CHAPTER 12

SOLUTIONS TO PROBLEMS

12.1 We can reason this from equation (12.4) because the usual OLS standard error is an estimate of $\sigma / \sqrt{SST_x}$. When the dependent and independent variables are in level (or log) form, the AR(1) parameter, ρ, tends to be positive in time series regression models. Further, the independent variables tend to be positive correlated, so $(x_t - \bar{x})(x_{t+j} - \bar{x})$ – which is what generally appears in (12.4) when the $\{x_t\}$ do not have zero sample average – tends to be positive for most t and j. With multiple explanatory variables the formulas are more complicated but have similar features.

If $\rho < 0$, or if the $\{x_t\}$ is negatively autocorrelated, the second term in the last line of (12.4) could be negative, in which case the true standard deviation of $\hat{\beta}_1$ is actually less than $\sigma / \sqrt{SST_x}$.

12.3 (i) Because U.S. presidential elections occur only every four years, it seems reasonable to think that the unobserved shocks – that is, elements in u_t – in one election have pretty much dissipated four years later. This would imply that $\{u_t\}$ is roughly serially uncorrelated.

(ii) The t statistic for H_0: $\rho = 0$ is $-.068/.240 \approx -.28$, which is very small. Further, the estimate $\hat{\rho} = -.068$ is small in a practical sense, too. There is no reason to worry about serial correlation in this example.

(iii) Because the test based on $t_{\hat{\rho}}$ is only justified asymptotically, we would generally be concerned about using the usual critical values with $n = 20$ in the original regression. But any kind of adjustment, either to obtain valid standard errors for OLS as in Section 12.5 or a feasible GLS procedure as in Section 12.3, relies on large sample sizes, too. (Remember, FGLS is not even unbiased, whereas OLS is under TS.1 through TS.3.) Most importantly, the estimate of ρ is *practically* small, too. With $\hat{\rho}$ so close to zero, FGLS or adjusting the standard errors would yield similar results to OLS with the usual standard errors.

12.5 (i) There is substantial serial correlation in the errors of the equation, and the OLS standard errors almost certainly underestimate the true standard deviation in $\hat{\beta}_{EZ}$. This makes the usual confidence interval for β_{EZ} and t statistics invalid.

(ii) We can use the method in Section 12.5 to obtain an approximately valid standard error. [See equation (12.43).] While we might use $g = 2$ in equation (12.42), with monthly data we might want to try a somewhat longer lag, maybe even up to $g = 12$.

12.7 (i) The usual Prais-Winsten standard errors will be incorrect because the usual transformation will not fully eliminate the serial correlation in u_t. The regression of the OLS residuals on a single lag can be expected to consistently estimate the correlation coefficient, but the transformed variables $u_t - \rho u_{t-1}$ will have serial correlation (where ρ is the correlation between u_t and u_{t-1} under stationarity).

(ii) Equation (12.32) shows the transformed equation used by the Prais-Winsten method (without the first time period, which is unimportant for this discussion). If the AR(1) model is misspecified, then the error term in (12.32), e_t, is still serially correlated. But the Newey-West method works, in theory, for any kind of serial correlation. Therefore, we can just apply Newey-West directly to (12.32). Note that if we use the first observation, our estimates of the β_j are still the P-W estimates. We are simply obtaining standard errors that are robust to other forms of serial correlation. If we drop the first observation, then we obtain the Cochrane-Orcutt estimates with standard errors robust to arbitrary serial correlation.

It may make sense to try to improve over OLS – assuming strict exogeneity, of course – by estimating a simple model of serial correlation to remove "most" of the serial correlation. But then we might want our inference to be robust to more general forms of serial correlation.

(iii) If we drop the homoskedasticity assumption, then the e_t in equation (12.32) exhibit heteroskedasticity as well as serial correlation. But Newey-West handles both problems at the same time. So we need do nothing different if we think the u_t were heteroskedastic. (Of course, as discussed in Section 12.6, we could model the heteroskedasticity, too, and try to adjust for both. See pages 438 and 439. Even if we undertake such an exercise, we can still use Newey-West standard errors on the transformed equation.)

SOLUTIONS TO COMPUTER EXERCISES

C12.1 Regressing $\hat{u}_t$ on $\hat{u}_{t-1}$, using the 69 available observations, gives $\hat{\rho} \approx .292$ and se($\hat{\rho}$) $\approx$ .118. The t statistic is about 2.47, and so there is significant evidence of positive AR(1) serial correlation in the errors (even though the variables have been differenced). This means we should view the standard errors reported in equation (11.27) with some suspicion.

C12.3 (i) The test for AR(1) serial correlation gives (with 35 observations) $\hat{\rho} \approx -.110$, se($\hat{\rho}$) $\approx$ .175. The t statistic is well below one in absolute value, so there is no evidence of serial correlation in the accelerator model. If we view the test of serial correlation as a test of dynamic misspecification, it reveals no dynamic misspecification in the accelerator model.

(ii) It is worth emphasizing that, if there is little evidence of AR(1) serial correlation, there is no need to use feasible GLS (Cochrane-Orcutt or Prais-Winsten).

C12.5 (i) Using the data only through 1992 gives

$$\widehat{demwins} = \underset{(.107)}{.441} - \underset{(.354)}{.473}\, partyWH + \underset{(.205)}{.479}\, incum + \underset{(.036)}{.059}\, partyWH \cdot gnews$$
$$- \underset{(.029)}{.024}\, partyWH \cdot inf$$

$$n = 20, \quad R^2 = .437, \quad \bar{R}^2 = .287.$$

The largest *t* statistic is on *incum*, which is estimated to have a large effect on the probability of winning. But we must be careful here. *incum* is equal to 1 if a Democratic incumbent is running and –1 if a Republican incumbent is running. Similarly, *partyWH* is equal to 1 if a Democrat is currently in the White House and –1 if a Republican is currently in the White House. So, for an incumbent Democrat running, we must add the coefficients on *partyWH* and *incum* together, and this nets out to about zero.

The economic variables are less statistically significant than in equation (10.23). The *gnews* interaction has a *t* statistic of about 1.64, which is significant at the 10% level against a one-sided alternative. (Since the dependent variable is binary, this is a case where we must appeal to asymptotics. Unfortunately, we have only 20 observations.) The inflation variable has the expected sign but is not statistically significant.

(ii) There are two fitted values less than zero, and two fitted values greater than one.

(iii) Out of the 10 elections with *demwins* = 1, 8 of these are correctly predicted. Out of the 10 elections with *demwins* = 0, 7 are correctly predicted. So 15 out of 20 elections through 1992 are correctly predicted. (But, remember, we used data from these years to obtain the estimated equation.)

(iv) The explanatory variables are *partyWH* = 1, *incum* = 1, $partyWH \cdot gnews$ = 3, and $partyWH \cdot inf$ = 3.019. Therefore, for 1996,

$$\widehat{demwins} = .441 - .473 + .479 + .059(3) - .024(3.019) \approx .552.$$

Because this is above .5, we would have predicted that Clinton would win the 1996 election, as he did.

(v) The regression of $\hat{u}_t$ on $\hat{u}_{t-1}$ produces $\hat{\rho} \approx$ -.164 with a heteroskedasticity-robust standard error of about .195. (Because the LPM contains heteroskedasticity, testing for AR(1) serial correlation in an LPM generally requires a heteroskedasticity-robust test.) Therefore, there is little evidence of serial correlation in the errors. (And, if anything, it is negative.)

(vi) The heteroskedasticity-robust standard errors are given in [·] below the usual standard errors:

$$
\begin{aligned}
\widehat{demwins} = {} & \underset{\substack{(.107)\\ [.086]}}{.441} - \underset{\substack{(.354)\\ [.301]}}{.473}\, partyWH + \underset{\substack{(.205)\\ [.185]}}{.479}\, incum + \underset{\substack{(.036)\\ [.030]}}{.059}\, partyWH\cdot gnews \\
& - \underset{\substack{(.028)\\ [.019]}}{.024}\, partyWH\cdot inf
\end{aligned}
$$

$$n = 20, \quad R^2 = .437, \quad \bar{R}^2 = .287.$$

In fact, all heteroskedasticity-robust standard errors are less than the usual OLS standard errors, making each variable more significant. For example, the t statistic on $partyWH\cdot gnews$ becomes about 1.97, which is notably above 1.64. But we must remember that the standard errors in the LPM have only asymptotic justification. With only 20 observations, it is not clear we should prefer the heteroskedasticity-robust standard errors to the usual ones.

C12.7 (i) The iterated Prais-Winsten estimates are given below. The estimate of ρ is, to three decimal places, .293, which is the same as the estimate used in the final iteration of Cochrane-Orcutt:

$$
\begin{aligned}
\widehat{\log(chnimp)} = {} & \underset{(22.78)}{-37.08} + \underset{(.63)}{2.94}\log(chempi) + \underset{(.98)}{1.05}\log(gas) + \underset{(.51)}{1.13}\log(rtwex) \\
& - \underset{(.319)}{.016}\, befile6 - \underset{(.322)}{.033}\, affile6 - \underset{(.342)}{.577}\, afdec6
\end{aligned}
$$

$$n = 131, \quad R^2 = .202.$$

(ii) Not surprisingly, the C-O and P-W estimates are quite similar. To three decimal places, they use the same value of $\hat{\rho}$ (to four decimal places it is .2934 for C-O and .2932 for P-W). The only practical difference is that P-W uses the equation for $t = 1$. With $n = 131$, we hope this makes little difference.

C12.9 (i) Here are the OLS regression results:

$$
\widehat{\log(avgprc)} = \underset{(.115)}{-.073} - \underset{(.0014)}{.0040}\, t - \underset{(.1294)}{.0101}\, mon - \underset{(.1273)}{.0088}\, tues + \underset{(.1257)}{.0376}\, wed + \underset{(.1257)}{.0906}\, thurs
$$

$$n = 97, \quad R^2 = .086.$$

The test for joint significance of the day-of-the-week dummies is $F = .23$, which gives p-value = .92. So there is no evidence that the average price of fish varies systematically within a week.

(ii) The equation is

$$\widehat{\log(avgprc)} = \underset{(.190)}{-.920} - \underset{(.0014)}{.0012}\, t - \underset{(.1141)}{.0182}\, mon - \underset{(.1121)}{.0085}\, tues + \underset{(.1117)}{.0500}\, wed + \underset{(.1110)}{.1225}\, thurs$$

$$+ \underset{(.0218)}{.0909}\, wave2 + \underset{(.0208)}{.0474}\, wave3$$

$n = 97$, $R^2 = .310$.

Each of the wave variables is statistically significant, with *wave2* being the most important. Rough seas (as measured by high waves) would reduce the supply of fish (shift the supply curve back), and this would result in a price increase. One might argue that bad weather reduces the demand for fish at a market, too, but that would reduce price. If there are demand effects captured by the wave variables, they are being swamped by the supply effects.

(iii) The time trend coefficient becomes much smaller and statistically insignificant. We can use the omitted variable bias table from Chapter 3, Table 3.2 to determine what is probably going on. Without *wave2* and *wave3*, the coefficient on *t* seems to have a downward bias. Since we know the coefficients on *wave2* and *wave3* are positive, this means that the wave variables are negatively correlated with *t*. In other words, the seas were rougher, on average, at the beginning of the sample period. (You can confirm this by regressing *wave2* on *t* and *wave3* on *t*.)

(iv) The time trend and daily dummies are clearly strictly exogenous, as they are just functions of time and the calendar. Further, the height of the waves is not influenced by past unexpected changes in log(*avgprc*).

(v) We simply regress the OLS residuals on one lag, getting $\hat{\rho} = .618, \text{se}(\hat{\rho}) = .081, t_{\hat{\rho}} = 7.63$. Therefore, there is strong evidence of positive serial correlation.

(vi) The Newey-West standard errors are $\text{se}(\hat{\beta}_{wave2}) = .0234$ and $\text{se}(\hat{\beta}_{wave3}) = .0195$. Given the significant amount of AR(1) serial correlation in part (v), it is somewhat surprising that these standard errors are not much larger compared with the usual, incorrect standard errors. In fact, the Newey-West standard error for $\hat{\beta}_{wave3}$ is actually smaller than the OLS standard error.

(vii) The Prais-Winsten estimates are

$$\widehat{\log(avgprc)} = \underset{(.239)}{-.658} - \underset{(.0029)}{.0007}\, t + \underset{(.0652)}{.0099}\, mon + \underset{(.0744)}{.0025}\, tues + \underset{(.0746)}{.0624}\, wed + \underset{(.0621)}{.1174}\, thurs$$

$$+ \underset{(.0174)}{.0497}\, wave2 + \underset{(.0174)}{.0323}\, wave3$$

$n = 97$, $R^2 = .135$.

The coefficient on *wave2* drops by a nontrivial amount, but it still has a t statistic of almost 3. The coefficient on *wave3* drops by a relatively smaller amount, but its t statistic (1.86) is borderline significant. The final estimate of ρ is about .687.

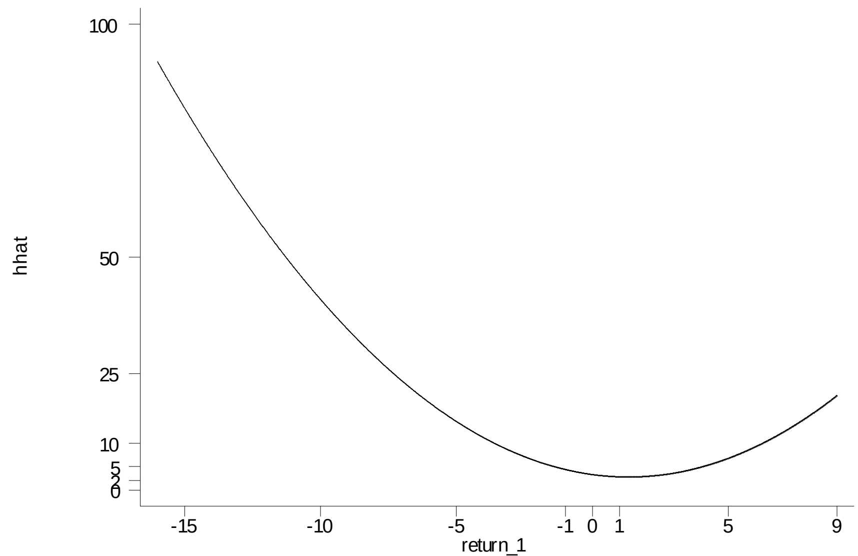
C12.11 (i) The average of $\hat{u}_t^2$ over the sample is 4.44, with the smallest value being .0000074 and the largest being 232.89.

(ii) This is the same as C12.4, part (ii):

$$\underset{}{\hat{u}_t^2} = \underset{(0.44)}{3.26} - \underset{(.196)}{.789}\, return_{t-1} + \underset{(.036)}{.297}\, return_{t-1}^2 + residual_t$$

$$n = 689,\ R^2 = .130,\ \bar{R}^2 = .128.$$

(iii) The graph of the estimated variance function is



The variance is smallest when $return_{-1}$ is about 1.33, and the variance is then about 2.74.

(iv) No. The graph in part (iii) makes this clear, as does finding that the smallest variance estimate is 2.74.

(v) The R-squared for the ARCH(1) model is .114, compared with .130 for the quadratic in $return_{-1}$. We should really compare adjusted R-squareds, because the ARCH(1) model contains only two total parameters. For the ARCH(1) model, $\bar{R}^2$ is about .112; for the model in part (ii), $\bar{R}^2$ = .128. Therefore, after adjusting for the different *df*, the quadratic in $return_{-1}$ fits better than the ARCH(1) model.

(vi) The coefficient on $\hat{u}_{t-2}^2$ is only .042, and its t statistic is barely above one (t = 1.09). Therefore, an ARCH(2) model does not seem warranted. The adjusted R-squared is about .113, so the ARCH(2) fits worse than the model estimated in part (ii).

C12.13 (i) The regression $\hat{u}_t$ on $\hat{u}_{t-1}$, $\Delta unem_t$ gives a coefficient on $\hat{u}_{t-1}$ of .073 with t = .42. Therefore, there is very little evidence of first-order serial correlation.

(ii) The simple regression $\hat{u}_t^2$ on $\Delta unem_t$ gives a slope coefficient of about .452 with t = 2.07, and so at the 5% significance level, we find that there is heteroskedasticity. The variance of the error appears to be larger when the change in unemployment is larger.

(iii) The heteroskedasticity-robust standard error is about .223, compared with the usual OLS standard error of .182. So, the robust standard error is more than 20% larger than the usual OLS one. Of course, a larger standard error leads to a wider confidence interval for β_1.

C12.15 (i) If there is in fact AR(1) serial correlation in the errors, then the OLS standard errors are invalid. With positive AR(1) serial correlation, the OLS estimates usually have a downward bias: they are too small, on average.

(ii) Using the command newey in Stata, with a lag of four, the standard error for the *lchempi* variable is about .65, which is notably higher than the usual OLS standard error, .48. Interestingly, it is only slightly higher than the usual GLS (Prais-Winsten) standard error, .63.

For *afdec*6, the Newey-West standard error is about .248, which is actually below the usual OLS standard error (.286) – something that is not ruled out by either the algebra or the theory – and well below the P-W standard error (.342). That the P-W standard error is actually larger suggests that perhaps the AR(1) model of serial correlation is not correct. However, we should be cautious with conclusions like that because we only have 131 observations.

(iii) Interestingly, as we go from g = 4 to g = 12, the standard error on *lchempi* increases to .741 while that on *afdec*6 decreases rather sharply, to .195. In any case, we can conclude that on the key policy variable, *afdec*6, the OLS standard error actually seems to be conservative.

[Instructor’s Note: This might be a good application to try the recommendation in Problem 12.7: Estimate an AR(1) model by P-W and then apply Newey-West to the transformed equation. The instructor suspects the robust standard error for P-W is below the usual GLS standard error.]