

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



How to create residual scattergram using
Stata

Chair of Economic History

2019

Manual: Step by Step

- I. First, run a regression of the dependent variable on all explanatory variables except the one which interests us most, and save the residuals with a Stata® option predict variable name, *resid*.
- II. Run the second regression with the dependent variable (is in this case the explanatory variable which interests us most) and explanatory variables in this regression are the other explanatory variables. For example, if one is studying the growth effects of development aid, he/she would run a regression of GDP growth on all other variables which could cause GDP growth such as assassinations, health environment, institutional quality and so on.
- III. Save the residuals.
- IV. Take the main explanatory variable of interest (which in this case is development aid) and regress them also on assassinations, health environment, and institutional quality.
- V. Save these residuals as well.
- VI. Then 2 sets of residuals are shown in the scatter gram. The basic idea of such a residual scatter gram is that we can compare the dependent variable from which the influence of all the other variables is removed with the main explanatory variable from which all the other influences are removed.

Here is an example of another residual scatter gram:

```
use "e:\a\re\re121221.dta", clear
xi: regress abccall popdens2372_1000 capital prot i.co if co=="ES" |
co=="IT" | co=="HU" | co=="RU", r
predict numres if e(sample), residuals
xi: regress lnratiolargeh_agricl popdens2372_1000 capital prot i.co if
co=="ES" | co=="IT" | co=="HU" | co=="RU", r
predict ineqres if e(sample), residuals
scatter numres ineqres, yt("Resid. numeracy") xt("Resid. land ineq.")
```