# Problem Set 2: Model specification & Dynamics

## 1) Diagnostic Tests of the Model

Recall the model you estimated yesterday:

$$\log(\text{CO2})_t = \hat{\beta}_1 + \hat{\beta}_2 \log(\text{RGDP})_t \tag{1}$$

1. Comment on the diagnostic tests of the model - is the model well specified?

## 2) Expanding the Model

The environmental Kuzents curve suggests that pollution follows a U-pattern as income increases. This is not well captured by our simple linear model. Consider the following extension of the model where we include a squared term of real GDP per capita.

$$\log(\text{CO2})_t = \beta_1 + \beta_2 \log(\text{RGDP})_t + \beta_3 \log(\text{RGDP})_t^2 + \epsilon_t \tag{2}$$

Table 1: Example Regression Results

| **Variable** | **Estimated Coefficient** |
|---:|---:|
| Constant | 0.5 (0.05)** |
| L(RGDP) | ... |
| L(RGDP)$^2$ | ... |
| | |
| $R^2$ | 0.86 |
| T | 46 |
| | |
| AR 1-2 Test | 104.3 [0.00]** |
| ARCH 1-1 Test | 10.5 [0.00]** |
| Normality Test | ... |
| Hetero Test | ... |
| RESET Test | ... |

1. What is now the marginal effect of log(RGDP per capita) on log(CO2 per capita)? What if $\beta_2$ is positive and $\beta_3$ is negative? (Hint: differentiate equation (2) with respect to log(RGDP)).

2. Use the *Algebra* menu to construct the variable $(\log(\text{RGDP}))^2$

3. Estimate model (2) and comment on the results. Report the estimated results in a table showing the coefficients (with standard errors in parentheses), together with the number of observations (T), $R^2$, and diagnostic tests (see Table 1 for an example of how results should be presented). In particular comment on the signs of $\hat{\beta}_2$, $\hat{\beta}_3$ and what this means in the context of the environmental Kuznets curve.

4. Are $\log(\text{RGDP})_t$ and $\log(\text{RGDP})_t^2$ individually significant? Are they jointly significant? (hint: construct an F-test for joint-significance through the *Test* menu by selecting *Exclusion Restrictions*)

## 3) Time series Properties of the Data

Now we turn to the time series properties of the data:

1. Plot the Auto-correlation and Partial auto-correlation functions of log(CO2) and log(RGDP) and comment on the plots. What lag length would you suggest is optimal?

2. Estimate an auto-regressive model with 1-lag (AR-1 models) for log(CO2) report the results in equation format and comment on the output.

$$\log(\text{CO2})_t = \alpha_0 + \alpha_1 \log(\text{CO2})_{t-1} + \epsilon_t \tag{3}$$