# Useful software for econometric beginners

Šárka Lejnarová, Adéla Rácková,

University of Economics, Prague, Department of econometrics, Nám. W. Churchilla 4,
Praha 3, 130 67, Czech Republic
Sarka.Lejnarova@vse.cz, Adela.Rackova@vse.cz

**Abstract.** In this paper we analyzed econometric software suitable for basic econometric lessons. We were limited by software availability at the University of Economics in Prague and that is why the following econometric software was selected for our analysis. We discussed advantages and disadvantages of GiveWin, Gretl and SAS Enterprise Guide and we evaluated them according to their convenience for econometric beginners. Using multiple criteria decision methods, Gretl was chosen as the most suitable econometric software for basic econometric lessons.

**Keywords:** Gretl, SAS, GiveWin, teaching, basic econometrics

## 1  Introduction

The importance of the science we all know as econometrics has risen during last years. Econometric lectures were involved in many majors and nowadays not only students with econometric majors but even students with major of economics, finance, national economy, accounting, marketing and management have to deal with this science. Such an expansion of lectures among students who are highly educated in economics but have very basic knowledge of statistics, math and relevant software programs causes many troubles during econometric lessons. The main goal of basic econometric lessons is to introduce students into the ideas of econometric and statistical analysis on real economic dates. This is crucial for the possibility to use a simple econometric analysis in students' final thesis to support their research.

After few lessons we recognized that usage of relevant software during lectures is one of the foundation-stones in the successful education of econometrics. It is not a problem to choose one program and present it to our students but the question is which software is the most appropriate. Majority of students is worried about the software and they concentrate on how to be familiar with the program and the fundamental of the lecture – in the meaning of studying econometrics – is hidden behind this program. That is why we would like to compare used and accessible programs at our university and decide which one is the best or the most acceptable.

At the University of Economics in Prague, we have the possibility to teach econometrics in three programs – in GiveWin, SAS and Gretl. That is the reason why

we chose these three programs for our comparison in this thesis. It is not our goal to make a comparison at the whole market of accessible econometric programs because this market is very wide and the fact is that not all programs are appropriate for students starting with the study of econometrics.

The confrontation of these programs was accomplished on the basis of several different criteria. One of them was the financial availability of the program. We explored if there is any freeware or limited free of charge version of program because we expect that students are not willing to buy the necessary program which they usually need only a half-year of their five years studies. And then we tried to take into account the spread of these programs in the use or in the other lectures.

The other criterion was how the work with the program is intuitive - based on the knowledge of commonly used software like MS Office, and how easy it is to familiarize students with the program in eleven lessons we have during the term. If the program is not complicated, students are more focused on econometrics and then they gain required knowledge of this science.

We compared the heftiness of model creation, the graphic interface of the program and other functions like the graphic or statistic analysis. The important criteria were the ability of the programs to run different data files, their collaboration with MS Office (in the meaning of data export and import) or if the data samples are included in the program or on the program's web pages. We confronted the results provided with the programs and if they are intelligible and well-arranged even for econometric beginners. We focused on whether software displays all statistical tests important for the basic statistical and econometric verification of the model.

In this paper we discussed advantages and disadvantages of mentioned programs and tried to decide which one of these three programs is the most suitable for lessons of econometrics especially from the point of view of econometric students.


## 2   The illustrative example

A model of dependency of inflation rate on unemployment rate was selected as an illustrative example. This dependency was modeled on data sets from the Czech Republic during time period 1993 – 2007.

The basic Phillips curve describes the inverse relationship between the rate of wage inflation and unemployment rate. We estimated the modified Phillips curve which was published by P. A. Samuelson and R. M. Solow and it represents the inverse relationship between the unemployment rate and inflation rate. Friedman and Phelps criticized the long-term Phillips curve and they proved that the negative slope of the Phillips curve represents the money illusion of the workers and that this illusion is only short-term and there is no tradeoff between the inflation and unemployment and the long-term Phillips curve is vertical (Mach, 2001).

Students have to prove the inverse relationship between inflation and unemployment known as the Phillips curve. Furthermore they have to figure out the goodness of fitted model, test classical linear regression model assumption and then discuss the economic interpretation of the model and model's compliance with the known economic theory.

## 2.1  GiveWin

GiveWin is the software containing econometric techniques, from single equation methods to advanced cointegration and volatility models. GiveWin contains different program units for different models.

GiveWin was chosen for this analysis because it is the only software used for econometric lessons and modeling of time-series at our university. The interface of the program is similar to the other programs that are familiar to our students.

It is supposed that the user of this program have basic knowledge of statistics to understand the displaying of the output. It is possible to gain an example library with data sets from software websites.

The disadvantage of this software is that there is no full version as freeware and it causes that students can not work with the program on their personal computers without breaking of a copyright. The students can only use a limited student's version.

There is the possibility to create own database in GiveWin but it is not very user-friendly and it is better to export data from other software product although there can be problems with the definition of decimal separators.

The program uses two types of menus. The main menu is the same for all program units and it contains basic operations with data needed for calculations, graphic analysis etc. The second menu is diverse for every program unit and it is used to create the particular model. Switching between main menu of the program and menu in the program unit can be little bit confusing for students.

The standard output involves typical statistical values and tests. The output is divided into two tracks. The first one contains the values for every variable separately and the characteristics of the whole model are included in the second part. The following figure represents the standard output of the program.

```
EQ( 1) Modelling Inflation by OLS (using Data1)
       The estimation sample is: 1993 to 2007

                    Coefficient  Std.Error  t-value  t-prob Part.R^2
Constant               19.2151       3.882     4.95   0.000   0.6534
Unemployment          -2.03631      0.5781    -3.52   0.004   0.4883

sigma                  3.98826  RSS                 206.781255
R^2                   0.488309  F(1,13) =      12.41 [0.004]**
log-likelihood        -40.9612  DW                        1.09
no. of observations         15  no. of parameters            2
mean(Inflation)        6.03333  var(Inflation)         26.9409
```

**Fig. 1.** The output from GiveWin.

The graphical analysis is not involved in this output and it must be done separately. The following figure represents some particular parts of graphical analysis.
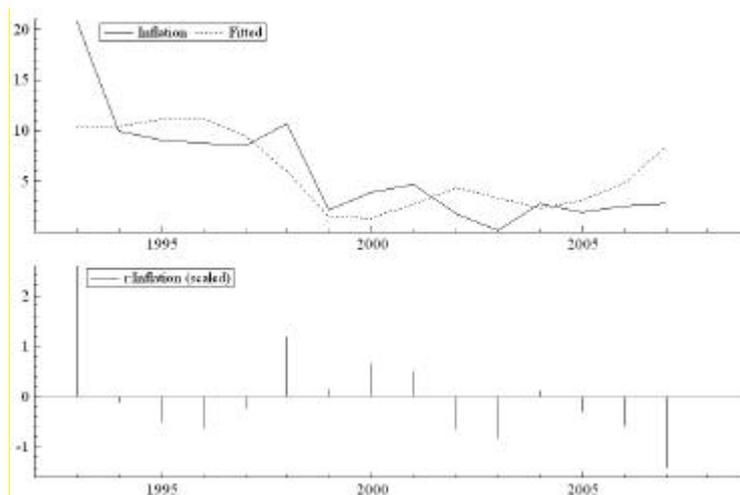
**Fig. 2.** Graphical analysis in GiveWin.

The graphical analysis is automatically used for predictions. The user has to change this setting if writing results instead of graphing is required.

The evaluation of autocorrelation based on the output can be misleading especially for econometric beginners because only Durbin-Watson statistic is displayed – no matter if it can or can not be used. We discovered problems with entering seasonality into the model.

## 2.2  Gretl

Gretl is the free and open-source software for econometric analysis with a large database of data sets on web sites.

Gretl was chosen for this analysis because it is downloadable, user-friendly and very intuitive software. Though we did not present this program to our students and we only told them where to download it, they were able to elaborate their half-yearly thesis in this software without further tutorial.

To run a regression, we need to import data files or create a new data file. The program offers a change of data definition during communication with user through simple dialog windows. The file with data does not need to be only in a standard database format and we discovered that Gretl communicates very well with excel and other database files.

The operating with Gretl is similar to the operating with E-views which is the other used statistical and econometric software. This can be an important advantage for future praxis of the students.

The software guides the user through well-arranged and intelligible windows. Each window contains menus which can be used for evaluation, graphic analysis, and advancement of the model or further diagnostic tests on the model.

The basic regression output involves a part with parameters' estimators and variables' characteristics. The characteristics of the model are specified below the first part and appear in the standard format. The figure 3 represents the standard regression output.

```
Model 1: OLS estimates using the 15 observations 1993-2007
Dependent variable: Inflation

                coefficient   std. error   t-ratio   p-value
---------------------------------------------------------------
const             19,2151      3,88155       4,950    0,0003  ***
Uneployment       -2,03631     0,578135     -3,522    0,0038  ***

Mean of dependent variable = 6,03333
Standard deviation of dep. var. = 5,37264
Sum of squared residuals = 206,781
Standard error of the regression = 3,98826
Unadjusted R-squared = 0,48831
Adjusted R-squared = 0,44895
Degrees of freedom = 13
Durbin-Watson statistic = 1,08853
First-order autocorrelation coeff. = 0,142501
Log-likelihood = -40,9612
Akaike information criterion (AIC) = 85,9223
Schwarz Bayesian criterion (BIC) = 87,3384
Hannan-Quinn criterion (HQC) = 85,9072
```

**Fig. 3.** The output from Gretl.

Students appreciate that the model's characteristics are not named in shortcuts. Using complete names makes the output more intelligible especially to econometric beginners who are not familiar with statistical shortcuts. If necessary for statistical verification, the regression output is extended with additional statistics (e.g. Durbin's $h$) and this is very beneficial and helpful during econometric lessons.
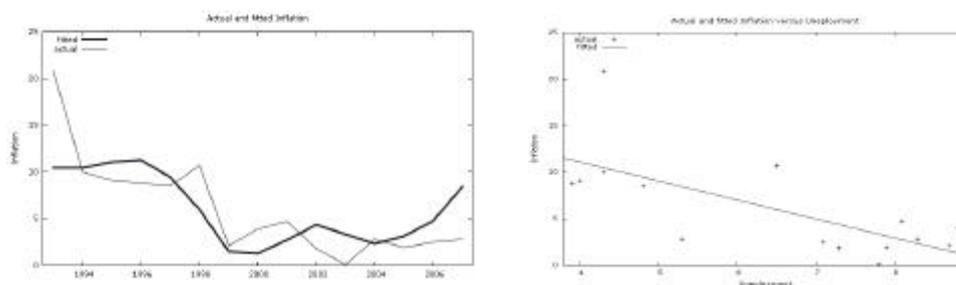


**Fig. 4.** Graphic analysis in Gretl.

Gretl offers wide range of graphs, two examples are shown on figure 4. This software allows the best editing of graphs among all the compared programs, e.g. saving graphs directly in PDF format.

For econometric analysis and procedures the command-line client program can be used, sample functions and commands are implemented in user's guide.

## 2.3  SAS Enterprise Guide

SAS Enterprise Guide is a graphical interface of SAS base, which is one of the best known statistical software. SAS Enterprise Guide contains statistical functions, from descriptive statistics to multivariate analysis. SAS Enterprise Guide was chosen for this analysis because it is the software that each student of our university knows from basic statistical lessons. Therefore students are aware of features and settings of SAS Enterprise Guide and they can more concentrate on econometric problems and new functions. The other advantages are the user-friendly interface and the possibility to display the output in html format. This format is more readable for non-statistical users. SAS Enterprise Guide allows writing code in SAS language. Products of SAS Company are widely used in praxes. SAS Enterprise Guide contains an example library with data sets.

A disadvantage of this software is the non econometric character. Also there is no freeware of this software and this causes that students cannot work with it on their personal computers for free.

SAS Enterprise Guide allows you to read datasets in different formats by function *import data*. For running a linear regression there is a simple choice *analyze – regression - linear.* In the window of linear regression you can choose task roles (dependent, explanatory variables), their types (numeric, character), model selection method (stepwise selection), model fit statistics (Akaike's information criterion, Sawa's Bayesian information criterion),  additional statistics (confidence interval of parameter estimates on different significance level, partial correlation) and also autocorrelation test and heteroskedasticity test. In the same menu you can choose different types of charts.
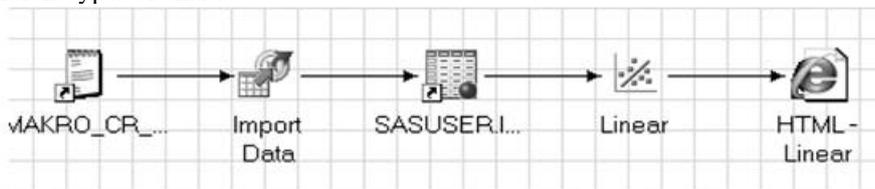


**Fig. 5.** Project diagram in SAS Enterprise Guide.

The output is automatically displayed in html format and it contains classical statistical test for significance of variables and of model and selected types of plots. This output provides basic information about fulfillment of Gauss-Markov theorem.

In the last tables of output, it is tested whether error terms are independent and identically distributed. Thee last table provides information about autocorrelation and displays Durbin-Watson statistics and autocorrelation coefficient *?*.

| Analysis of Variance | | | | | |
|---|---|---|---|---|---|
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 197.33208 | 197.3320 | 12.41 | 0.0038 |
| Error | 13 | 206.78126 | 15.90625 | | |
| Corrected Total | 14 | 404.11333 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 3.98826 | R-Square | 0.4883 |
| Dependent Mean | 6.03333 | Adj R-Sq | 0.4489 |
| Coeff Var | 66.10382 | | |

| Parameter Estimates | | | | | |
|---|---|---|---|---|---|
| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| |
| Intercept | 1 | 19.21505 | 3.88155 | 4.95 | 0.0003 |
| Unemployement | 1 | −2.03631 | 0.57813 | −3.52 | 0.0038 |

| Test of First and Second Moment Specification | | |
|---|---|---|
| DF | Chi- | Pr > ChiSq |
| 2 | 3.41 | 0.1815 |

| | |
|---|---|
| Durbin-Watson D | 1.089 |
| Number of Observations | 15 |
| 1st Order Autocorrelation | 0.121 |

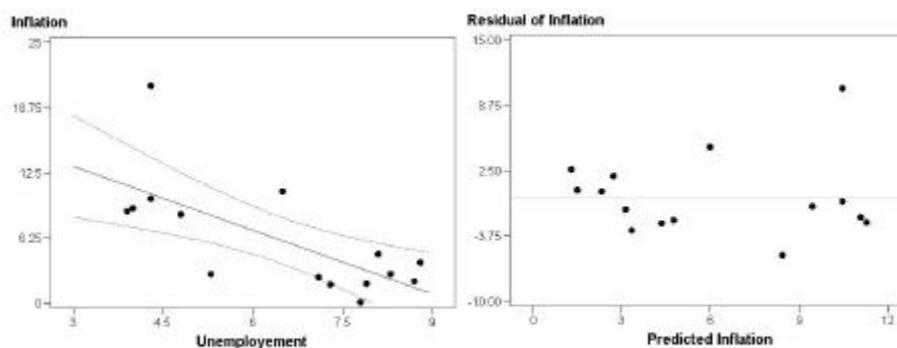**Fig. 6**. Output from SAS Enterprise Guide.



**Fig. 7**. Residuals in SAS Enterprise Guide.

Working in SAS Enterprise Guide allows you to do a quick and readable analysis. This characteristic is very useful for students working on their final thesis and doing basic analysis of economic variables. For further econometric analysis it is possible to

use other tests by "writing code", samples codes are implemented in SAS Enterprise Guide help.

## 3 Decision analysis

We applied methods of multiple criteria decision making to choose one of three programs. Each program was evaluated according to seven criteria: Financial availability, Intuitiveness of program, Heftiness of model creation, Graphic interface, Using in praxes, Import of data files, Econometrics function. Methods of multiple criteria decision making are searching a compromise solution by trying "maximize" or "minimize" different criteria. The list of the criteria and the values used in our analysis are in the Table 1.

**Table 1.** Criteria.

| Criteria | Label | | GRETL | GiveWin | SAS | Scale |
|----------|-------|-----|-------|---------|-----|-------|
| DAT | *Run different data files* | MAX | 4 | 3 | 5 | 5 - runs different kinds of data files, 0 - just one type of data files |
| ECO | *Econometrics function* | MAX | 5 | 5 | 3 | 5 - large scale, 0 - restricted |
| FIN | *Financial availability* | MAX | 2 | 1 | 0 | 2 - freeware, 1 - restricted student version, 0 - no freeware |
| GRA | *Graphic interface* | MAX | 4 | 2 | 5 | 5 - advanced graphic interface, 0 - restricted graphic interface |
| HEF | *Heftiness of model creation* | MAX | 5 | 2 | 4 | 5 - easy fitting, 0 - complecated fitting |
| INT | *Intuitiveness of program* | MAX | 5 | 2 | 5 | 5 - user friedly, 0 - user unfriendly |
| PRA | *Using in praxes* | MAX | 1 | 2 | 5 | 5 - widly use, 0 - no use |

The criteria have a different importance in this analysis. We used Fuller method to define weights of criteria. Fuller method uses the couple comparison of criteria. There is the graphical scheme of comparison called Fuller's triangle.
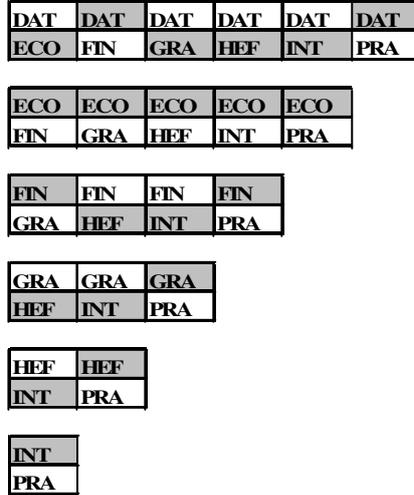
| DAT | DAT | DAT | DAT | DAT | DAT |
|-----|-----|-----|-----|-----|-----|
| ECO | FIN | GRA | HEF | INT | PRA |

| ECO | ECO | ECO | ECO | ECO |
|-----|-----|-----|-----|-----|
| FIN | GRA | HEF | INT | PRA |

| FIN | FIN | FIN | FIN |
|-----|-----|-----|-----|
| GRA | HEF | INT | PRA |

| GRA | GRA | GRA |
|-----|-----|-----|
| HEF | INT | PRA |

| HEF | HEF |
|-----|-----|
| INT | PRA |

| INT |
|-----|
| PRA |

**Fig. 8.** Fuller's triangle

We computed weights like

$$v_i = \frac{n_i}{\sum_{i=1}^{k} n_i} = \frac{n_i}{N}, \tag{1}$$

where $v_i$ denotes weights of criteria $i$,
$n_i$ denotes how many times was criteria $i$ selected and
$N$ is total number of comparison obtain by

$$N = \binom{k}{2} = \frac{k(k-1)(k-2)!}{2!(k-2)!} = \frac{k(k-1)}{2}, \tag{2}$$

where $k$ is number of criteria.

The method of weight average was chosen to determinate the compromise solution. Normalized values and weights are in Table 2. The highest weight average has Gretl.

**Table 2.** Normalized values of criteria and weights.

| Criteria | Label | Weights | Normalized values | | |
|---|---|---|---|---|---|
| | | | GRETL | GiveWin | SAS |
| DAT | *Run different data files* | *0,095* | 0,500 | 0,000 | 1,000 |
| ECO | *Econometrics function* | *0,286* | 1,000 | 1,000 | 0,000 |
| FIN | *Financial availability* | *0,095* | 1,000 | 0,500 | 0,000 |
| GRA | *Graphic interface* | *0,143* | 0,667 | 0,000 | 1,000 |
| HEF | *Heftiness of model creation* | *0,190* | 1,000 | 0,000 | 0,667 |
| INT | *Intuitiveness of program* | *0,190* | 1,000 | 0,000 | 1,000 |
| PRA | *Using in praxes* | *0,000* | 0,000 | 0,250 | 1,000 |
| Weight average | | | 0,905 | 0,333 | 0,556 |

## 4 Conclusion

In this paper we compared three different programs used for econometric analysis with the goal to decide which of mentioned software would be the best for basic econometric students at the University of Economics in Prague. We were limited especially with software availability at our university. We decided according to different criteria using multiple criteria decision methods. We discussed the advantages and disadvantages of the programs with respect to usage during our elementary econometric lessons. We did not discuss advanced econometric procedures and methods like reaction functions or panel-data analysis. Those are not a content of basic econometric lectures.

Based on this analysis we recommend using Gretl as the fundamental software for our lectures even if this analysis can not be generalized due to specific conditions as the selection of criteria and the criteria weights used in this paper.

# References

1. GRETL webpage: http://gretl.sourceforge.net/
2. Guaranti D.N,: Basic econometrics. McGraw-Hill, Inc.,USA (1988)
3. Hušek, R.: Ekonometrická Analýza. Ekopress, Praha (1999)
4. Jablonský, J.: Operacní výzkum: kvantitativní modely pro ekonomické rozhodování.
5. Professional Publishing, Praha (2007).
6. Mach M.: Makroekonomie II pro magisterské (inženýrské) studium 1. a 2. cást, p. 268-269. Melandrium, Slaný (2001)
7. OxMetrix webpage: http://www.oxmetrics.net/
8. SAS webpage: http://www.sas.com/technologies/bi/query_reporting/guide/