Exam in: ECON 3150/4150: Introductory Econometrics. ANNOTATED VERSION

Day of exam: 15 May 2013

Time of day: 14:30-17:30

This is a 3 hour school exam.

Guidelines:
In the grading, question A will count 1/3 and question B will count 2/3.

Question A (1/3)

1. Consider the econometric model specified by (1)-(3)

   (1) $$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, 2, ..., n$$

   (2) $$E(\varepsilon_i \mid X_i) = 0, \text{ for all } i$$

   (3) $$E(\varepsilon_i \varepsilon_j \mid X_i,\, X_j) = \begin{cases} \sigma^2, \text{ for } i = j \\ 0, \text{ for } i \neq j \end{cases}$$

   where the notation $\mid X_i$ means "conditional on any value that the variable X_i can take". Show that β_0 and β_1 can be interpreted as parameters of the conditional expectation function of Y_i given X_i.

   Answer note: $E(Y_i \mid X_i = x_i) = E(\beta_0 + \beta_1 X_i + \varepsilon_i \mid X_i = x_i) = \beta_0 + \beta_1 x_i$ by the use of (3). This gives the expectation of Y_i as a function of the values of X_i. The coefficients β_0 and β_1 are parameters of that function.

2. Assume that $\{X_i, Y_i\}$ $(i = 1, 2, \ldots, n)$ have identical and independent joint probability density functions $f(x_i, y_i)$. How will you extend the model specification (1)-(3) if $f(x_i, y_i)$ $(i = 1, 2, \ldots, n)$ are binormal probability densities?

   Answer note: In this case we can show that Y_i is normally distributed with expectation $E(Y_i \mid X_i = x_i)$ and variance σ^2. It follows that $\varepsilon_i = Y_i - E(Y_i \mid X_i = x_i)$ is normally distributed with expectations 0 and variance σ^2. Hence

   $$\varepsilon_i \sim N(0, \sigma^2 \mid X_i = x_i)$$

   is the extension of the model that we would typically use. (The explicit conditioning can be dropped if it is clear that the point about Y_i being conditionally normal distributed has been understood).This is the "baseline" answer. Some students may have picked up that if joint normality, then the properties of the disturbances are implications not assumptions.

3. Let $\hat{\beta}_1$ denote the OLS estimator of the parameter β_1 in (1). Make use of the *Law of iterated expectations* to show that $E(\hat{\beta}_1) = \beta_1$ under the assumptions specified in (1)-(3)

Answer note:

$$E(\hat{\beta}_1) = E\left[E(\hat{\beta}_1 \mid X)\right] = E(\beta_1) = \beta_1$$

where we first find the expectation of the function where X is treated as a parameter, namely $E(\hat{\beta}_1 \mid X) = \beta_1$.

4. Assume that a second model is formulated where $Var(\varepsilon_i)$ is assumed to be proportional to the variable Z_i. In all other respects the second model is identical to (1)-(3): What is the expression for the BLUE estimator of β_1 for the second model?

   Answer note: We interpret this as: $Var(\varepsilon_i) = \sigma^2 Z_i$, heteroskedasticity. Now the OLS estimators are not BLUE. But the OLS estimators of

$$\frac{Y_i}{\sqrt{Z_i}} = \beta_0 \frac{1}{\sqrt{Z_i}} + \beta_1 \frac{X_i}{\sqrt{Z_i}} + \varepsilon_i^*$$

   where $\varepsilon_i^* = \frac{\varepsilon_i}{\sqrt{Z_i}}$ are BLUE since ε_i^* has classical properties. These OLS estimators are called weighted-least-squares or Generalized Least Squares.

5. Consider a third model for Y_i:

(4) $$Y_i = \gamma_0 + \gamma_1 X_i + \gamma_2 Z_i + e_i \quad i = 1, 2, ..., n$$

(5) $$E(e_i \mid X_i, Z_i) = 0, \text{ for all } i$$

(6) $$E(e_i e_j \mid X_i,\ Z_j, X_i, Z_j) = \begin{cases} \tau^2, \text{ for } i = j \\ 0, \text{ for } i \neq j \end{cases}$$

   (a) Express the probability limit (*plim*) of the OLS estimator $\hat{\beta}_1$ of model (1)-(3) in terms of parameters of the model (4)-(6).

   (b) How is the precision of the OLS estimate of γ_1 affected by the degree of multi-collinearity? Explain briefly.

   (c) Can we maintain that (1)-(3) is a valid regression model for Y_i, if the true model of Y_i given X_i and Z_i is (4)-(6) with $\gamma_2 \neq 0$? Explain briefly.

   Answer note a:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2} = \gamma_1 + \gamma_2 \frac{\sum (X_i - \bar{X}) Z_i}{\sum (X_i - \bar{X})^2} + \frac{\sum (X_i - \bar{X}) \varepsilon_i}{\sum (X_i - \bar{X})^2}$$

$$\begin{aligned} \text{plim}(\hat{\beta}_1) &= \gamma_1 + \gamma_2 \text{plim}\left(\frac{\sum (X_i - \bar{X}) Z_i}{\sum (X_i - \bar{X})^2}\right) + \text{plim} \frac{\sum (X_i - \bar{X}) \varepsilon_i}{\sum (X_i - \bar{X})^2} \\ &= \gamma_1 + \gamma_2 \text{plim}\left(\frac{\sum (X_i - \bar{X}) Z_i}{\sum (X_i - \bar{X})^2}\right) \end{aligned}$$

   b. Make reference to the formula for the estimated variance of $\hat{\gamma}_1$.

   c Yes, because both models are based on valid conditioning. (1)-(3) is a valid regression model for Y_i given $X_{i.}$. (4)-(6) is a valid regression model for Y_i given $X_{i.}$and Z_i.

Question B (2/3)

In this problem, we model CPI inflation in Norway with the use of Phillips-curve models. We use annual data for the years 1981 to 2008. INF_t is the annual percentage change in the Consumer Price Index. U_t is the unemployment rate in percent.

When we estimate the static Phillips-curve model

$$INF_t = \beta_0 + \beta_1 U_t + \varepsilon_t$$

by OLS, we obtain the result

$$\widehat{INF}_t = \overbrace{\underset{(1.59)}{8.69}}^{\hat{\beta}_0} + \overbrace{\underset{(0.465)}{-1.42}}^{\hat{\beta}_1} U_t \qquad (7)$$

$$1981 - 2008 \ (T = 28),\ R^2 = 0.263 \ \hat{\sigma} = 2.86$$

$\widehat{INF}_t$ denotes the fitted values of inflation. Estimated standard-errors are in round brackets below the estimates of the intercept and the slope coefficient. Below the estimated equation we have included information about the sample size (T), the coefficient of determination (R^2) and the estimated standard error of the disturbances ($\hat{\sigma}$).

1. If you use (7) to estimate the natural rate of unemployment, U^{nat}, the answer is $\hat{U}^{nat} = 6.12$ %. The estimated covariance between $\hat{\beta}_0$ and $\hat{\beta}_1$ in (7) is -0.70. Use the "delta-method" to show that an approximate standard error for the estimated natural rate is 1.01.

   Hint: Writing the non-linear estimator of the natural rate in the form: $\hat{U}^{nat} = \frac{\hat{\theta}_1}{\hat{\theta}_2}$, the delta-method formula for the standard error $\hat{U}^{nat}$ can be expressed as:

$$Var(\hat{U}^{nat}) = Var(\left(\frac{\hat{\theta}_1}{\hat{\theta}_2}\right) \approx \left(\frac{1}{\hat{\theta}_2}\right)^2 \left[Var(\hat{\theta}_1) + \left(\frac{\hat{\theta}_1}{\hat{\theta}_2}\right)^2 Var(\hat{\theta}_2) - 2\left(\frac{\hat{\theta}_1}{\hat{\theta}_2}\right) Cov(\hat{\theta}_1, \hat{\theta}_2)\right]$$

   Answer note:

$$\hat{U}^{nat} = \frac{8.69}{1.42} = 6.1197$$

   $Var(\hat{\theta}_1) = 1.59^2 = 2.5281$, $Var(\hat{\theta}_2) = 0.465^2 = 0.21623$ and $Cov(\hat{\theta}_1, \hat{\theta}_2) = Cov(\hat{\beta}_0, -\hat{\beta}_1) = -(-0.7) = 0.7$.

   The most important element in the answer is to use this informtion correctly in the formulae. If that is right, any errors in the computations should not be penalized.

$$Var(\hat{U}^{nat}) = Var(\frac{\hat{\theta}_1}{\hat{\theta}_2}) \approx \left(\frac{1}{-1.42}\right)^2 * \left[1.59^2 + \left(\frac{8.69}{-1.42}\right)^2 * 0.465^2 - 2 * \left(\frac{8.69}{1.42}\right) * (0.70)\right]$$
$$= \left(\frac{1}{-1.42}\right)^2 * \left[2.5281 + \left(\frac{8.69}{-1.42}\right)^2 * 0.21623 - 2 * \left(\frac{8.69}{1.42}\right) * (0.70)\right]$$
$$= \left(\frac{1}{-1.42}\right)^2 * 2.0585 = 0.49593 * 2.0585 = 1.0209.$$

$$\sqrt{Var(\hat{U}^{nat})} \approx \sqrt{1.0209} = 1.0104$$

8Not asked for.: Approximate confidence interval: $[6.1197 - 2 * 1.01; 6.1197 + 2 * 1.01]$, [4.1 %: 8.14%])

2. Using the residuals $\hat{\varepsilon}_t$ ($t = 1982 - 2000$) from (7), we formulate the following auxiliary regression:

$$\hat{\varepsilon}_t = \underset{(0.085)}{0.764}\ \hat{\varepsilon}_{t-1} - \underset{(0.707)}{2.736} + \underset{(0.204)}{0.062}\ U_t \tag{8}$$
$$1982 - 2008\ (T = 27),\ R^2 = 0.778$$

Use this result to test the null-hypothesis of no first order autocorrelation in the disturbances of the static Phillips curve model.

Answer note: The null is given, but is good if the students are precise about the specification of the alternative (one-sided or two sided). Use R^2 from this auxiliary regression to calculate the LM-type test: $T * R^2$ and compare with percentiles from $\chi^2(1)$ distribution, e.g. 3.8 (0.95) ,5.02 (0.975) or 6.64 (0.99) The "degrees of freedom corrected test" uses the t-ratio 0.763645/0.08458 which is clearly significant here when compared to percentiles from $t(25)$ distribution Alternatively $F(1, 25)$ of course.

3. What is the implication of the result of the autocorrelation test for the reliability of the inference that is based on the static Phillips curve model (7)?

Answer note: If the model is in other respects correctly specified, the consequences are that the OLS estimators of the coefficients are inefficient, and that the standard errors of the estimated coefficients are wrong. They are underestimated in the case of positive autocorrelation that we have here. These consequences carry over to the estimate of the parameter called U^{nat}. Hence the inference is unreliable.

4. A near-to-hand extension of (7) is the following dynamic Phillips-curve model:

$$\widehat{INF}_t = \underset{(0.985)}{2.736} + \underset{(0.080)}{0.75879}\ INF_{t-1} - \underset{(0.238)}{0.5998}\ U_t \tag{9}$$
$$1981 - 2008\ (T = 28),\ R^2 = 0.839,\ \hat{\sigma} = 1.361$$

What is the intuitive explanation for why the estimated coefficient for U_t is markedly lower in absolute value in (9) than in (7)?

Answer note: The estimated β_1 in the static model was -1.42, which we take as proof that $Cov(INF, U)$ is negative. From (9) we see that γ_2 is however positive and therefore we expect the estimated net-coefficient γ_1 to be smaller in absolute value than the estimated gross-coefficient of -1.42.

(The algebra is not asked for, but for reference; From OLS algebra (used in A6c for example), we know that β_1 from the static PCM is a gross-coefficient. In fact:

$$\beta_1 = \gamma_1 + \gamma_2 \frac{Cov(INF, U)}{Var(U)}$$

Hence

$$\gamma_1 = \beta_1 - \gamma_2 \frac{Cov(INF, U)}{Var(U)}$$

)

5. When we test for mis-specification of (9), there is no evidence of first order autocorrelation. Neither do the standard tests for heteroskedasticity and for departures from normality indicate model mis-specification. (We do not include the results of these tests in order to save notation and time). On this basis we will use (9) to test hypotheses about the properties of the Norwegian Phillips-curve.

    (a) Explain briefly why the OLS estimators of the regression coefficients in (9) are consistent (but biased in finite samples) under the assumption of no-autocorrelation in the disturbances.

    Answer note: Under these assumptions, INF_{t-1} is a predetermined variable. Maintaining (as we can) that U_t is a regressor, then the OLS estimators are consistent. The explanation is that, although INF_{t-1} is correlated with past disturbances, INF_{t-1} is uncorrelated with the current disturbance and *all* future disturbances. (We take that stationary and causal case as given). It is the uncorrelatedness with future disturbances that drives the biases of the OLS estimatoers to zero asymptotically.

    (b) Test the hypothesis that the long-run Norwegian Phillips-curve is vertical. What is the estimated slope of the long-run Phillips-curve.

    Answer note: $(1 - 0.75879)/0.080 = 3.0151$. Reject the H_0 of a vertical long Phillips curve at the 5 % level. The estimated slope is $-0.5998/(1 - 0.75879) = -2.4866$.

    (c) When the sample period is extended to $1981 - 2012$, the residual sum of squares of the dynamic Phillips curve is $RSS = 53.6490523$. When you include four indicator variables (dummies), one for each of the four new years in the sample, the residuals sum of squares is $RSS = 46.3102852$. Test the hypothesis that there is no joint significance of these four indicator variables.

    Answer note: $F(4, 25) = [(53.6490523 - 46.310285)/46.310285]*(25/4) = 0.15847 * 25/4 = 0.99044[0.4309]$. The direct interpretation here is that we test the null that the intercept of the regression is the same on the

1981 − 2008 sample and the (extended) 1981 − 2012, sample. However, the inclusion of one year-dummy for each new observations means we are effectively testing whether *all the coefficients* of the dynamic NPC are constant when the sample is extended by the four new years. It is a bonus if the students notes this. (In fact, this is the "1-sample Chow test", which is not presented very clearly in HGL).

6. Imagine that you have a friend who is a business school student, and who has estimated a model for the change in inflation: $\Delta INF_t = INF_t - INF_t$. He has used exactly the same data for INF_t and U_t as you have used and the sample period 1981 − 2008. The equation he has estimated by OLS is:

$$\widehat{\Delta INF}_t = \underset{(1.891)}{2.736} - \underset{(0.080)}{0.24121}\ INF_{t-1} - \underset{(0.238)}{0.5998}\ U_t \tag{10}$$

$$1981 - 2008\ (T = 28),\ R^2 = 0.314,\ \hat{\sigma} = 1.361$$

Your friend cannot understand why he gets the same estimates for the intercept and for the coefficient of U_t, but a different estimate of the coefficient of INF_{t-1}. He is also worried that R^2 in (10) is lower than in your model (9). Can you resolve the puzzles for him? Explain briefly how.

Answer note: The puzzles are resolved by noting that (10) is a re-parameterization of (9). If written out in regression model form, it is evident that the disturbances are the same and that the intercept and slope coefficients are unaffected. The coefficient of INF_{t-1} is the original coefficient minus one of course. The students should know that R^2 is not invariant to such transformations when the LHS variable is affected. More concretely, in this case: We have that RSS is unchanged (can be confirmed from observing that the estimated standard errors are the same), but TSS will be smaller as a result of the differencing. Hence, R^2 is lower in this regression.

7. Your friend also suggests that you should try another estimation method than OLS. You agree that U_{t-1} may be a relevant instrumental variable for U_t, and decide to estimate the relationship between INF_t and U_t by the Methods of moments estimator, also called the Instrumental variable estimator. You decide to re-estimate the static Phillips curve in (7). The results are

$$\widehat{INF}_t = \underset{(1.120)}{10.84} - \underset{(0.560)}{2.082}\ U_t \tag{11}$$

$$1981 - 2008\ (T = 28),\ \hat{\sigma} = 2.970$$

(there is no generally accepted "coefficient of determination" for the IV-estimator, which is why R^2 does not appear). The results for the regression between U_t and U_{t-1} is:

$$\widehat{U}_t = \underset{(0.316)}{0.6188} + \underset{(0.0941)}{0.8199}\ U_{t-1} \tag{12}$$

$$1981 - 2008\ (T = 28),\ R^2 = 0.731, \hat{\sigma} = 0.606$$

Compare the results in (11) with the results in (7) and comment in particular on the differences that you would expect to find based on your knowledge about the properties of the two estimators. Why is it relevant to test for instrumental variable strength, and what is the result in this case?

Answer note: Comparison of (7) with (11) show that both coefficients have different have different point estimates when IV is used instead of OLS, not only the slope coefficient β_1. But the main theoretical point to note is that the IV standard errors are larger. The students have some background for understanding this, see HGL p 411. They also know about the importance of strong instruments (instrument relevance) and how they can test for weak-instrument using (12). The HGL book does not mention the $F > 10$ rule-of-thumb, so the general idea about testing for significance in (12) will do.