EE325 Section 2 HW4 answers ☺

8.31

a) A priori, one would **expect a positive relationship between CM and TFR**, for the larger the number children born to a woman the greater is the likelihood of increased mortality due to health and other reasons.

b) The coefficients of PGNP are not very different, but that of FLR look different. To see if the difference is real, we can use the t test. Suppose we use Eq. (1) and hypothesize that the true coefficient of PGNP is –1.7680. We can now use the t test as follows:

$$\mathrm{t} = \frac{-2.2316 - (-1.7680)}{0.2099} = -2.2086$$

This t value exceeds 2 in absolute terms, so can reject the hypothesis that the true coefficient is –1.7680. Note here we have used the 2-t rule of thumb since the number of observations is reasonably high.

c) We can treat model (1) as the restricted version of model (2). Hence, we can use the R^2version of the F test given in (8.7.10), since the dependent variables in the two models are the same. The resulting F statistic is as follows:

$$F = \frac{(0.7474 - 0.7077)/1}{(1 - 0.7474)/(64 - 4)} = 9.4523$$

Under the standard assumptions, this has the F distribution with 1 and 60 df in the numerator and denominator, respectively. The 1% critical F for these dfs is 7.08. Since the computed F exceeds this critical value, we can reject the restricted model (1) and conclude that the TFR variable belongs in the model.

d) Recall that
$F_{1,k} = t_k^2$. Therefore, taking the (positive) square root of the F value given in (c) above, we find:

$$t = \sqrt{9.4523} = 3.0744, approx.$$

Therefore, under the null hypothesis that the true value of coefficient of TFR in model (2) is zero, we can obtain the standard error of the estimated TFR coefficient by dividing the estimated coefficient by the preceding t value, which gives

$$\mathrm{se} = \frac{12.8686}{3.0744} = 4.1857, approx.$$

9.2

a) As per economic theory, **the coefficients of X_2, X_5 are expected to be positive and that of X_3, X_8, and X_9 are expected to be negative**. **The coefficient of X_4 could be positive or negative, depending on wife's age and the number of children**. Perhaps an interactive dummy of age and children under 6 or between 6 and 13 might shed more light on the relationship between age and desired hours of work.

b) Holding all other factors constant, one would expect that desired hours of work would be higher than the (common) intercept value of 1,286 hours. This coefficient, however, has a negative sign. However, since it is not statistically significant, we can say little about the impact of X_6 on (average) Y. As for X_7, its coefficient is expected to be positive, which it is. Not only that, it Type equation here.is statistically significant, as the t value is quite high.

c) Perhaps, this is due to collinearity between age and education, as well as collinearity of these variables with number of children. Also, notice that the model does not include years of schooling completed by husband.

9.3

a) The relationship between the two variables **is expected to be negative**, for **if the unemployment rate is high, indicating slackness in the labor market, employers are less likely to advertise job vacancies**.

b) It is 3.8998 (=2.7491+1.1507). Since the dummy coefficient is statistically significant, the unemployment rate post 1966 4th quarter is statistically higher than it was in the pre-1964 4th quarter period.

c) Since the differential dummy coefficient is just about significant at the 5% level, we could say that the slopes of the regression function in the two periods are different.

9.21

a) Since the dummy enters in the log form, and since the log of zero is undefined, by redefining the dummy as 1 and 10, we can obtain logs of these numbers.
b) The regression results are as follows (*t* values in parentheses):

$$\ln(Savings)_t = -0.1589 + 0.6695\ln Income_t - 0.00029\ln D_t$$
$$t = (-0.2074) \quad (6.2362) \qquad (-0.00505)$$
$$R^2 = 0.8780$$

Since **the dummy coefficient is not statistically significant**, for all practical purposes the two intercept terms are the same. The interpretation of the intercept coefficient of –0.1589 is that it represents the value of log of savings when all the regressors take a value of zero. Taking the antilog of this value, we obtain the value of 0.8531 (billions of dollars).

It may be interesting to compare the preceding regression results with the following results, which allow for the interaction effect:

$$\ln(Savings)_t = -2.0048 + 0.9288\ln Income_t + 2.3278\ln D_t - 0.2985(\ln D_t * Income_t)$$
$$t = (-0.26528) \quad (8.7596) \qquad (3.9696) \qquad (-3.9820)$$
$$R^2 = 0.9291$$

Now you get an entirely different picture, for the differential intercept and slope dummies are both significant. **For the 1982-1995 period, the MPS (marginal propensity to save) is 0.6303, whereas for the earlier period it is 0.9288.** By the same token, **the intercept term for the first period is negative but it is positive for the second period.**

9.22

a) b) and c) The "slope" coefficients are in fact differential intercepts, with first quarter as the reference quarter. Only the 4th quarter dummy for washing machines is statistically significantly different from the first quarter; suggesting that only washing machines exhibit some type of seasonality.

Since there is no statistically visible seasonality in dishwasher and disposers sales, there is no need for deseasonalizing the data. For washing machines, the residuals from that regression will represent deseasonalized time series.

10.3

a) Although the numerical values of the intercept and the slope coefficients of PGNP and FLR have changed, their signs have not. Also, these variables are still statistically significant. These changes are due to the addition of the TFR variable, suggesting that there may be some collinearity among the regressors.

b) Since the t value of the TFR coefficient is very significant (the p value is only .0032), it seems TFR belongs in the model. The positive sign of this coefficient also makes sense in that the larger the number of children born to a woman, the greater the chances of increased child mortality.

c) This is one of those "happy" occurrences where despite possible collinearity, the individual coefficients are still statistically significant.

10.28

Since the R2 values in all the auxiliary regressions are uniformly high, it seems the data suffer from the multicollinearity problem.

There are probably too many substitute good variables in the equation. One could use only the composite substitute good price, price of chicken and disposable income as regressors. This was already done in Problem 7.19.

Creating a relative price variable, say the price of beef divided by the price of pork, might alleviate the collinearity problem.