BOSTON COLLEGE
Department of Economics
EC 228 Econometrics, Prof. Baum, Mr. Barbato, Spring 2003
Problem Set 6.
Answer sheet

Problem 10.2

We follow the hint and write

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

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

$$\begin{aligned} int_t &= \gamma_o + \gamma_1(\alpha_o + \delta_0 int_{t-1} + \delta_1 int_{t-2} + \mu_{t-1} - 3) + \upsilon_t \\ &= (\gamma_0 + \gamma_1\alpha_0 - 3\gamma_1) + \gamma_1\delta_0 int_{t-1} + \gamma_1\delta_1 int_{t-2} + \gamma_1\mu_{t-1} + \upsilon_t \end{aligned}$$

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

$$Cov(int, \mu_{t-1}) = E(int_t \bullet \mu_{t-1}) = \gamma_1 E(\mu_{t-1}^2) > 0$$

because $\gamma_1 > 0$. If $\sigma_\mu^2 = E(\mu_t^2)$ for all t then $Cov(int, \mu_{t-1}) = \gamma_1\sigma_\mu^2$. This violates the strict exogeneity assumption, TS.2. While μ_t is uncorrelated with int_t, int_{t-1}, and so on, μ_t is correlated with $int_{t+1.}$

Problem 10.5

The functional form was not specified, but a reasonable one is

$$\log(hsestrts_t) = \alpha_0 + \alpha_1 t + \delta_1 Q2_t + \delta_2 Q3_t + \delta_3 Q4_t + \beta_1 int_t + \beta_2 \log(pcinc_t) + \mu_t,$$

Where $Q2_t, Q3_t,$ and $Q4_t$ are quarterly dummy variables (the omitted quarter is the first) and the other variables are self-explanatory. The inclusion of the linear time trend allows the dependent variable and $\log(pcinc_t)$ to trend over time (int_t robably does not contain a trend), and the quarterly dummies allow all variables to display seasonality. The β_2 is an elasticity and $100 \bullet \beta_1$ is a semi-elasticity.

Problem 10.8

(i) Adding a linear time trend to (10.22) gives

$$\begin{aligned} \widehat{\log(chnimp)} =& \underset{(20.78)}{-2.37} - \underset{(1.240)}{.686} \log(chempi) + \underset{(.876)}{.466} \log(gas) + \underset{(.472)}{.078} \log(rtwex) + \underset{(.251)}{.090}\ befile6 \\ & + \underset{(.257)}{.097}\ affile6 - \underset{(.282)}{.351}\ afdec6 + \underset{(.004)}{.013}\ t \\ & n = 131,\ R^2 = .362,\ \overline{R^2} = .325 \end{aligned}$$

Only the trend is statistically significant. In fact, in addition to the time trend, which has a t statistic over three, only *afdec*6 has a t statistic bigger than one in absolute value. Accounting for a linear trend has important effects on the estimates.

(ii) The F statistic for joint significance of all variables except the trend and intercept, of course, is about .54. The *df* in the F distribution are 6 and 123. The p-value is about .78, and so the explanatory variables other than the time trend are jointly very insignificant. We would have to conclude that once a positive linear trend is allowed for, nothing else helps to explain log(*chnimp*). This is a problem for the original event study analysis.

(iii) Nothing of importance changes. In fact, the p-value for the test of joint significance of all variables except the trend and monthly dummies is about .79. The 11 monthly dummies themselves are not jointly significant: p-value $\approx$ .59.

Problem 10.9

Adding log(*prgnp*) to equation (10.38) gives

$$\widehat{\log(prepop_t)} = \underset{(1.26)}{-6.66} - \underset{(.040)}{.212} \log(\min cov_t) + \underset{(.222)}{.486} \log(usgnp_t) + \underset{(.080)}{.285} \log(prgnp_t) - \underset{(.005)}{.027}\ t$$

$$n = 38,\ R^2 = .889,\ \overline{R^2} = .876$$

The coefficient on log($prgnp_t$) is very statistically significant (t statistic ≈ 3.56). Because the dependent and independent variable are in logs, the estimated elasticity of *prepop* with respect to *prgnp* is .285. Including log(*prgnp*) actually increases the size of the minimum wage effect: the estimated elasticity of *prepop* with respect to *mincov* is now -.212, as compared with -.169 in equation (10.38).

Problem 10.13

(i) 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 ≈ 8.5).

(ii) 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 signiicance 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 lags in consumption.

(iii) 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.

Problem 10.17

(i) The variable *beltlaw* becomes one at $t = 61$, which correspnods to January, 1986. The variable *spdlaw* goes from zero to one at $t = 77$, which corresponds to May, 1987.

(ii) The OLS regression gives

$$\widehat{\log(totacc)} = \underset{(.019)}{10.469} + \underset{(.00016)}{.00275}\ t - \underset{(.0244)}{.0427}\ feb + \underset{(.0244)}{.0798}\ mar + \underset{(.0245)}{.0185}\ apr + \underset{(.0245)}{.0321}\ may + \underset{(.0245)}{.0202}\ jun + \underset{(.0245)}{.0376}\ jul$$

$$+ \underset{(.0245)}{.0540}\ aug + \underset{(.0245)}{.0424}\ sep + \underset{(.0245)}{.0821}\ oct + \underset{(.0245)}{.0713}\ nov + \underset{(.0245)}{.0962}\ dec$$

$$n = 108,\ R^2 = .797$$

When multiplied by 100, the coefficient on t gives roughly the average monthly percentage growth in *totacc*, ignoring seasonal factors. In other words, once seasonality is eliminated, *totacc* grew by about .275% per month over this period, or, $12(.275) = 3.3\%$ at an annual rate.

There is pretty clear evidence of seasonality. Only February has a lower number of total accidents than the base month, January. The peak is in December:roughly, there are 9.6% more accidents in December than January in the average year. The F statistic for joint significance of the monthly dummies is F = 5.15. With 11 and 95 *df*, this gives a p-value essentially equal to zero.

(iii) I will report only the coefficients on the new variables:

$$\widehat{\log(totacc)} = \underset{(.063)}{10.640} + ... + \underset{(.00378)}{.00333}\ wkends - \underset{(.0034)}{.0212}\ unem - \underset{(.0126)}{.0538}\ spdlaw + \underset{(.0142)}{.0954}\ beltlaw$$

$$n = 108,\ R^2 = .910$$

The negative coefficient on *unem* makes sense if we view *unem* as a measure of economic activity . As economic activity increases - *unem* decreases - we expect more driving, and therefore more accidents. The estimate is that a one percentage point increase in the unemployment rate reduces total accidents by about 2.1%. A better economy does have costs in terms of traffic accidents.

(iv) At least initially, the coefficients on *spdlaw* and *beltlaw* are not what we might expect. The coefficient on *spdlaw* implies that accidents dropped by about 5.4% *after* the highway speed limit was increased from 55 to 65 miles per hour. There are at least a couple of possible explanations. One is that people become safer drivers after the increased speed limiting, recognizing that they must be more cautious. It could also be that some other change - other than the increased speed limit or the relatively new seat belt law - caused a lower total number of accidents, and we have not properly accounted for this change.

The coefficient on *beltlaw* also seems counterintuitive at first. But, perhaps people became less cautious once they were forced to wear seatbelts.

(v) The average of *prcfat* is about .886, which means, on average, slightly less than one percent of all accidents result in a fatality. The highest value of *prcfat* is 1.217, which means there was one month where 1.2% of all accidents resulted in a fatality.

(vi) As in part (iii), I do not report the coefficients on the time trend and seasonal dummy variables:

$$\widehat{prcfat} = \underset{(.103)}{1.030} + ... + \underset{(.00616)}{.00063}\ wkends - \underset{(.0055)}{.0154}\ unem + \underset{(.0206)}{.0671}\ spdlaw - \underset{(.0232)}{.0295}\ beltlaw$$

$$n = 108,\ R^2 = .717$$

Higher speed limits are estimated to increase the percent of fatal accidents, by .067 percentage points. This is a statistically significant effect. The new seat belt law is estimated to decrease the percent of fatal accidents by about .03, but the two-sided *p*-value is about .21.

Interestingly, increases economic activity also increases the percent of fatal accidents. This may be because more commercial trucks are on the roads, and these probably increase the chance that an accident results in a fatality.