelevant variable (T), a possible irrelevant variable (CO), two important but basically redundant variables (POP and Y), and a potentially weak (at least theoretically) variable (PR) that is also multicollinear with POP and Y. One of the nicest attributes of this exercise is that when you make a specification change, most statistics (from SEs to $R^2$ s) act the way they “should;” believe us, finding such examples is not easy. With respect to serial correlation, all classes so far have ended up choosing as their final specification regression runs one of #26, #25, #24, #11, or #8, of which only #26 has a DW below the critical $d_L$ value (#11 is on the borderline). A GLS version of equation #26 appears as equation #27.
Once the class has chosen its “best” specification, refer to the following key to find the number of the regression run chosen. This interactive exercise has been used extensively in class, and the following 26 runs contain all the choices that have been chosen during that time, but there are a number of other specifications that could also be chosen. If the class picks one of these, shift to the closest alternative.

<table>
<thead>
<tr>
<th>Run #</th>
<th># of Xs</th>
<th>KEY to Included Independent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>All</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>All but POP</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>All but T</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>All but R</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>All but PR</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>All but CO</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>All but Y</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>All but POP and T</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>All but Y and T</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>All but R and T</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>All but PR and T</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>All but CO and T</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>All but R and PR</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>All but R and CO</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>All but R and Y</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>All but PR and CO</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>All but PR and Y</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>All but CO and Y</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>All but POP and R</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>All but POP and PR</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>All but POP and CO</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>POP, R, and CO are included</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>POP, R, and PR are included</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>R, CO, and Y are included</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>R, PR, and Y are included</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>POP and R are included</td>
</tr>
</tbody>
</table>

Regression Run # 1

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>7.6</td>
<td>0.07</td>
<td>0.79</td>
<td>-0.33</td>
<td>-63.4</td>
<td>-7.20</td>
<td>1.08</td>
<td>1.91</td>
<td>.84</td>
</tr>
<tr>
<td>SE</td>
<td>0.46</td>
<td>0.52</td>
<td>0.11</td>
<td>68.9</td>
<td>4.49</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>0.2</td>
<td>1.5</td>
<td>-3.0</td>
<td>-0.9</td>
<td>-1.6</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To drop POP see #2; to drop PR see #5; otherwise see key.
Regression Run #2

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>15.0</td>
<td>0.80</td>
<td>-0.32</td>
<td>-73.5</td>
<td>-6.92</td>
<td>1.21</td>
<td>1.89</td>
<td>.85</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.51</td>
<td>0.10</td>
<td>20.9</td>
<td>3.98</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>1.6</td>
<td>-3.2</td>
<td>-3.5</td>
<td>-1.7</td>
<td>6.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add POP see #1; to drop T see #8; otherwise see key.

Regression Run #3

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>46.0</td>
<td>0.18</td>
<td>-0.35</td>
<td>-52.6</td>
<td>-6.90</td>
<td>0.92</td>
<td>1.64</td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.47</td>
<td>0.11</td>
<td>70.6</td>
<td>4.62</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>0.4</td>
<td>-3.1</td>
<td>-0.7</td>
<td>-1.5</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add T see #1; to drop POP see #8; otherwise see key.

Regression Run #4

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>54.0</td>
<td>-0.45</td>
<td>0.99</td>
<td>-132.0</td>
<td>-5.44</td>
<td>2.01</td>
<td>1.27</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.50</td>
<td>0.61</td>
<td>76.4</td>
<td>5.24</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>-0.9</td>
<td>1.6</td>
<td>-1.7</td>
<td>-1.0</td>
<td>2.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add R see #1; to drop POP see #19; otherwise see key.

Regression Run #5

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>-39.0</td>
<td>0.47</td>
<td>0.74</td>
<td>-0.36</td>
<td>-9.63</td>
<td>0.28</td>
<td>2.10</td>
<td>.85</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.14</td>
<td>0.52</td>
<td>0.10</td>
<td>3.62</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>3.3</td>
<td>1.4</td>
<td>-3.5</td>
<td>-2.7</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #1; to drop T see #11; otherwise see key.

Regression Run #6

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>51.0</td>
<td>-0.24</td>
<td>0.75</td>
<td>-0.31</td>
<td>-128.0</td>
<td>1.78</td>
<td>1.38</td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.43</td>
<td>0.54</td>
<td>0.11</td>
<td>57.9</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>-0.5</td>
<td>1.4</td>
<td>-2.7</td>
<td>-2.2</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add CO see #1; to drop POP see #21; otherwise see key.

Regression Run #7

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>-55.0</td>
<td>0.62</td>
<td>0.71</td>
<td>-0.38</td>
<td>20.0</td>
<td>-9.92</td>
<td>2.10</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.09</td>
<td>0.53</td>
<td>0.10</td>
<td>10.0</td>
<td>3.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>6.6</td>
<td>1.3</td>
<td>-3.6</td>
<td>2.0</td>
<td>-2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add Y see #1; to drop T see #9; otherwise see key.
### Regression Run #8

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>68.0</td>
<td>–0.33</td>
<td>–78.4</td>
<td>–6.14</td>
<td>1.26</td>
<td>1.58</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.10</td>
<td>21.5</td>
<td>4.09</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>–3.2</td>
<td>–3.7</td>
<td>–1.5</td>
<td>7.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add POP see #3; to drop CO see #25; otherwise see key.

### Regression Run #9

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>–11.0</td>
<td>0.64</td>
<td>–0.39</td>
<td>18.4</td>
<td>–9.27</td>
<td>1.82</td>
<td>1.82</td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.09</td>
<td>0.11</td>
<td>10.1</td>
<td>4.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>6.9</td>
<td>–3.7</td>
<td>1.8</td>
<td>–2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add Y see #3; to drop PR see #22; otherwise see key.

### Regression Run #10

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>107.0</td>
<td>–0.36</td>
<td>–124.0</td>
<td>–4.90</td>
<td>1.88</td>
<td>1.19</td>
<td>.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.51</td>
<td>79.0</td>
<td>5.43</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>–0.7</td>
<td>–1.6</td>
<td>–0.9</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add R see #3; to add T see #4; otherwise see key.

### Regression Run #11

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>5.5</td>
<td>0.51</td>
<td>–0.38</td>
<td>–8.95</td>
<td>0.25</td>
<td>1.81</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.14</td>
<td>0.11</td>
<td>3.67</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>3.6</td>
<td>–3.6</td>
<td>–2.4</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #3; to drop Y see #22; otherwise see key.

### Regression Run #12

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>86.0</td>
<td>–0.12</td>
<td>–0.33</td>
<td>–115.0</td>
<td>1.59</td>
<td>1.16</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.43</td>
<td>0.12</td>
<td>58.0</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>–0.3</td>
<td>–2.8</td>
<td>–2.0</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add CO see #3; to drop POP see #25; otherwise see key.

### Regression Run #13

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>–43.0</td>
<td>0.36</td>
<td>0.93</td>
<td>–10.7</td>
<td>0.35</td>
<td>1.62</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.17</td>
<td>0.64</td>
<td>4.5</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>2.1</td>
<td>1.4</td>
<td>–2.4</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add R see #5; to add PR see #4; otherwise see key.
Regression Run #14

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>84.0</td>
<td>–0.66</td>
<td>0.95</td>
<td>–178.</td>
<td>2.50</td>
<td>1.03</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.46</td>
<td>0.61</td>
<td>62.4</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>–1.4</td>
<td>1.6</td>
<td>–2.9</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add R see #6; to add CO see #4; otherwise see key.

Regression Run #15

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>–65.0</td>
<td>0.55</td>
<td>0.89</td>
<td>24.5</td>
<td>–10.7</td>
<td>1.62</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.12</td>
<td>0.66</td>
<td>12.4</td>
<td>4.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>4.8</td>
<td>1.4</td>
<td>2.0</td>
<td>–2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add R see #7; to add Y see #4; otherwise see key.

Regression Run #16

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>–47.0</td>
<td>0.68</td>
<td>0.56</td>
<td>–0.39</td>
<td>0.03</td>
<td>1.19</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.13</td>
<td>0.59</td>
<td>0.12</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.1</td>
<td>1.0</td>
<td>–3.3</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #6; to add CO see #5; otherwise see key.

Regression Run #17

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>–56.0</td>
<td>0.75</td>
<td>0.58</td>
<td>–0.40</td>
<td>–3.73</td>
<td>1.37</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.07</td>
<td>0.56</td>
<td>0.11</td>
<td>2.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>10.0</td>
<td>1.0</td>
<td>–3.6</td>
<td>–1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #7; to drop T see #22; otherwise see key.

Regression Run #18

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>–50.0</td>
<td>0.71</td>
<td>0.55</td>
<td>–0.39</td>
<td>0.43</td>
<td>1.18</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.10</td>
<td>0.59</td>
<td>0.12</td>
<td>6.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>7.4</td>
<td>0.9</td>
<td>–3.3</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add CO see #7; to drop T see #23; otherwise see key.

Regression Run #19

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>$\hat{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>2.1</td>
<td>0.93</td>
<td>–66.3</td>
<td>–7.35</td>
<td>1.15</td>
<td>1.38</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.61</td>
<td>25.1</td>
<td>4.79</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>1.5</td>
<td>–2.6</td>
<td>–1.5</td>
<td>5.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add R see #2; to add POP see #4; otherwise see key.
Regression Run #20

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.20</td>
<td>1.06</td>
<td>0.40</td>
<td>2.37</td>
<td>0.07</td>
<td>0.64</td>
<td>2.03</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.04</td>
<td>0.12</td>
<td>0.07</td>
<td>0.11</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.3</td>
<td>1.7</td>
<td>0.3</td>
<td>2.3</td>
<td>0.1</td>
<td>2.8</td>
<td>3.2</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #2; to drop T see #24; otherwise see key.

Regression Run #21

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.20</td>
<td>0.69</td>
<td>0.33</td>
<td>1.05</td>
<td>0.07</td>
<td>0.35</td>
<td>1.33</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.04</td>
<td>0.10</td>
<td>0.25</td>
<td>0.12</td>
<td>0.05</td>
<td>0.16</td>
<td>0.33</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.3</td>
<td>1.3</td>
<td>0.1</td>
<td>2.5</td>
<td>0.1</td>
<td>2.4</td>
<td>3.2</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

To add CO see #2; to drop T see #25; otherwise see key.

Regression Run #22

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.20</td>
<td>0.76</td>
<td>0.33</td>
<td>0.35</td>
<td>0.07</td>
<td>0.35</td>
<td>1.35</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.04</td>
<td>0.11</td>
<td>0.25</td>
<td>0.12</td>
<td>0.05</td>
<td>0.16</td>
<td>0.33</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.3</td>
<td>1.3</td>
<td>0.1</td>
<td>2.5</td>
<td>0.1</td>
<td>2.4</td>
<td>3.2</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #9; to drop CO see #26; otherwise see key.

Regression Run #23

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.20</td>
<td>0.73</td>
<td>0.40</td>
<td>0.35</td>
<td>0.07</td>
<td>0.35</td>
<td>1.35</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.04</td>
<td>0.11</td>
<td>0.25</td>
<td>0.12</td>
<td>0.05</td>
<td>0.16</td>
<td>0.33</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.3</td>
<td>1.3</td>
<td>0.1</td>
<td>2.5</td>
<td>0.1</td>
<td>2.4</td>
<td>3.2</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

To add CO see #9; to drop PR see #26; otherwise see key.

Regression Run #24

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.20</td>
<td>0.33</td>
<td>0.33</td>
<td>0.35</td>
<td>0.07</td>
<td>0.35</td>
<td>1.35</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.04</td>
<td>0.11</td>
<td>0.25</td>
<td>0.12</td>
<td>0.05</td>
<td>0.16</td>
<td>0.33</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.3</td>
<td>1.3</td>
<td>0.1</td>
<td>2.5</td>
<td>0.1</td>
<td>2.4</td>
<td>3.2</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #8; to add POP see #11; otherwise see key.

Regression Run #25

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.20</td>
<td>0.33</td>
<td>0.33</td>
<td>0.35</td>
<td>0.07</td>
<td>0.35</td>
<td>1.35</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.04</td>
<td>0.11</td>
<td>0.25</td>
<td>0.12</td>
<td>0.05</td>
<td>0.16</td>
<td>0.33</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>5.3</td>
<td>1.3</td>
<td>0.1</td>
<td>2.5</td>
<td>0.1</td>
<td>2.4</td>
<td>3.2</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

To add CO see #8; to add POP see #12; otherwise see key.
## Regression Run #26

<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>DW</th>
<th>( \hat{R}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta} )</td>
<td>-16.0</td>
<td>0.73</td>
<td>-0.40</td>
<td>0.07</td>
<td>1.05</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.07</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>10.1</td>
<td>-3.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To add PR see #23; to add CO see #22; otherwise see key.


<table>
<thead>
<tr>
<th>#</th>
<th>INT</th>
<th>POP</th>
<th>T</th>
<th>R</th>
<th>PR</th>
<th>CO</th>
<th>Y</th>
<th>Rho</th>
<th>( \hat{R}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta} )</td>
<td>-11.0</td>
<td>0.68</td>
<td>-0.30</td>
<td>0.17</td>
<td>0.546</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.17</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>4.1</td>
<td>-2.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>