OECD
DATA SET

REGRESSION EXERCISES

c.dougherty@lse.ac.uk
March 2002

TABLE OF CONTENTS

DOWNLOADING THE DATA SET

You will find the OECD data set in three different formats at the Introduction to Econometrics web site, URL http://econ.lse.ac.uk/ie/. Stata format for Stata users, EViews format for EViews users, and comma-delimited ASCII (text format) for everybody else. It is also available in Excel format by email. To download, click on the filename and follow the instructions in the dialogue box. Note that you may wish to download to a directory other than that specified as the default in the dialogue box. If you are using the Stata file, the file will be called OECD.dta and it will be ready for use. If you are using the EViews file, it may download as OECD_wf1.bin. It should have extension wf1, not bin, so rename it as OECD.wf1. To do this, go to My Computer, browse until you find the downloaded file, click on File and then Rename. You will now be able to delete the .bin extension and to replace _wf1 with .wf1.

1. DESCRIPTION OF THE DATA SET

The data set has been compiled from various issues of OECD *Economic Outlook* over the period 1990-2000. All the data are in the form of annual averages for the period 1988-1997. Missing values have been coded -9999, except in the Stata data set, where the missing value code has been used.

Variables

WAGES	average annual rate of growth of nominal wages
PRICES	average annual rate of growth of prices
GDP	average annual rate of growth of real GDP
EMPLOY	average annual rate of growth of employment
MONEY1	average annual rate of growth of money and quasi-money
MONEY2	average annual rate of growth of money and quasi-money, alternative data compiled from individual country tables in the IMF *International Financial Statistics Yearbook 2000*.
UNEMPLOY	average rate of unemployment

2. EXERCISES

Note: These two exercises merely replicate regressions in the text and are intended as a starting point for investigating possible relationships among the variables in the data set. One motive for constructing this data set was to provide an opportunity to investigate macroeconomic relationships using cross-section data.

In regressions involving *EMPLOY*, the observation for Mexico should be excluded because the figure has been distorted by special circumstances. With the NAFTA agreement, US firms started moving their manufacturing plants to low-wage Mexico, recruiting workers many of whom had previously been employed in the informal sector. The official employment statistics, collected by social security, measure only employment in the formal sector and therefore grossly overestimate the net increase in employment. In 1997 alone, employment increased by 13.3 percent according to the official figures, clearly nonsensical. Over the whole period, the average employment growth rate was greater than the average GDP growth rate, also nonsensical.

Exercise 1 A simple regression

How does the growth rate of employment vary with the growth rate of GDP?

Regress *EMPLOY* on *GDP* and perform appropriate statistical tests. [Exercises 2.1, 3.13, and 3.20.]

Exercise 2 Nonlinear regression specifications

Is the relationship between employment growth and GDP growth nonlinear?

Plot a scatter diagram of *EMPLOY* and *GDP*, and investigate whether a nonlinear specification might be superior to a linear one. [Exercise 5.5.]

3. INSTRUCTIONS FOR DOING THE EXERCISES USING STATA OR EVIEWS

Stata

Exercise 1:
```
reg EMPLOY GDP
```
Exercise 2:
```
gen GDPSQ = GDP*GDP
reg EMPLOY GDP GDPSQ
gen LGGDP=ln(GDP)
reg EMPLOY LGGDP
gen GDPREC = 1/GDP
reg EMPLOY GDPREC
```
(These are just some suggestions. The present writer was unable to find out how to run the final EViews specification using Stata.)

EViews (from the command line)

Exercise 1:
```
LS EMPLOY C GDP
```
Exercise 2:
```
GENR GDPSQ = GDP*GDP
LS EMPLOY C GDP GDPSQ
GENR LGGDP=ln(GDP)
LS EMPLOY C LGGDP
GENR GDPREC = 1/GDP
LS EMPLOY C GDPREC
LS EMPLOY=C(1)+C(2)/(GDP+C(3))
```