FACING THE FOUR-F TEST: GRETL´S ACHIEVEMENTS AND CHALLENGES

Ana Jesús López-Menéndez, anaj@uniovi.es
Rigoberto Pérez Suárez, rigo@uniovi.es
University of Oviedo, SPAIN

Abstract
Schools of Economics and Business face the challenge of training individuals capable of analysing, understanding and explaining the functioning of the economy. With this aim, Statistics and Econometrics can be considered as strategic tools, providing competences referred to the ability to search, process and analyse information, and identify, pose and solve problems.
Since the skills in the use of Information and Communication Technologies (ICT) are increasingly important and the educational software plays an outstanding role in the teaching-learning process, in this paper we focus on the adequacy of software packages, proposing the "Four-F" test and analysing Gretl as a case study.
The Four-F test includes the requirements of Freedom, Flexibility, Functionality and Friendliness. In order to test these hypotheses empirical evidence is collected through online students´ surveys and the obtained results confirm that Gretl is perceived as a free, flexible and friendly software package, while some difficulties appear with regard to the most demanding functionality requirements.

Keywords: Software, test, Freedom, Flexibility, Functionality, Friendliness, FOSS, Gretl, Econometrics, online surveys, hypothesis testing

1. Introduction. The Four-F Teaching Test

The effectiveness of computer-based training has been confirmed by several studies and the evaluation of instructional software is a key success factor. More specifically, focusing on Statistics and Econometrics courses, several authors as Hsu et al (2009) and Sosa et al. (2011) pay attention to different attributes (such as technology types, student´s engagement, control over the learning process or feedback) that could account for differences in the effectiveness.

In a broad sense, three main factors can be considered in order to evaluate instructional software: accessibility, usability and understanding facilitation.

The **accessibility** requirement suggests the convenience of using free open source software (FOSS), whose licenses give users the freedom to run the program for any purpose, to study and modify the program, and to redistribute copies of either the original or modified program, without having to pay royalties to previous developers. During the last decades FOSS has risen to great prominence and its advantages have been shown by many authors, as Murphy (1995) and Wheeler (2011).

Usability mainly refers to the quality of human-machine interface for its intended user. Then, in order to be considered "highly usable" a program should be easy to learn and easy to use, and both requirements must be evaluated through hands-on testing.

Finally, **understanding** is closely related to usability and, since the demonstration of this requirement is not easy, some authors suggest a list of abilities that might be helpful as an intuitive guide for determining the level of understanding. More specifically, this guide includes the ability of providing examples, explaining results, making analogies, repairing malfunctions or predicting the effects of a change.

An empirical investigation by Hsu et al. (2009) shows that both computer attitude and statistical software self-efficacy have significant positive effects on students´ perceived usefulness. In addition, it was also found that perceived usefulness and perceived ease of use positively influence learners' intentions to use statistical software, whereas their anxiety with statistics has a significant, negative impact on perceived usefulness, perceived ease of use and behavioural intentions.

Testing accessibility, usability and understanding facilitation is not an easy goal, suggesting the need of a more concrete specification of the required conditions. With this aim, in the next section we describe the "Four F" test including the hypotheses of Freedom, Flexibility, Functionality and Friendliness.

The implementation of this test on Gretl is described in section 3 and section 4 summarises the main empirical findings. Finally, the paper ends with some concluding remarks.

2. The Four-F Teaching Test

The need of testing the accessibility, usability and understanding facilitation of statistical and econometric software suggests the convenience of providing a more specific setting of these concepts. Following the approach represented in figure 1, we set the hypotheses of Freedom, Flexibility, Functionality and Friendliness, closely related to the previously described factors and easier to be expressed in a more "objective" way, allowing to perform the "Four-F" statistical test.

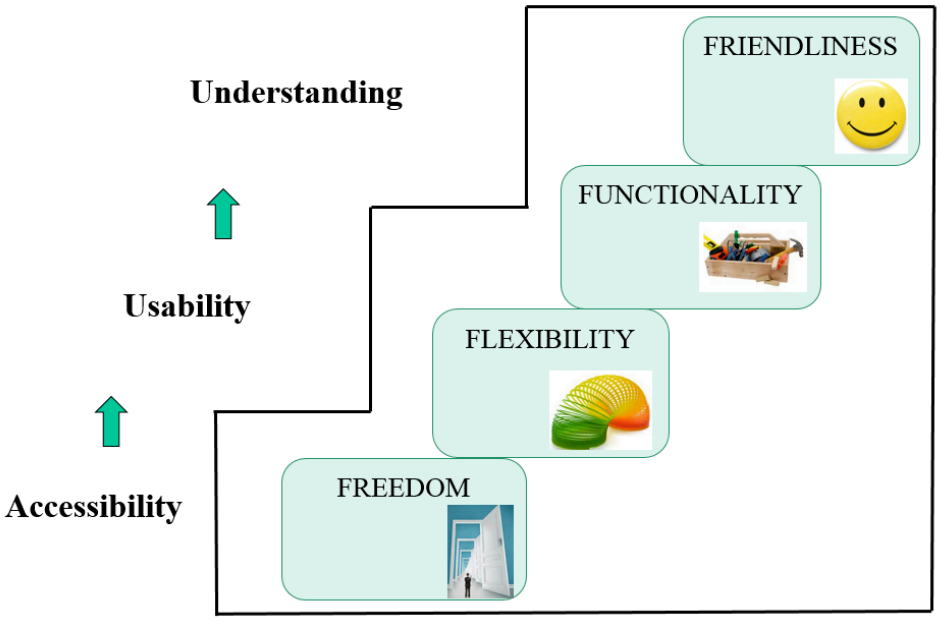

Figure 1. The Four-F tests

The first hypothesis is *Freedom,* mainly related to the access and also understood as openness. The inclusive term FOSS (Free Open Source Software) refers to software that is both free (libre) and open source, meaning that it is liberally licensed to grant users the right to use, copy, study, change, and improve its design through the availability of its source code. Thus, the "free" term refers to the freedom to copy and re-use the software, rather than to the price of the software, and it is quite easy to test.

The hypothesis of *Flexibility* is connected to both access and usability. Since flexibility refers to the ability to be easily modified, this condition should be understood in a broad sense including menu options, languages and many other software attributes.

Functionality refers to the quality of being suited to serve a purpose well. According to this third hypothesis, the considered software should provide a wide range of operations adapted to the needs of its potential users, thus facilitating the understanding of different concepts and methods.

Finally, last but not least, we set the hypothesis of *Friendliness.* Since the quality of being friendly affects the software capability to facilitate understanding, this requirement is particularly important for instructional purposes, and especially with regard to the less advanced users.

In order to provide a more detailed specification of these hypotheses, table 1 summarises the related requirements

Table 1. The Four-F Test Requirements

Hypothesis	**Requirements Description**
Freedom	Openness (FOSS)
Flexibility	Available in different languages and platforms
	Supporting different data formats
	Suitable for different potential users
Functionality	Providing a wide range of operations
	Interacting with different statistical software
	Allowing different levels of use
	Providing functional output
Friendliness	Intuitive and Friendly interface
	Understanding facilitation through graphs, tables, database, …
	Learning materials available

Although some of these requirements, such a those related to freedom and flexibility, can be easily checked on objective information, most of them should be tested according to users' subjective opinions, that can be easily collected through online surveys. With this regard, main attention should be paid to the survey design, including specific items referred to the different software options (functionality), their flexibility and their contribution to understanding (friendliness).

Regarding the methodology, the commonly used Likert scales allow to measure respondents' attitudes by asking the extent to which they agree or disagree with a particular question or statement and, if the typical five point scale is considered (1-strongly disagree, 2-disagree, 3- not sure/undecided, 4-agree, 5-strongly agree), the collected information allows the setting and testing of hypotheses related to the population proportion (H_0: $p>0.5$) and/or the population mean (H_0: $\mu>3$), leading to the corresponding critical levels (p-values) and conclusions.

3. Empirical Application to Gretl

The implementation of the Four-F test on Gretl starts by collecting objective evidences referred to the different requirements. Some of these characteristics have been stressed as Gretl´s advantages by Baiocchi & Distaso (2003), Mixon & Smith (2006), Yalta & Yalta (2007), Rosembladt (2008), Cottrell (2009), Falat & Panciková (2012), Cottrell & Luchetti (2014) and Adkins (2014), among others.

Table 2. The Four-F Test on Gretl. Objective Evidence

Requirement	Evidences
Freedom	Open-source statistical software Developed in Linux. Available on Microsoft Windows and Mac OS X
Flexibility	Available in several languages Wide variety of supported data formats Tailored for a wide variety of potential users
Functionality	Command Line Interface and Gretl Console Integrated powerful scripting language Import data from several file formats Interaction with R Output models as LaTeX files, in tabular or equation format Command loop structure for simulations and iterative estimation procedures User's Guide available
Friendliness	Intuitive and Friendly interface Graphs and Icons facilities Large database and textbooks available

Although freedom and flexibility can be mainly tested from objective information, our main aim is to analyse users´ subjective opinions in order to confirm if they perceive these advantages. Therefore, we have implemented online student´s surveys including three blocks of questions, summarised in table 3, respectively referred to the user`s personal characteristics, their level of agreement with specific aspects and their perception of Gretl´s ease of use and its comparison with alternative software. Furthermore, since some of the considered requirements are related to both functionality and friendliness, the items referred to the most intuitive facilities (icons and graphs) have been assigned to friendliness hypothesis while the remaining ones have been related to functionality.

Table 3. Online Questionnaires Structure

	Items	Hypotheses
Personal Characteristics	Level of expertise using Gretl (time) Language Operating System Location (home, work, university, …) Purpose (individual work, team project, research)	Freedom-Flexibility
Level of specific agreement	Workfile options Output facilities Console Functions User´s Guide and Help facilities	Functionality
	Icons Graphs	Friendliness
Level of global agreement	Ease of use Preference to alternative software	

The wide variety of advantages offered by Gretl make it a powerful tool for teaching econometrics at different levels, ranging from undergraduate to master and doctorate courses. Some teaching experiences with this software can be found, among others, in Mixon & Smith (2006), Pérez & López (2009), Lejnarová & Rácková (2009), López & Pérez (2011, 2012) and Falat & Panzicová (2012).

In order to approach the perception of a wide variety of students, encompassing different academic levels and ways of working with Gretl, this online survey has been implemented in several university courses, as summarised in table 4. More specifically:

- The compulsory second year course *Statistical and Econometric Methods* uses Gretl in the lab sessions, making an intensive use of several point-and-click menu options (probability, distributions, tests, sampling, models...) complemented with some specific functions (confidence intervals, combinatorial, …).

-The third year course *Econometrics* makes and intensive use of Gretl, using this software not only in lab sessions but also in the development of a team project, enabling students *learning by doing* Econometrics, as described in López & Pérez (2012).

-*Economic Forecasting* is an optional course providing a more specialised used of Gretl, mainly focusing on time series facilities (Filters, ARIMA, VAR, Cointegration ...). In this case students are asked to develop a personal project.

- The online course *Forecasting for Economics and Business* included in a Shared Virtual Campus uses Gretl as a main teaching tool for problem solving and assessment. This experience, described with more detail in López, Pérez & Moreno (2010), shows some differential characteristics referred to both the online teaching method and the diversity of students, coming from different universities, degrees and courses.

- Finally, Gretl has also been used in some post-graduate courses, including "*Time Series*" (Master in Finance, University of León), and "*Applied Econometrics*" (Ph D in Applied Economics, University of Oviedo) where the menu options are complemented with an intensive use of the Gretl´s console and some specific functions.

Table 4. Courses Description

Course	Description	Method	Gretl Use
Statistical and Econometric Methods	Compulsory, Second year, Degree in Economics	Blended Learning	Computer sessions Team Project
Econometrics	Compulsory, Third year, Degree in Economics	Blended Learning	Computer sessions Team Project
Economic Forecasting	Optional, Fourth year, Degree in Economics	Blended Learning	Computer sessions Personal Project
Forecasting for Business and Economics	Optional, Free-election course for different degrees of nine Spanish Universities (G9 Virtual Network)	E-Learning (G9 Shared Virtual Campus)	Online Materials Online Questionnaires
Time Series, Applied Econometrics	Master and Doctorate Courses, Economics and Finance (Universities of Oviedo and Leon)	Face to face	Practical Sessions Personal Project

4. Gretl´s Achivements and Challenges

The empirical evidence, based on a sample of 145 students and summarised in table 5, shows that Gretl clearly fullfils the requirements of flexibility and friendliness, while the conclusions referred to functionality differ depending on the considered items.

Table 5. The Four-F Test on Gretl. Evidence from online surveys

		Critical level (p-value)	
Hypothesis	**Item**	**H_0:p>0.5**	**H_0:μ>3**
FREE-FLEXIBLE	Use from different locations	1	
	Use for different purposes	1	
FUNCTIONAL	Agreement with workfile options	0.01	0
	Agreement with output facilities	1	1
	Agreement with Gretl console	0	0
	Agreement with Gretl functions	0.4	0.42
	Agreement with Gretl User´s guide and help facilities	0.5	0
FRIENDLY	Agreement with Gretl icons	0.98	0.99
	Agreement with graphs	1	1
	Ease of use	0.2	0.08
	Preference to alternative software	1	1

Regarding Flexibility, most students declare to use Gretl from different locations (home, school, work, …) and for several purposes (clases, autonomous work, team projects, …), thus failing to reject the null hypotesis p>0.5.

Furthermore, as table 5 shows, the functionality and friendliness requirements have been anlyzed through different ítems, running the corresponding statistical tests for both the population proportion (H_0:p>0.5) and the population mean (H_0:μ>3). The obtained results show that students strongly agree with the most intuitive Gretl facilities (icons and graphs), also confirming Gretl´s ease of use and considering this software better than another alternatives. These results are particularly important since, according to several studies, perceived usefulness and perceived ease of use positively influence learning.

With regard to functionality, students show a quite strong desagreement with the workfile options and Gretl´s console, leading to the rejection of the corresponding hypotheses. However, they agree with another functionality ítems such as Gretl´s output facilities, functions and –at least partially- User´s guide (in this last item most students agree to some extent, although the average score does not reach the required 3 points).

In general terms, the obtained results suggest that students consider Gretl a flexible and friendly teaching tool, while they face some functionality difficulties, mainly related to the most demanding options, such as using the console or working with datafiles (importing, compacting, sampling,…).

In order to analyse the robustness of these conclusions we have considered sub-samples of students with different levels of expertise, confirming the previous conclusions. However, we have found that the level of Gretl users expertise significantly affect their agreement with the funcionalitiy ítems, leading to a slight increase in the p-values.

5. Concluding remarks

In a context where educational software plays an increasingly outstanding role, this paper has tried to provide further evidence about some attributes that could significantly foster learning. With this aim we have focused on software access, usability and understanding facilitation, and we have developed the "Four-F" test, including the hypotheses of Freedom, Flexibility, Functionality and Friendliness.

This test has been implemented on the free-open source software Gretl, using both objective information and subjective users´ opinions collected from online students´ surveys.

The empirical evidence have helped us identifying the main weak and strong points of Gretl. According to the obtained results, Gretl fullfils the requirements of freedom, flexibility and friendliness, while some difficulties are found with regard to its functionality.

The online surveys show that the most positive scores correspond to some interesting aspects, such as the flexibility, the perceived usefulness and the perceived ease of use which are expected to positively influence learning.

On the other hand, since according to students´ perceptions the levels of desagreement are referred to the most demanding Gretl options (such as the console or workfile manipulation), further efforts should be made in order to facilitate their knowledge, also emphasizing their many advantages.

REFERENCES

Adkins, L.C. (2014). *Using Gretl for Principles of Econometrics*, Oklahoma State University, http://www.learneconometrics.com/gretl/using_gretl_for_POE4.pdf

Baiocchi, G, & Distaso, W. (2003). Gretl: Econometric Software for the Gnu Generation. *Journal of Applied Econometrics*, 18, 105-110.

Baytiyeh, H., & Pfaffman, J. (2010). Open source software: A community of altruists, *Computers in Human Behavior*, 26, 1345-1354.

Cottrell, A. (2009). Gretl: Retrospect, Desing and Prospect, In *Econometrics with Gretl. Proceedings of the Gretl Conference 2009*, Universidad del País Vasco, 3-13.

Cottrell, A., & Luchetti, R. (2014). Gretl User´s Guide. Gnu Free Documentation License, http://ricardo.ecn.wfu.edu/pub//gretl/manual/en/gretl-guide-a4.pdf

Falat, L., & Panciková, L. (2012). Introduction to Statistical Modeling in Gretl and its Applications in Econometrics. *Jounal of Information, Control and Management Systems*, 10 (2), 165-172.

Hsu, M.K., Wang, S.W., & Chiu, K.K. (2009). Computer Attitude, statistics anxiety and self-efficacy on statistical software adoption behavior: An empirical study of online MBA learners, Computers in Human Behavior, 25, 412-420.

Lejnarová, S., & Rácková, A. (2009). Useful Sofware for Econometric Beginners, In *Econometrics with Gretl. Proceedings of the Gretl Conference 2009*, Universidad del País Vasco, 179-189.

Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 140, 1-55.

López, A.J., & Pérez, R. (2011). Blended Learning in Econometrics. Crossing borders between Learning and Doing, *China-USA Business Review* 10 (1), 73-80.

López, A.J., & Pérez, R. (2012). University Challenges and Free Open Source Software (FOSS). Opening Minds, Closing Gaps. *Proceedings of the Edulearn12 Conference,*

Barcelona.
López, A.J., & Pérez, R. (2013). Challenges and achievements of Econometric software. The "Four-F" test. *6th International Conference of Education, Research and Innovations*, ICERI, Sevilla.
Luchetti, R. (2009). Who uses Gretl? An Analysis of the SourceForge Download Data. In *Econometrics with Gretl. Proceedings of the Gretl Conference 2009*, Universidad del País Vasco, 45-55.
Mixon J.W. (2010). GRETL: An econometrics package for teaching and research. Managerial Finance, 36 (1), 71-81.
Mixon J.W., & Smith, R.J. (2006). Teaching undergraduate Econometrics with Gretl. *Journal of Applied Econometrics*, 21, 1103-1107.
Murphy, C.A. (1995). An Evaluation Format for "Open" Software Tools*, Computers in Human Behavior*, 11 (3-4), 619-631.
Pérez, R., & López, A.J. (2009). Teaching and Learning Econometrics with Gretl. Summarising some experiences. In *Econometrics with Gretl. Proceedings of the Gretl Conference 2009*, Universidad del País Vasco, 191-201.
Roseenblad, A. (2008). Gretl 1.7.3. *Journal of Statistical Software*, 25 (1), 1-14.
Schmid, R.F., Bernard, R.M., Borokhovski, E., Tamim, R., Abrami, P.C., Wade, C.A., Surkes, M.A., & Lowerison, G. (2009). Technology´s effect on achievement in higher education: a Stage I meta-analysis of classroom applications, *Journal of Computing in Higher Education*, 21, 95-109.
Sosa, G.W., Berger, D.E.; Saw, A.; Mary, J.C. (2011). Effectiveness of Computer-Assisted Instruction in Statistics: A Meta-Analysis, *Review of Educational Research*, 81 (1), 97–128.
Wheeler, D.A. (2011). How to Evaluate Open Source Software/Free Software (OSS/FS) Programs, Available online: http://www.dwheeler.com/oss_fs_eval.html, Last accessed on February 2015.
Yalta, A.T., & Yalta, A.Y. (2007). Gretl 1.6.0 and its Numerical Accuracy. *Journal of Applied Econometrics*, 22 (4), 849-854.