BOSTON COLLEGE
Department of Economics
EC 228 01 Econometric Methods
Fall 2008, Prof. Baum, Ms. Phillips (tutor), Mr. Dmitriev (grader)
Problem Set 6

Problem sets should be your own work. You may work together with classmates, but if you're not figuring this out on your own, you will eventually regret it.

1. (**9.1**) (10 marks)There is functional form misspecification if $\beta_6 \neq 0$ or $\beta_7 \neq 0$, where these are the population parameters on *ceoten*^{2} and *comten*^{2}, respectively. Therefore, we test the joint significance of these variables using the R-squared form of the F-test: $F = [(.375-.353)/(1-.375)][(177-8)/2] \approx 2.97$. With 2 and ∞ df, the 10 % critical value is 2.30 awhile the 5 % critical value is 3.00. Thus, the p-value is slightly above .05, which is reasonable evidence of functional form misspecification. (Of course, whether this has a practical impact on the estimated partial effects for various levels of the explanatory variables is a different matter.)

2. (**C9.3**)

(i) (5 marks) If the grants were awarded to firms based on firm or worker characteristics, *grant* could easily be correlated with such factors that affect productivity. In the simple regression model, these are contained in u.

(ii) (5 marks) The simple regression estimates using the 1988 data are

$$\widehat{\log(scrap)} = \underset{(.241)}{.409} + \underset{(.406)}{.057}\ grant$$

$$n = 54, R^2 = .0004.$$

The coefficient on *grant* is actually positive, but not statistically different from zero.

(iii) (5 marks) When we add $\log(scrap_{87})$ to the equation, we obtain

$$\widehat{\log(scrap_{88})} = \underset{(.089)}{.021} - \underset{(.147)}{.254}\ grant_{88} + \underset{(.044)}{.831}\ \log(scrap_{87})$$

$$n = 54, R^2 = .873,$$

where the year subscripts are for clarity. The t-statistic for $H_0 : \beta_{grant} = 0$ is $-.254/.147 \approx -1.73$. We use the 5 % critical value for 40 df in Table G.2: -1.68. Because $t = -1.73 < -1.68$, we reject H_0 in favor of $H_1 : \beta_{grant} < 0$ at the 5 % level.

(iv) (5 marks) The t-statistic is $(.831 - 1)/.044 \approx -3.84$, which is a strong rejection of H_0.

(v) (5 marks) With the heteroskedasticity-robust standard error, the t-statistic for $grant_{88}$ is $-.254/.142 \approx -1.79$, so the coefficient is even more significantly less than zero when we use the heteroskedasticity-robust standard error. The t-statistic for $H_0 : \beta_{\log(scrap_{87})} = 1$ is $(.831 - 1)/.0735 \approx -2.29$, which is notably smaller than before, but it is still pretty significant.

3. (**C9.4**)

(i) (10 marks) Adding DC to the regression in equation (9.37) gives

$$\begin{array}{lcccccc} \widehat{infmort}) = & \underset{(12.42)}{23.95} & - & \underset{(1.641)}{.567} & \log(pcinc) - & \underset{(1.19)}{2.74} & \log(physic) \\ & & + & \underset{(.191)}{.629} & \log(popul) + & \underset{(1.77)}{16.03} & DC \end{array}$$

$$n = 51, R^2 = .691, \bar{R}^2 = .664.$$

The coefficient on DC means that even if there was a state that had the same per capita income, per capita physicians, and population as Washington D.C., we predict that D.C. has an infant mortality rate that is about 16 deaths per 1000 live births higher. This is a very large "D.C. effect."

(ii) (10 marks) In the regression from part (i), the intercept and all slope coefficients, along with their standard errors, are identical to those in equation (9.38), which simply excludes D.C. [Of course, equation (9.38) does not have DC in it, so we have nothing to compare with its coefficient and standard error.] Therefore, for the purposes of obtaining the effects and statistical significance of the other explanatory variables, including a dummy variable for a single observation is identical to just dropping that observation when doing the estimation. The R-squareds and adjusted R-squareds from (9.38) and the regression in part (i) are not the same. They are much larger when DC is included as an explanatory variable because we are predicting the infant mortality rate perfectly for D.C. You might want to confirm that the residual for the observation corresponding to D.C. is identically zero.

4. (**10.2**) (10 marks) We follow the hint and write

$$gGDP_{t-1} = \alpha_0 + \delta_0 int_{t-1} + \delta_1 int_{t-2} + u_{t-1}$$

and plug this into the right-hand-side of the int_t equation:

$$int_t = \gamma_0 + \gamma_1(\alpha_0 + \delta_0 int_{t-1} + \delta_1 int_{t-2} + u_{t-1} - 3) + v_t$$
$$int_t = (\gamma_0 + \gamma_1\alpha_0 - 3\gamma_1) + \gamma_1\delta_0 int_{t-1} + \gamma_1\delta_1 int_{t-2} + \gamma_1 u_{t-1} + v_t$$

Now by assumption, u_{t-1} has zero ,mean and is uncorrelated with all right-hand-side variables in the previous equation, except itself of course. So

$$Cov(int, u_{t-1}) = E(int_t \cdot u_{t-1}) = \gamma_1 E(u_{t-1}^2)$$

because $\gamma_1 > 0$. If $\sigma_u^2 = E(u_t^2)$ for all t then $Cov(int, u_{t-1}) = \gamma_1\sigma_u^2$. This violates the strict heterogeneity assumption, TS.2. While u_t is uncorrrelated with int_t, int_{t-1}, and so on, u_t is correlated with int_{t+1}

5. (**C10.7**)

(i) (5 marks) The estimated equation is

$$\widehat{gc}_t = \underset{(.0019)}{.0081} + \underset{(.067)}{.571}\ gy_t$$

$$n = 36, R^2 = .679.$$

This equation implies that if income growth increases by one percentage point, consumption growth increases by .571 percentage points. The coefficient on gy_t is very statistically significant (t-statistic $\approx$ 8.5).

(ii) (5 marks) Adding gy_{t-1} to the equation gives

$$\widehat{gc}_t = \underset{(.0023)}{.0064} + \underset{(.070)}{.552}\, gy_t + \underset{(.069)}{.096}\, gy_{t-1}$$

$$n = 35, R^2 = .695.$$

The t-statistic on gy_{t-1} is only about 1.39, so it is not significant at the usual significance levels. (It is significant at the 20 % level against a two-sided alternative.) In addition, the coefficient is not especially large. At best there is weak evidence of adjustment lags in consumption.

(iii) (5 marks) If we add $r3_t$ to the model estimated in part (i) we obtain

$$\widehat{gc}_t = \underset{(.0020)}{.0082} + \underset{(.072)}{.578}\, gy_t - \underset{(.00063)}{.00021}\, r3_t$$

$$n = 36, R^2 = .680.$$

The t-statistic on $r3_t$ is very small. The estimated coefficient is also practically small: a one-point increase in $r3_t$ reduces consumption growth by about .021 percentage points.

6. (**C10.9**)

(i) (5 marks) The sign of β_2 is fairly clear-cut: as interest rates rise, stock returns fall, so $\beta_2 < 0$. Higher interest rates imply that T-bill and bond investments are more attractive, and also signal a future slowdown in economic activity. The sign of β_1 is less clear. While economic growth can be a good thing for the stock market, it can also signal inflation, which tends to depress stock prices.

(ii) (5 marks) The estimated equation is

$$\widehat{rsp500}_t = \underset{(3.27)}{18.84} + \underset{(.129)}{.036}\, pcip_t - \underset{(.54)}{1.36}\, i3_t$$

$$n = 557, R^2 = .012.$$

A one percentage point increase in industrial production growth is predicted to increase the stock market return by .036 percentage points (a very small effect). On the other hand, a one percentage point increase in interest rates decreases the stock market return by an estimated 1.36 percentage points.

(iii) (5 marks) Only $i3$ is statistically significant with t-statistic $\approx$ -2.52.

(iv) (5 marks) The regression in part (i) has nothing directly to say about predicting stock returns because the explanatory variables are dated contemporaneously with $rsp500$. In other words, we do not know $i3_t$ before we know $rsp500_t$. What the regression in part (i) says is that a change in $i3$ is associated with a contemporaneous change in $rsp500$.