EC 327 PROBLEM SET 3
Prof. Baum, Mr. Dmitriev
11.7
(i) (5 marks) We plug the first equation into the second to get

$$y_t - y_{t-1} = \lambda(\gamma_0 + \gamma_1 x_t + e_t - y_{t-1}) + a_t$$

and rearranging,

$$y_t = \lambda\gamma_0 + (1-\lambda)y_{t-1} + \lambda\gamma_1 x_t + a_t + \lambda e_t \equiv \beta_0 + \beta_1 y_{t-1} + \beta_2 x_t + u_t$$

where $\beta_0 \equiv \lambda\gamma_0$, $\beta_1 \equiv (1-\lambda), \beta_2 \equiv \lambda\gamma_1$, and $u_t \equiv a_t + \lambda e_t$.

(ii) (10 marks) An OLS regression of y_t on y_{t-1} and x_t produces consistent, asymptotically normal estimators of the β_j. Under $E(e_t|x_t, y_{t-1}, x_{t-1}, ...) = E(a_t|x_t, y_{t-1}, x_{t-1}, ...) = 0$ it follows that $E(u_t|x_t, y_{t-1}, x_{t-1}, ...) = 0$, which means that the model is dynamically complete [see equation (11.37)]. Therefore, the errors are serially uncorrelated. If the homoskedasticity assumption $Var(u_t|x_t, y_{t-1}) = \sigma^2$ holds, then the usual standard errors, t statistics and F statistics are asymptotically valid.

(iii) (5 marks) Because $\beta_1 = (1-\lambda)$, if $\widehat{\beta}_1 = 0.7$, then $\widehat{\lambda} = 0.3$. Further, $\hat{\beta}_2 = \hat{\lambda}\hat{\gamma}_1$, or $\hat{\gamma}_1 = \hat{\beta}_2/\hat{\lambda} = \frac{0.2}{0.3} \approx 0.67$.

11.8

(i) (10 marks) Sequential exogeneity does not rule out correlation between, say, u_{t-1} and x_{ij} for any regressors $j = 1, 2, ..., k$. The differencing generally induces correlation between the differenced errors and the differenced regressors. To see why, consider a sing explanatory variable, x_t. Then $\delta u_t = u_t - u_{t-1}$ and $\delta x_t = x_t - x_{t-1}$. Under sequential exogeneity, u_t is uncorrelated with x_t and x_{t-1}, and u_{t-1} is uncorrelated with x_{t-1}. But u_{t-1} can be correlated with x_t, which means that δu_t and δx_t are generally correlated. In fact, under sequential exogeneity, it is always true that $Cov(\delta x_t, \delta u_t) = -Cov(x_t, u_{t-1})$.

(ii) (10 marks) Strict exogeneity of the regressors in the original equation is sufficient for OLS on the first-differenced equation to be consistent. Remember, stric exogeneity implies that the regressors in any time period are uncorrelated with the errors in any time period. Of course, we could make the weaker assumption: for any t, u_t is uncorrelated with $x_{t-1,j}, x_{tj}$, and $x_{t+1,j}$ for all $j = 1, 2, ..., k$. The strengthening beyond sequential exogeneity is the assumption that u_t is uncorrelated with all of next period's outcomes on all regressors. In practice, this is probably similar to just assuming strict exogeneity.

(iii) (10 marks) If we assume sequential exogeneity in a static model, the condition can be written as

$$E(y_t|z_t, z_{t-1}, z_{t-2}, ...) = E(y_t|z_t)$$

which means that, once we control for the current (contemporaneous) values of all explanatory variables, no lags matter. Although some relationships in economics are purely static, many admit distributed lag dynamics. Therefore, if one wants to capture lagged effects, it is a good idea to explore distributed lag models - whether or not we think there might be feedback from u_t to future values of the explanatory variables.

18.6

(i) (5 marks) This is given by the estimated intercept, 1.54. Remember, this is the percentage growth at an annualized rate. It is statistically different from zero since t = 1.54/.56 = 2.75.

(ii) (5 marks) 1.54 + .031(10) = 1.85. As an aside, you could obtain the standard error of this estimate by running the regression.

pcipt on pcipt-1, pcipt-2, pcipt-3, (pcspt-1 – 10),

and obtaining the standard error on the intercept.

(iii) (5 marks) Growth in the S&P 500 index has a statistically significant effect on industrial production growth – in the Granger causality sense – because the t statistic on pcspt-1 is about 2.38. The economic effect is reasonably large.

C18.2

(i)(5 marks) We run the regression

ginvpct= –.786 – .956 log(invpct-1) + .0068 t
(.170) (.198) (.0021)

+ .532 ginvpct-1 + .290 ginvpct-2
(.162) (.165)

n = 39, R2 = .437,

where ginvpct = log(invpct) – log(invpct-1). The t statistic for the augmented Dickey-Fuller unit root test is –.956/.198 –4.82, which is well below –3.96, the 1% critical value obtained from Table 18.3. Therefore, we strongly reject a unit root in log(invpct). (Incidentally, remember that the t statistics on the intercept and time trend in this estimated equation to not have approximate t distributions, although those on ginvpct-1 and ginvpct-2 do under the usual null hypothesis that the parameter is zero.)

(ii)(5 marks) When we apply the regression to log(pricet) we obtain

tgprice = –.040 – .222 log(pricet-1) +.00097 t
(.019) (.092) (.00049)

+ .328 gpricet-1 + .130 gpricet-2
(.155) (.149)

n = 39, R2 = .200,

Now the Dickey-Fuller t statistic is about –2.41, which is above –3.12, the 10% critical value from Table 18.3. [The estimated root is 1 – .222 = .778, which is much larger than for log(invpct).] We cannot reject the unit root null at a sufficiently small significance level.

(iii) (5 marks) Given the very strong evidence that log(invpct) does not contain a unit root, while log(pricet) may very well, it makes no sense to discuss cointegration between the two. If we take any nontrivial linear combination of an I(0) process (which may have a trend) and an I(1) process, the result will be an I(1) process (possibly with drift).

C18.5

(i) (5 marks) The estimated equation is

= .078 + 1.027 hy36thyt-1 - 1.021*change*hy3t - .085 *change*hy3t-1
(.028) (0.016) (0.038) (.037)
- .104 *change*hy3t-2
(.037)

n = 121, R2 = .982, se = .123.

The t statistic for H0: $B = 1$ is $(1.027 - 1)/.016 = 1.69$. We do not reject H0: $B = 1$ at the 5% level against a two-sided alternative, although we would reject at the 10% level.

(ii) (5 marks) The estimated error correction model is

hy6t= .070 + 1.259 *change*hy3t-1 - .816 (hy6t-1 − hy3t-2)
(.049) (.278) (.256)
+ .283 *change*hy3t-2 + .127 (hy6t-2 − hy3t-3)
(.272) (.256)

n = 121, R2 = .795.

Neither of the added terms is individually significant. The F test for their joint significance gives F = 1.35, p-value = .264. Therefore, we would omit these terms and stick with the error correction model estimated in (18.39).

C18.11

(i) (5 marks) For lsp500, the ADF statistic without a trend is t = -.79; with a trend, the t statistic is -2.20. This are both well above their respective 10% critical values. In addition, the estimated roots are quite close to one. For lip, the ADF statistic without a trend is -1.37 without a trend and -2.52 with a trend. Again, these are not close to rejecting even at the 10% levels, and the estimated roots are very close to one.

(ii) (5 marks) The simple regression of lsp500 on lip gives

lip500lsp= -2.402 + 1.694
(.095) (.024)

n = 558, R2 = .903

The t statistic for lip is over 70, and the R-squared is over .90. These are hallmarks of spurious regressions.

(iii) (5 marks) Using the residuals obtained in part (ii), the ADF statistic (with two lagged changes) is -1.57, and the estimated root is over .99. There is no evidence of cointegration. (The 10% critical value is -3.04.) tu

(iv) (5 marks) After adding a linear time trend to the regression from part (ii), the ADF statistic applied to the residuals is -1.88, and the estimated root is again about .99. Even with a time trend there is no evidence of cointegration.

(v) (5 marks) It appears that lsp500 and lip do not move together in the sense of cointegration, even if we allow them to have unrestricted linear time trends. The analysis does not point to a long-run equilibrium relationship.

C18.13 (i) (5 marks) The DF statistic is about -3.31, which is to the left of the 2.5% critical value (-3.12), and so, using this test, we can reject a unit root at the 2.5% level. (The estimated root is about .81.)

(ii) (5 marks) When two lagged changes are added to the regression in part (i), the t statistic becomes -1.50, and the root is larger (about .915). Now, there is little evidence against a unit root.

(iii) (5 marks) If we add a time trend to the regression in part (ii), the ADF statistic becomes -3.67, and the estimated root is about .57. The 2.5% critical value is -3.66, and so we are back to fairly convincingly rejecting a unit root.

(iv) (5 marks) The best characterization seems to be an I(0) process about a linear trend. In fact, a stable AR(3) about a linear trend is suggested by the regression in part (iii).

(v) (5 marks) For prcfatt, the ADF statistic without a trend is -4.74 (estimated root = .62) and with a time trend the statistic is -5.29 (estimated root = .54). Here, the evidence is strongly in favor of an I(0) process whether or not we include a trend.