

Gretl Workshop 4A

Multiple Regression

Cross country growth regressions
And tests for poverty trap

Data entry, descriptives

- Start gretl;
- Click on File/Open data/sample file and choose **gretl tab** and **mrw** data file.
- Look at variable definitions
- Select Data/Select all, then View/Summary statistics
 - What was average rate of growth in gdp?
 - Report lowest and median values for gdp60

- Summary Statistics, using the observations 1 - 121
- (missing values were skipped)

Variable	MEAN	MEDIAN	MIN	MAX
gdp60	3681.8	1962.0	383.00	77881
gdp85	5683.3	3484.5	412.00	25635
gdpgrow	4.0940	3.9000	-0.90000	9.2000
popgrow	2.2794	2.4000	0.30000	6.8000

Variable	S.D.	C.V.	SKEW	EXCSKURT
gdp60	7492.9	2.0351	8.5155	81.404
gdp85	5688.7	1.0010	1.2304	0.60692
gdpgrow	1.8915	0.46201	0.28524	0.0050989
popgrow	0.99875	0.43816	0.45339	2.7397

Multiple regressions

- From main window choose Model/ordinary least squares.
- Choose gdpgrow as dependent variable
- Choose gdp60, popgrow, inv, school as explanatory variables
- Test hypothesis that coefficient on popgrow equals 1.

Restricted regression

- In main window choose *Add/define new variable* and enter equation

$$\text{gdppcgrow} = \text{gdpgrow} - \text{popgrow}$$

- Since we cannot reject that coefficient on *popgrow* equals 1, impose this restriction in new regression:
 - dependent variable as *gdppcgrow*
 - omit *popgrow* as explanatory variable

- **Model 2: OLS estimates using 105 observations from 1-121**
- **Missing or incomplete observations dropped: 16**
- **Dependent variable: gdpgrow**

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
const	-0.528668	0.600139	-0.881	0.38048
gdp60	-9.20330E-05	2.14791E-05	-4.285	0.00004 ***
popgrow	1.03244	0.163820	6.302	<0.00001 ***
inv	0.120069	0.0245230	4.896	<0.00001 ***
school	0.0916699	0.0575799	1.592	0.11453

- **Mean of dependent variable = 4.0581**
- **Standard deviation of dep. var. = 1.87928**
- **Sum of squared residuals = 217.07**
- **Standard error of residuals = 1.47333**
- **Unadjusted R-squared = 0.409005**
- **Adjusted R-squared = 0.385366**
- **F-statistic (4, 100) = 17.3016 (p-value < 0.00001)**
-

name	Descriptive label
	auto-generated constant
	dummy: = 1 for non-oil producers
	dummy: = 1 for countries with fairly good data
	dummy: = 1 for OECD countries
	GDP per adult, 1960
	GDP per adult, 1985
	Average growth of GDP, 1960 to 1985
	Average growth of working-age population, 1960-85
	Average ratio of investment to GDP, percent
	Percent of the working-age population in secondary school

etl: add var

Enter formula for new variable:

grow=gdpgrow-popgrow

Help Cancel OK

Restricted regression

- **What is the meaning of gdppcgrow?**
- **Regress gdppcgrow on gdp60, inv, school**
- **Test the significance of the coefficient on inv. Is a one-tailed test appropriate? Why or why not? What is the smallest significance level to reject null hypothesis?**
- **Is the negative coefficient on gdp60 reasonable? What is the theory behind this? Minimize this window.**

- **Model 3: OLS estimates using 105 observations from 1-121**
- **Missing or incomplete observations dropped: 16**
- **Dependent variable: gdppcgrow**

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
const	-0.433339	0.356623	-1.215	0.22715
gdp60	-9.04908E-05	1.99219E-05	-4.542	0.00002 ***
inv	0.118964	0.0237665	5.006	<0.00001 ***
school	0.0901343	0.0567833	1.587	0.11556

- **Mean of dependent variable = 1.79048**
- **Standard deviation of dep. var. = 1.90362**
- **Sum of squared residuals = 217.155**
- **Standard error of residuals = 1.4663**
- **Unadjusted R-squared = 0.423794**
- **Adjusted R-squared = 0.406679**
- **F-statistic (3, 101) = 24.7616 (p-value < 0.00001)**

Interpretations

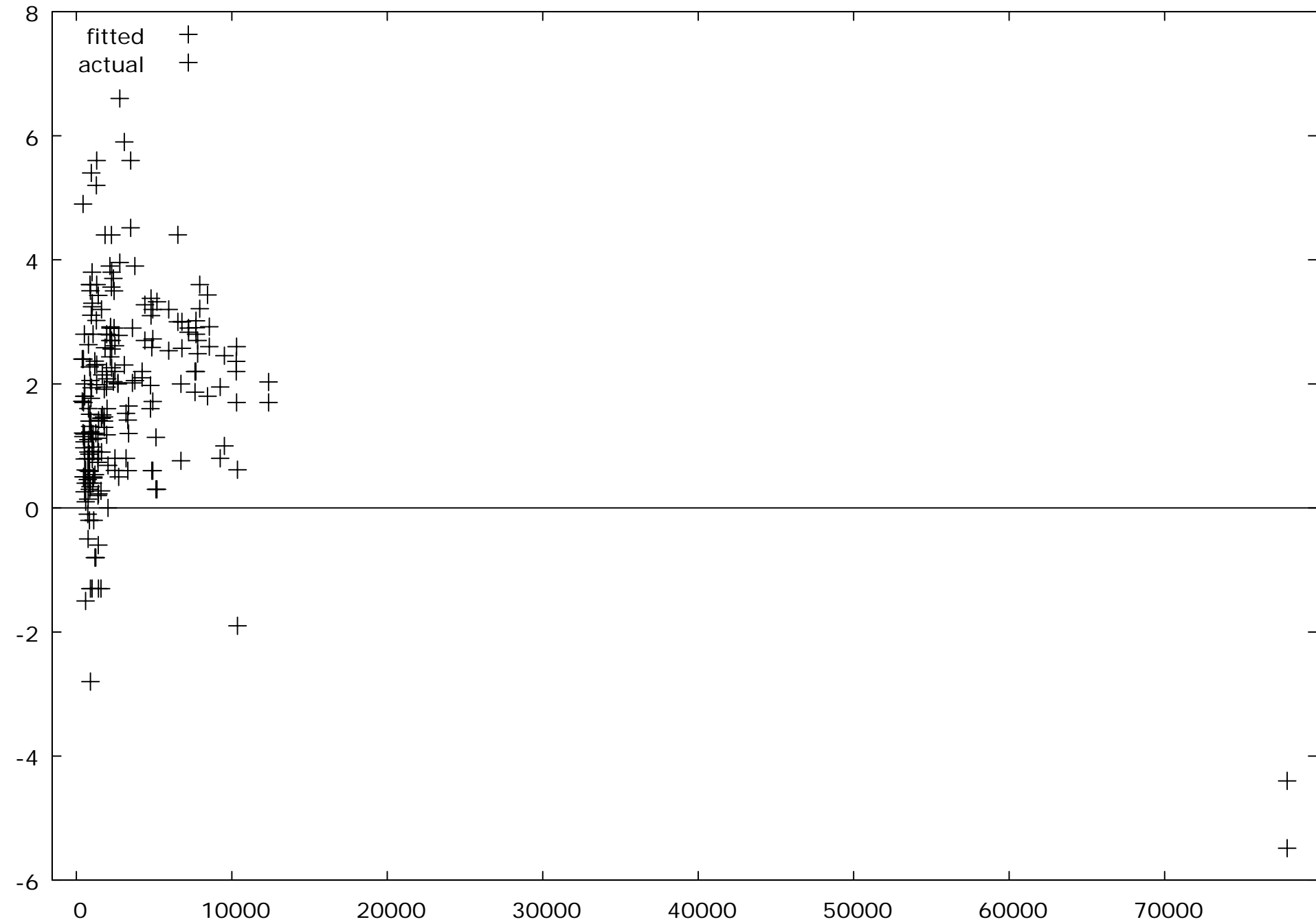
- Coefficient on inv. If inv increases by 10 percentage points, what happens to the expected average rate of growth in per capita GDP?
- Exercise. What is the difference between the lowest and median values for gdp60? What is the difference in their expected growth rates?

Outlier Problem

Examine graph of actual, fitted against gdp60. (From equation output window choose Graphs/Fitted actual plot/Against gdp60)

- Do you see any problem here?
- In main gretl window double click on gdp60. Which country is the outlier?

Actual and fitted gdp60 versus gdp60



Restrict sample

- Eliminate this outlier by restricting the sample. Choose
sample/restrict based on criteria

And enter `gdp60<60000`

To define a new sample. Now repeat the regression.

Restricted Sample

- **Model 4: OLS estimates using 104 observations from 1-115**
- **Missing or incomplete observations dropped: 11**
- **Dependent variable: gdpccgrows**

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
• const	-0.536432	0.347512	-1.544	0.12584
• gdp60	-0.000267769	6.70994E-05	-3.991	0.00013 ***
• inv	0.131662	0.0234806	5.607	<0.00001 ***
• school	0.165775	0.0614666	2.697	0.00821 ***

- **Mean of dependent variable = 1.85**
- **Standard deviation of dep. var. = 1.81199**
- **Sum of squared residuals = 201.798**
- **Standard error of residuals = 1.42056**
- **Unadjusted R-squared = 0.403282**
- **Adjusted R-squared = 0.38538**
- **F-statistic (3, 100) = 22.5278 (p-value < 0.00001)**
-

Test for Nonlinear Convergence

- Create `sq_gdp60` by highlighting `gdp60` and choosing:
 - Add/squares of selected variables.
- Add `sq_gdp60` to the regression.
- Does the result fit the hypothesis?

Model 3: OLS estimates using 104 observations from 1-115

Missing or incomplete observations dropped: 11

Dependent variable: gdp60

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
const	-0.468065	0.370435	-1.264	0.20936
gdp60	-0.000385639	0.000225655	-1.709	0.09059 *
sq_gdp60	1.06179E-08	1.94013E-08	0.547	0.58542
inv	0.134296	0.0240499	5.584	<0.00001 ***
school	0.176844	0.0649144	2.724	0.00762 ***

Unadjusted R-squared = 0.405082

Poverty Trap

- Hypothesis: poorest countries are in poverty trap and have low growth rates. Medium income countries are on high growth path as they move towards convergence in per capita income levels with slow-growing rich countries.
- Hypothesis predicts an inverted U-shaped relation between `gdppcgrow` and `gdp60`.

Poverty Trap

- Insignificance of sq_gdp60 inconsistent with hypothesis. Relation is not quadratic.
- Student's interpretation: poorest countries may grow slowly, due to low investment in education, physical capital; we already control for these variables in the model.
- Test this conjecture by excluding all other variables except $gdp60$ and sq_gdp60 .

Model 2: OLS estimates using 104 observations from 1-115

Missing or incomplete observations dropped: 11

Dependent variable: gdp_pcgrow

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
const	1.04238	0.372920	2.795	0.00621 ***
gdp60	0.000564069	0.000221316	2.549	0.01232 **
sq_gdp60	-5.20631E-08	2.14451E-08	-2.428	0.01696 **

Unadjusted R-squared = 0.0604747

Poverty Trap

- Compare these two models to test conjecture.
- Last model shows inverted U-shape, consistent with poverty trap hypothesis.
- Control for other variables and poverty trap disappears.
- Exercise. Use the estimates in the last regression to find the level of gdp60 at which gdppcgrow is maximized.

Lessons

- Single outlier can spoil estimates
- Interpretation of regression coefficients: effect of one variable on dependent variable, **controlling for all other variables in the model.**