Answers for Homeworks 3, 4, and 5 Economics 421/612

Because we did questions 1 and 4 from homework 3 in class I do not include answers for these problems.

HW 3 #2: Specification errors are of the following 4 types:

1) The model takes the incorrect functional form
2) The model omits relevant variables
3) The independent variables are incorrectly measured
4) The model includes irrelevant variables.

Each of the first three errors results in biased and inconsistent estimates. The fourth error does not bias coefficient estimates, but it does result in variance estimates which are larger than would be the case in a correctly specified model.

HW 3 #3: High but imperfect multicollinearity might be suspected if individually variables are not statistically significant but they have high explanatory power as a group. That is, small t-statistics combined with high R squared might signal a problem of high multicollinearity.

HW 4 #1: Given a Durbin-Watson statistic of 1.35, one can find an approximate value for the first-order autocorrelation coefficient by solving 1.35 = 2(1-r) for the value of r; r = 1- dw/2 = .325.

To test for first-order autocorrelation, determine the upper and lower bounds of the critical DW. Given 40 observations and 2 estimated coefficients (not including the intercept) the upper and lower bounds are 1.398 and 1.198, respectively. The computed dw falls within this range, so one must conclude that the presence of autocorrelation is indeterminate.

If autocorrelation is present, least squares estimates are unbiased but no minimum variance. In fact, the least squares formula for the variance of the coefficient estimates is incorrect.

HW 4 #2: Correct for autocorrelation in the question above by using the Cochrane Orcutt procedure. Specifically, use the estimated value of the correlation coefficient to construct new variables Cons*= $Cons_t$ - $rCons_{t-1}$, Income*= $Income_t$ - $rIncome_{t-1}$, and intrate* = $intrate_t$ - $rintrate_{t-1}$. Estimate using OLS the model with these variables in place of the true values of the data. Use this procedure again, after getting a new estimate of r, to get a more finely tuned set of estimates of the parameters of the original model.

HW 4 #3: The runs test requires that one compute the number of runs of errors of one sign, the number of positive errors and the number of negative errors. Using the formulas for the expected number of runs and the variance of the number of runs, one can compute these two statistics and construct a confidence interval. In the absence of autocorrelation, the observed number of runs shouls fall within the interval 95% of the time.

The formulas are: $E(n) = (2*N_1*N_2 - N)/N$ and $Var(n) = 2*N_1*N_2*E(n)/(N*(N-1))$ where N is the number of observations, N_1 is the number of positive errors, and N_2 is the number of negative errors.

HW 5 #1: You did this one as an exercise in class. Note that you may choose either the Koyck or the Almon lag approach. Which you choose depends upon what you believe about the pattern of the lags. For example, does the effect continuously decline or does it rise then decline? In the former case, choose the Koyck approach, in the latter the Almon approach is appropriate. Koyck may result in autocorrelated errors and correlation of the equation errors with an explanatory variable; Almon leads to multicollinearity problems and difficulties testing hypotheses on the paramaters of the original model. In the Almon model, the researcher must impose a lag length, in the Koyck model that is not the case.

HW 5 #2: Adaptive expectations models lead to a Koyck specification. This means the estimating equation is:

$$M_t = \alpha\delta + \beta_1 I_t + \beta_1 (\delta-1)I_{t-1} + \delta\beta_2 r_t + (1-\delta)M_{t-1} + (\mu_t - (1-\delta)\mu_{t-1})$$

The errors are correlated over time and correlated with the lagged value of the dependent variable.

HW 5 #3: The first equation is not identified, the second and third equations are each over identified.

HW 5 #4: The approach to estimating this model is to use the logit model. The dependent variable will be the natural log of the ratio of the proportion of yes votes to the proportion of no votes in a neighborhood. The explanatory variables are the average income in the neighborhood, the proportion of households in each neighborhood which own their own home and the number of school age children in each neighborhood. Parameter estimates are interpreted as the effect of a change in the exogenous variable on the log of the odds ratio.