



Journal of Statistical Software

March 2008, Volume 25, Software Review 1.

<http://www.jstatsoft.org/>

Reviewer: Andreas Rosenblad
Uppsala University

gretl 1.7.3

The **gretl** Team, Winston-Salem, NC. Open source.
<http://gretl.sourceforge.net/>

Introduction

gretl, short for *GNU Regression, Econometrics and Time-series Library*, is an open-source statistical software package with the main focus being on statistical methods for econometric analyses. Its origin is the **Econometrics Software Library** (ESL) package written by Ramu Ramanathan, Professor Emeritus at the Economics Department, University of California, San Diego, who later made it open source and handed it over to Professor Allin Cottrell at the Department of Economics, Wake Forest University. The latter continued to develop the package, and after releasing a few development versions the first stable version, **gretl** 1.0, was released to the public on November 15, 2002 (<http://gretl.sourceforge.net/ChangeLog.html>). It is in continuous development, with the latest version being 1.7.4, released on March 21, 2008. Besides Professor Cottrell the main developer is Riccardo “Jack” Lucchetti, Associate Professor of Econometrics at the Dipartimento di Economia, Università Politecnica delle Marche, Ancona, Italy. **gretl**’s motto is “by econometricians, for econometricians”.

Originally developed on Linux, **gretl** is now available also on Microsoft Windows and Mac OS X. It is written in the C programming language and available for free according to the GNU General Public License (<http://www.gnu.org/licenses/gpl.html>). The Free Software Foundation (<http://www.fsf.org/>) has also adopted **gretl** as a GNU program. Besides English, **gretl** is available in Basque, French, German, Italian, Polish, Portuguese and Spanish.

Like many other open-source software projects, **gretl** is hosted on SourceForge (<http://sourceforge.net/>), where the development of the program is coordinated through the CVS (Cederqvist et al. 2006) version control system. The source code can be found at <http://gretl.cvs.sourceforge.net/gretl/gretl/>. To discuss the development of **gretl**, a mailing list (<http://lists.wfu.edu/pipermail/gretl-devel/>) is used.

Earlier reviews of **gretl** include Baiocchi and Distaso (2003), Mixon and Smith (2006), and Yalta and Yalta (2007), which reviewed versions 0.997, 1.5.1, and 1.6.0, respectively. The conclusion of Baiocchi and Distaso (2003, p. 110) was that

The overall quality, free availability, and ease of use of **gretl** should encourage econometricians to use it in teaching and for some research purposes.

Mixon and Smith (2006) mainly discussed the usefulness of **gretl** for teaching undergraduate econometrics. Their conclusion (p. 1107) was that

gretl is a full-featured package. It is user-friendly for students taking an introductory econometrics course, and one that will provide most of these students all of the analytical firepower they will require as they proceed into their careers.

In the review by Yalta and Yalta (2007), the authors made a comprehensive testing of the numerical accuracy of **gretl**. Their conclusion (p. 853) was that

gretl is a high-quality, feature-rich, and accurate econometrics package. Most importantly, it is a public resource for all economists around the world. The intuitive and well laid out user interface makes it an ideal tool for learning and teaching econometrics, as discussed in detail by [Mixon and Smith \(2006\)](#). Furthermore, the program proves to be as good or even better in terms of numerical precision compared to other, widely used alternatives.

Supplementary materials to [Yalta and Yalta \(2007\)](#), where details from the comparison of numerical accuracy between **gretl** version 1.6.0, SAS 6.12 ([SAS Institute Inc. 1999](#)), SPSS 7.5 ([SPSS Inc. 1998](#)), S-PLUS 4.0 ([Insightful Corp. 1997](#)), Stata 7 ([StataCorp 2001](#)), and GAUSS 3.2 ([Aptech Systems, Inc. 1995](#)) are given, can be found at <http://www.econ.queensu.ca/jae/2007-v22.4/yalta-yalta/>.

The current review of **gretl** uses version 1.7.3, released on 2008-12-29, and was performed on Windows XP with Service Pack 2 using an AMD Athlon 64 X2 4400+ with 2.3 GHz CPU and 2.9 GB RAM.

Installation

gretl is available for download from its Web page <http://gretl.sourceforge.net/> as a prebuilt self-installing executable file for Microsoft Windows, disk image for Mac OS X, Debian GNU/Linux package, generic binary rpm for modern Linux systems, or as source files for the user who wants to compile from source. The installation on Windows XP is automatic and flawless. It was also installed on a PC using Windows Vista, which was equally easy.

For generating graphs **gretl** uses the open-source plotting program **gnuplot** ([gnuplot Team 2008](#))—for a review, see [Racine \(2006\)](#)—which is included in the executable file for Microsoft Windows and automatically installed together with **gretl**. After installing **gretl**, the user can also install two optional extra software packages, the **X-12-ARIMA** seasonal adjustment software of the [US Census Bureau \(2007\)](#) and the **TRAMO/SEATS** ([Gómez and Maravall 1997](#), Time series Regression with ARIMA Noise, Missing values and Outliers/Signal Extraction in ARIMA Time Series) software packages for analyzing monthly or lower frequency time series data. These are available for download from the **gretl** Web page as prebuilt self-installing executable files for Microsoft Windows or binary rpm packages for GNU/Linux, and can after

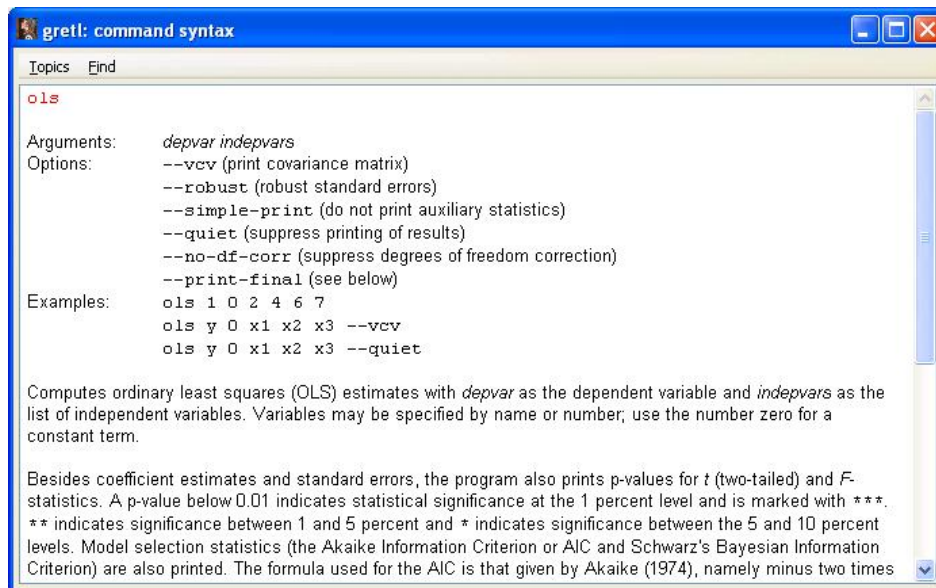


Figure 1: Browsable help files.

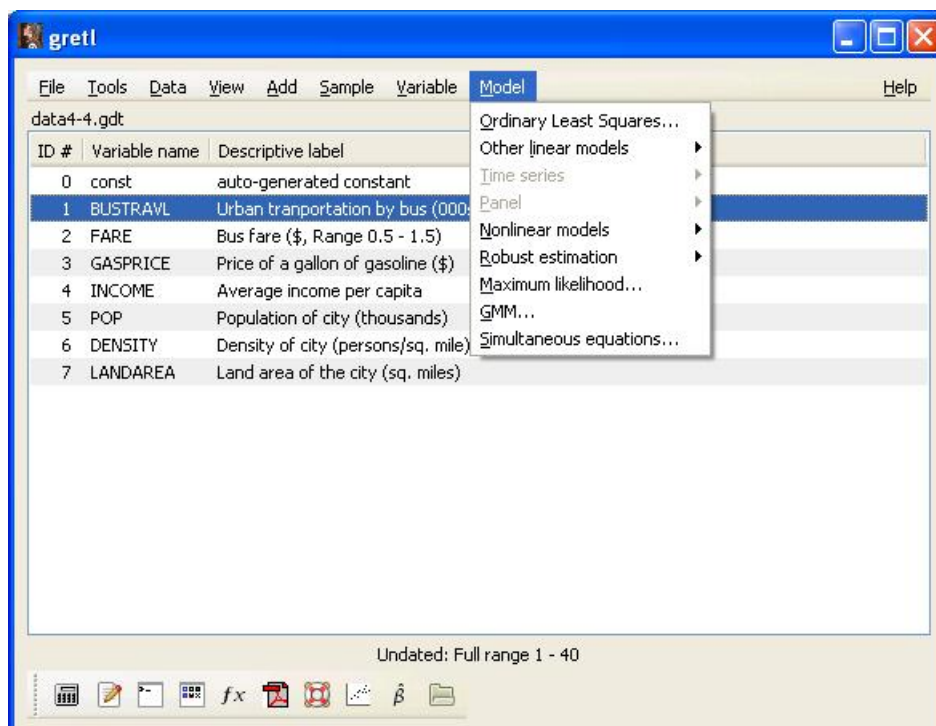


Figure 2: The main window in gretl with a menu opened.

installation be accessed from inside **gretl**. Finally, the **gretl** Web page also provides datasets for several popular econometrics textbooks, which can be installed into **gretl**.

Manuals and support

gretl is shipped with substantial pieces of documentation in PDF format, available via the **Help** menu: **gretl User's Guide** (Cottrell and Lucchetti 2008b), 207 pp, and **gretl Command Reference** (Cottrell and Lucchetti 2008a), 109 pp. Command and function references from the latter are also available as browseable help files (Figure 1). The *User's Guide* consist of four parts: Part I discusses the basics of running **gretl** and gives an overview of its graphical user interface (GUI) and command line interface (CLI), the handling of data files, and how to use **gretl**'s scripting language. Part II is devoted to the application of econometric methods in **gretl**, with detailed discussions of the estimation methods that are available in **gretl**. Part III gives some technical details, mainly for **gretl**'s L^AT_EX-related functionality. Part IV, finally, contains a few appendices with details about **gretl**'s data file format, how to build **gretl** from source, and **gretl**'s numerical accuracy. The *User's Guide* gives a good and detailed discussion of the program's features, including screenshots of input and output windows as well as discussions of syntaxes and formulas for the implemented features, with examples that makes it easy to understand how to use the program.

The *Command Reference* gives detailed descriptions of the commands and functions that can be used interactively with **gretl**'s command line interface or in script or batch files for non-interactive execution, together with discussions about options and arguments that can be used when invoking **gretl**. The reference manual is well-written and easy to understand, making it easy for the user who wants to use the CLI instead of the GUI to perform the analyses of his/her choice. Since **gretl** is a free, open-source software, the user cannot demand to get any help or support from the developers. However, they provide a mailing list for users (<http://lists.wfu.edu/pipermail/gretl-users/>), where one can get help and support from other users and also often from the developers.

Besides the two manuals provided by the developers of **gretl**, there is also a third party manual, *Using **gretl** for Principles of Econometrics, 3rd Edition* (Adkins 2007), 377 pp, written by Professor Lee C. Adkins at the Department of Economics and Legal Studies in Business, Oklahoma State University. It is written to accompany the third edition of *Principles of Econometrics* (Hill, Griffiths, and Lim 2008), and is available for free from <http://www.LearnEconometrics.com/gretl.html>. This manual focuses on the practical aspects of using **gretl** in econometrics, illustrated with numerical examples using both the GUI and the CLI. Although it is intended to be used with Hill *et al.* (2008), it can also be used independently of this book as a manual for performing econometrics using **gretl**.

User interface

Compared to other open source statistical programs, such as e.g. R (R Development Core Team 2008), one of **gretl**'s most attractive features is its intuitive and user friendly graphical user interface, built with the widely used open-source **GTK+** (Krause 2007) widget toolkit, which give the user access to all the program's features from menu entries and dialog boxes. Figure 2 shows **gretl**'s main window for a sample session. To estimate, e.g., a linear regression model using OLS one chooses `Model > Ordinary Least Squares...` from the menu bar,

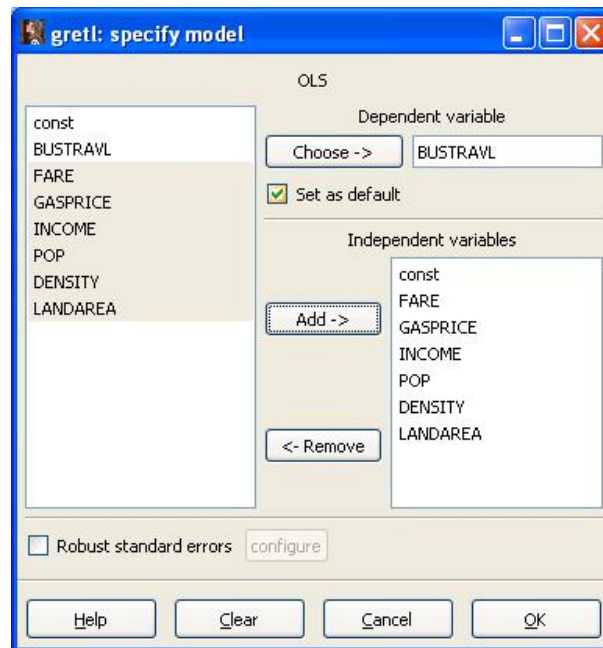


Figure 3: OLS estimation dialog box.

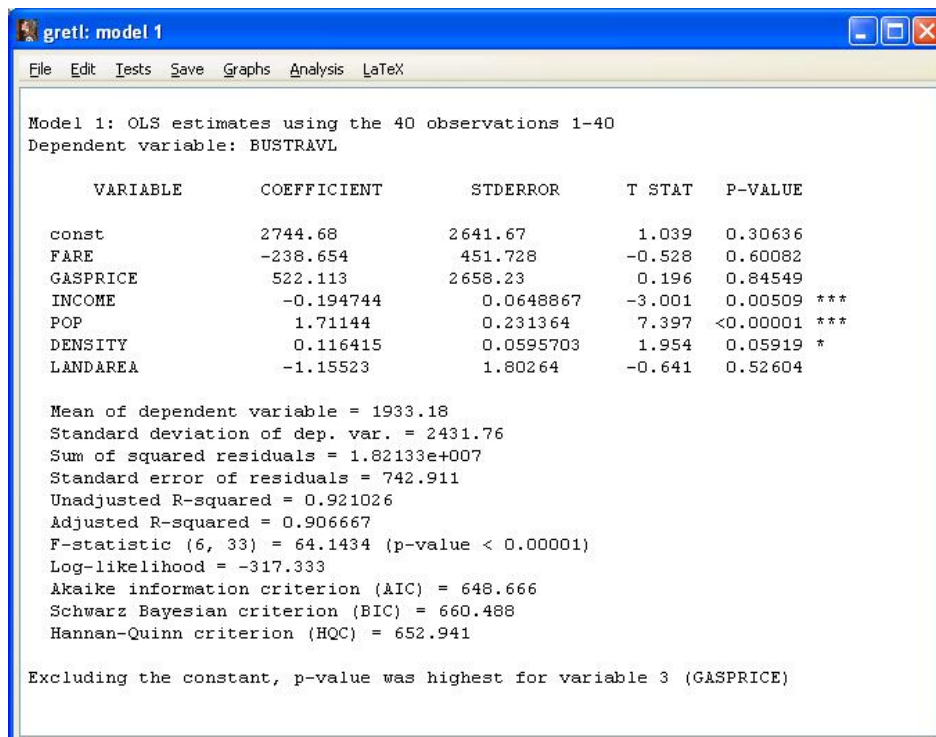


Figure 4: OLS estimation output window.

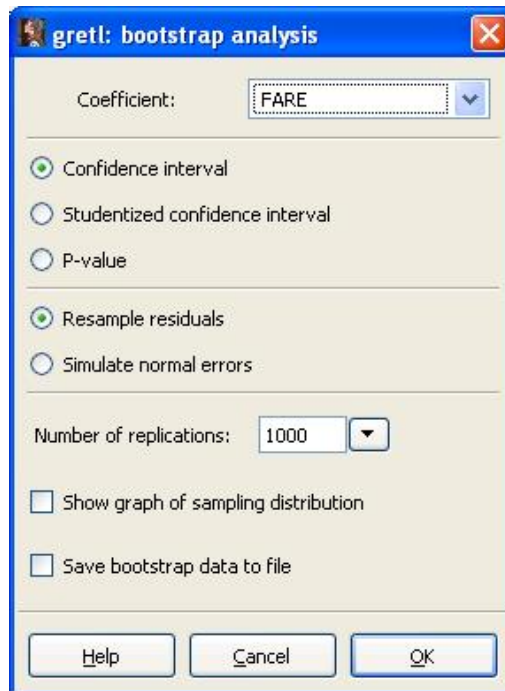


Figure 5: Bootstrap analysis dialog box for OLS estimation.

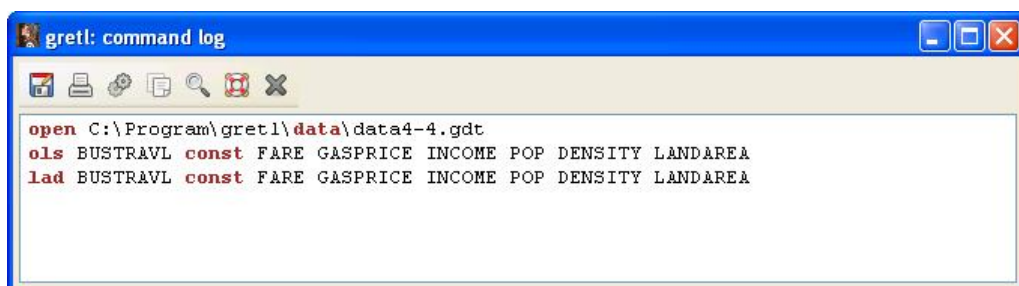


Figure 6: Command log.

and in the dialog box (Figure 3) selects the variables one wants to include in the model. The results of the OLS model estimation are shown in an output window (Figure 4). From this the user can access several new menus, where he/she can perform new tests or analyses of the data, plot graphs of the results, save the results (e.g. residuals and fitted values) to the current data set for later use, or use the results in other ways. For example, from the menu **Analysis > Bootstrap...** one can access a dialog box for bootstrap analysis of the model (Figure 5). All analyses that are performed via menu entries are also saved as scripts in a command log (Figure 6), which can then be rerun or saved and used for later analyses.

Besides the menu bar at the top, **gretl** also has a set of toolbar buttons at the bottom of the main window (see Figure 2). This gives the user faster access to some often used features, such

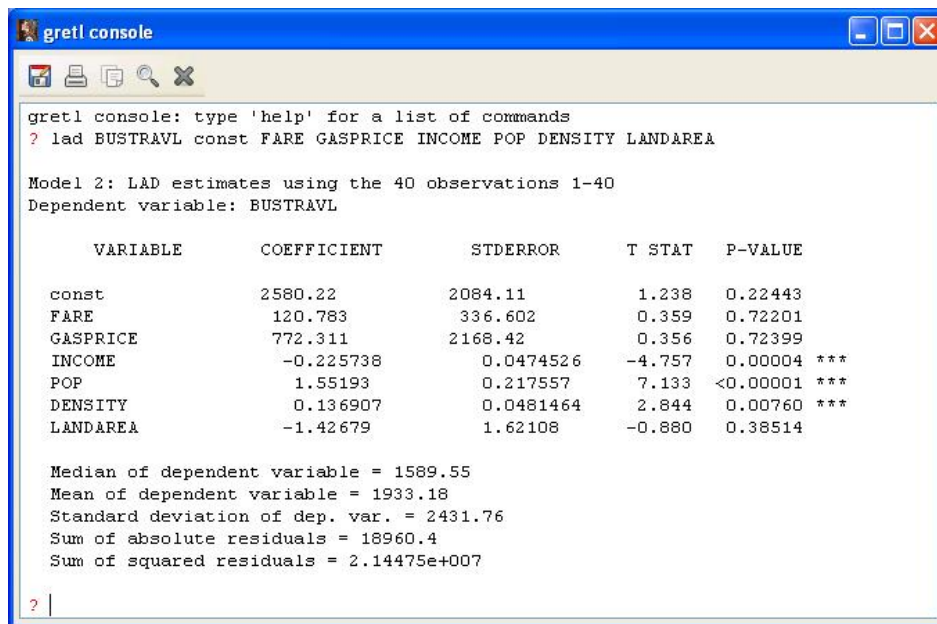


Figure 7: Console window with command line interface.

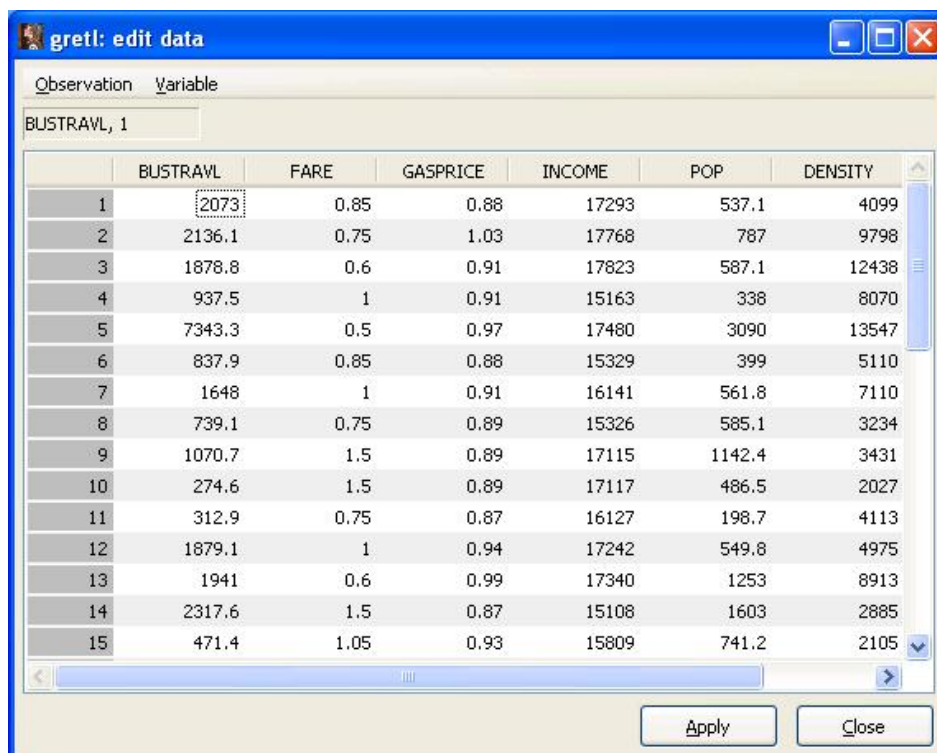


Figure 8: Edit data window.

as creating new script files, opening the session icon view (where the analyses from the current session are saved) or opening **gretl**'s console window (Figure 7). The latter is **gretl**'s CLI run from inside the GUI. To edit a data set one can open one or all variables in a spreadsheet-like window (Figure 8).

Features

gretl uses its own XML-based file format, but can import data from several other file formats, including comma-separated values (CSV), ASCII, Open Document Format (ODF) for spreadsheets, Microsoft **Excel**, **Gnumeric** (**Gnumeric Team 2008**), GNU Octave (**Eaton 2002**), **EViews** (**Quantitative Micro Software 2007**) workfiles, Stata and **JMulTi** (**Lütkepohl and Krätzig 2004**) file formats, as well as the RATS (**Estima 2007**) and **PcGive** (**Doornik and Hendry 2007**) database formats. It can also export data to CSV, R, GNU Octave, **JMulTi** and **PcGive** data formats.

gretl supports a wide range of statistical methods, with a focus on methods used in econometrics, especially regression methods. For cross-sectional data one can estimate several linear models, besides OLS also weighted least squares, two-stage least squares and least absolute deviation regression (LAD, see Figure 7), as well as many nonlinear models, such as logit, probit, tobit, heckit, Poisson, logistic and non-linear least squares. One can also specify and estimate maximum likelihood, generalized method of moments (GMM) and simultaneous equations models, with the latter including seemingly unrelated regressions (SUR), ordinary or weighted least squares, two- or three-stage least squares and full or limited information maximum likelihood estimators. For panel data one can estimate, e.g., models with fixed or random effects and dynamic panel models.

Likewise, for time series data a wide range of analysis methods are available, including filtering with simple and exponential moving average, the Hodrick-Prescott filter and Baxter-King band pass filter, ADF, ADF-GLS and KPSS unit root tests as well as Engle-Granger and Johansen cointegration tests, estimation using the Cochrane-Orcutt, Hildreth-Lu and Prais-Winsten estimators, and modeling with ARIMA, ARCH/GARCH, VAR and VECM. Besides these built-in methods, **gretl** is also integrated with the two external time series software packages **X-12-ARIMA** and **TRAMO/SEATS** and provides GUI frontends for these, accessible from inside **gretl**. A useful feature for the estimation of cross-sectional, time series and panel data models is that the output can also be viewed, copied or saved in L^AT_EX format.

Besides the above mentioned features, **gretl** also provides access to standard statistical methods such as descriptive statistics, cross tabulation and correlation measures as well as plotting of histograms, time series plots, scatter plots, boxplots and 3D plots, among others. The layout of the graphs can be changed to the user's taste with an advanced GUI for plot controls. Another useful feature is the **Tools** menu, where one can access menu entries for calculating critical values or P-values for several probability distributions and test statistics calculators for parametric and nonparametric tests. Figure 9 shows the test statistics calculator for parametric tests.

Although many of **gretl**'s users will probably mainly use the program through its GUI, it should be noted that it also has an advanced scripting language, which is very useful when one is doing research. This can be used by running script files with commands, but also interactively either from the console window (Figure 7) or by starting and running **gretl** in a

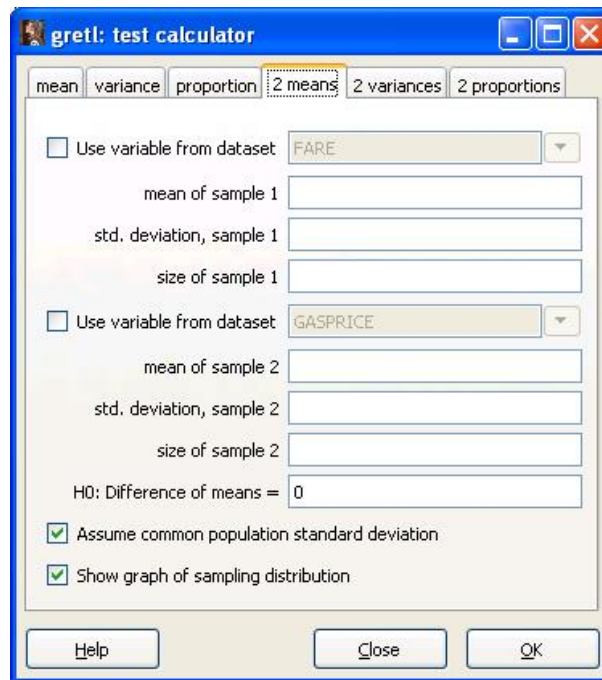


Figure 9: Test statistics calculator for parametric tests.

```

RGui - [R Console]
File Edit View Misc Packages Windows Help

R version 2.6.2 (2008-02-08)
Copyright (C) 2008 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> # load data from gretl
> library(stats)

> vnum <- as.double(R.version$major) + (as.double(R.version$minor) / 10.0)

> if (vnum > 2.41) library(utils)

> gretldata <- read.table("C:/Documents and Settings/uanka06/Application Data/gretl/Rdata.tmp", header=TRUE)

> attach(gretldata)
> |

```

Figure 10: R started from inside **gretl**, with the currently active **gretl** data set loaded into R.

pure CLI mode through the program's command-line client, **gretlcli**. The scripting language includes matrix operations, loop constructs, boolean operators, user-defined functions, and other features one should expect, besides commands for **gretl**'s built-in statistical methods.

Finally, a useful feature of **gretl** is its integration with R. Besides saving data in R format, if one has installed R and wants to use a statistical method that is not yet available in **gretl** but is available in R, one can click on the menu entry **Tools > Start GNU R**. This will export the currently active **gretl** data set to R format and open R with this data set loaded, ready for further analyses in R (Figure 10).

Extensions

gretl has many useful features for statistical analyses, especially in the areas of econometrics and time series analysis. To extend these capabilities, the users can write their own functions. A very useful feature is that these user-written functions can easily be provided with a GUI, packaged with **gretl** and distributed to other users.

Suppose that we want a function that gives the sample size n necessary to obtain a power $1 - \beta$ with a significance level α in detecting a difference $\delta = \mu_1 - \mu_2$ for a test of difference between two independent and normally distributed samples x_i and y_i with means μ_1 and μ_2 . The hypotheses of interest are thus $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$. Assuming that the variances σ_1^2 and σ_2^2 are known and that $n_1 = n_2 = n$, this can be tested with the test statistic

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}},$$

where $Z \sim N(0, 1)$ under H_0 . The sample size n is then given by the formula

$$n = \left\lceil \frac{(\sigma_1^2 + \sigma_2^2) (z_{\alpha/2} + z_{\beta})^2}{\delta^2} \right\rceil.$$

A **gretl** script for a function that calculates n is given by the following code:

```
function samplesize(scalar Alpha[0.05], scalar Power[0.80], \
scalar Variance1, scalar Variance2, scalar Difference)
n = ceil(((Variance1 + Variance2) * (critical(N, Alpha/2) + \
critical(N, 1-Power))^2) / Difference^2)
printf "n = %g\n", n      #print n and add a new line
end function
```

After running the function script, the function can be called from the **gretl** CLI:

```
? samplesize(0.05, 0.80, 1, 1, 1)
n = 16
```

With a few mouse clicks one can also package and save the function on the local machine and at the same time choose to upload the package to the **gretl** server so that it can be

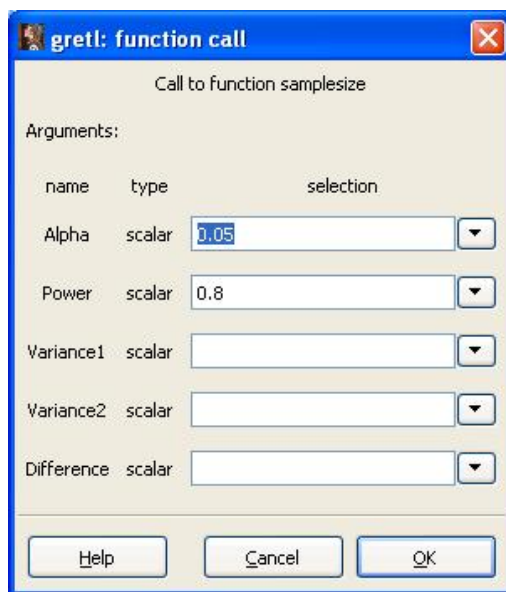


Figure 11: GUI for the `samplesize` function, automatically provided by **gretl**.

accessed by other users. In the packaging process the function is automatically provided with a GUI (Figure 11), which can be ran from the **gretl** server or from the local machine. This makes **gretl** a good choice for statisticians without knowledge of C or other general-purpose programming languages that want to write user friendly statistical programs with a GUI and make these available to the statistical community.

What is missing?

Although **gretl** has many features for performing econometrics and time series analyses, there are some features that are missing. These are chiefly methods for quantile regression, survival analysis (such as Cox regression) and multinomial logit regression. Also, the editing of data could be easier with e.g. adding the possibility to copy and paste arbitrary parts of the data set in the data editing window (currently only a single observation for a single variable can be copied and pasted) and the handling of string variables in the data sets, which currently is not possible.

The facility for adding user-written functions to **gretl**, useful as it is, could be polished and expanded to make it more useful still. For example, the GUI layout provided by **gretl** could be polished to look better and the user who writes the function should have the possibility to influence how it is laid out. Further, a feature to package several main functions into a joint package (as is possible in R), instead of distributing every function separately, would also be welcome. This would make **gretl** an even better environment for developing statistical programs.

Finally, in the opinion of **gretl**'s two main developers, more C coders are needed. The reviewer agrees with this opinion. Since the program is open source and totally volunteer-based, it is dependent on people contributing code to the project. While the two main developers has

done and continues to do a great job, for **gretl** to continue to develop and improve it would benefit from more C programmers to join the project and contribute code.

Conclusions

gretl is an easy-to-use statistical software that offers most of the features necessary for performing econometrics and time series analyses. Its user-friendliness and the fact that it is available as open source makes it an ideal software for teaching and learning econometrics and time series analysis. Since the students can freely download and install the software at their home computers, they need not rely on the teaching institutions computer labs, which should save money for the latter. Professional econometricians and economists will also find **gretl** useful. The facts that it is open source and programmed in the widely used C programming language make it easy for other statisticians and econometricians with programming skills to join in the development of the program, adding more features and making it even better. The future of **gretl** thus looks good. It can be recommended to both students and professionals.

Acknowledgments

Thanks to Professor Allin Cottrell, Dr. Riccardo “Jack” Lucchetti and the JSS editors for useful comments and suggestions.

References

- Adkins LC (2007). *Using **gretl** for Principles of Econometrics, 3rd Edition: Version 1.0*. URL <http://www.LearnEconometrics.com/gretl/ebook.pdf>.
- Aptech Systems, Inc (1995). *GAUSS Mathematical and Statistical System, Version 3.2*. Aptech Systems, Inc., Black Diamond, Washington. URL <http://www.Aptech.com/>.
- Baiocchi G, Distaso W (2003). “**gretl**: Econometric Software for the GNU Generation.” *Journal of Applied Econometrics*, **18**(1), 105–110. doi:10.1002/jae.704.
- Cederqvist et al P (2006). *Version Management with CVS*. Network Theory Ltd, Bristol. ISBN 0-9541617-1-8, URL <http://www.Network-Theory.co.uk/docs/CVSmanual/>.
- Cottrell A, Lucchetti R (2008a). ***gretl** Command Reference*. Winston-Salem, NC. URL <http://ricardo.ecn.wfu.edu/pub//gretl/manual/en/gretl-ref.pdf>.
- Cottrell A, Lucchetti R (2008b). ***gretl** User’s Guide*. Winston-Salem, NC. URL <http://ricardo.ecn.wfu.edu/pub//gretl/manual/en/gretl-guide.pdf>.
- Doornik JA, Hendry DF (2007). *PcGive 12*, volume I–IV. Timberlake Consultants Press, London. URL <http://www.PcGive.com/>.
- Eaton JW (2002). *GNU Octave Manual*. Network Theory Ltd, Bristol. ISBN 0-9541617-2-6, URL <http://www.Network-Theory.co.uk/docs/Octave/>.

- Estima (2007). *RATS 7.0*. Estima, Evanston, IL. URL <http://www.Estima.com/>.
- Gómez V, Maravall A (1997). “Programs **TRAMO** (Time Series Regression with ARIMA Noise, Missing Observations and Outliers) and **SEATS** (Signal Extraction in ARIMA Time Series): Instructions for the User (Beta Version: Junio 1997).” *Working Paper 970001*, Ministerio de Economía y Hacienda, Dirección General de Análisis y Programación Presupuestaria, Madrid. URL <http://www.bde.es/servicio/software/econome.htm>.
- Hill RC, Griffiths WE, Lim GC (2008). *Principles of Econometrics*. Wiley, Chichester, 3rd edition. URL <http://www.PrinciplesofEconometrics.com/>.
- Insightful Corp (1997). *S-PLUS 4.0*. Insightful Corp., Seattle, WA. URL <http://www.Insightful.com/>.
- Krause A (2007). *Foundations of GTK+ Development*. Apress, Berkeley, CA. ISBN 1-59059-793-1, URL <http://www.GTKbook.com/>.
- Lütkepohl H, Krätzig M (eds.) (2004). *Applied Time Series Econometrics*. Cambridge University Press, Cambridge. URL <http://www.JMulti.com/>.
- Mixon JW, Smith RJ (2006). “Teaching Undergraduate Econometrics with **gretl**.” *Journal of Applied Econometrics*, **21**(7), 1103–1107. doi:10.1002/jae.927.
- Gnumeric** Team (2008). *Gnumeric 1.8.2*. GNOME Foundation, Cambridge, MA. URL <http://www.GNOME.org/projects/Gnumeric/>.
- gnuplot** Team (2008). *gnuplot 4.2.3*. Hanover, NH. URL <http://www.gnuplot.info/>.
- Quantitative Micro Software (2007). *EViews 6*. Quantitative Micro Software, Irvine, CA. URL <http://www.EViews.com/>.
- Racine J (2006). “**gnuplot** 4.0: A Portable Interactive Plotting Utility.” *Journal of Applied Econometrics*, **21**, 133–141. doi:10.1002/jae.885.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- SAS Institute Inc (1999). *The SAS System, Release 6.12*. SAS Institute Inc., Cary, NC. URL <http://www.SAS.com/>.
- SPSS Inc (1998). *SPSS 7.5*. SPSS Inc., Chicago, IL. URL <http://www.SPSS.com/>.
- StataCorp (2001). *Stata 7*. StataCorp LP, College Station, TX. URL <http://www.Stata.com/>.
- US Census Bureau (2007). *X-12-ARIMA, Version 0.3*. US Census Bureau, Washington, DC. URL <http://www.census.gov/srd/www/x12a/>.
- Yalta AT, Yalta AY (2007). “**gretl** 1.6.0 and its Numerical Accuracy.” *Journal of Applied Econometrics*, **22**(4), 849–854. doi:10.1002/jae.946.

Reviewer:

Andreas Rosenblad
Uppsala University
Center for Clinical Research Västerås
Central Hospital
S-721 89 Västerås, Sweden
E-mail: andreas.rosenblad@ltv.se
URL: <http://www.ltv.se/ckfvasteras/>