

Inforum IT in 2014: Demonstrating the Collection, Compilation, and Publication of Information

Ronald Horst and Troy Wittek*

September 2014

Abstract

This work reviews the use of key software programs that are provided by Inforum. A variety of demonstration routines are described that indicate fundamental and advanced techniques and suggest suitable applications. In particular, new routines are introduced to demonstrate capabilities for collecting data from spreadsheets and compiling them in databases. The data then are displayed in tables and graphs that are produced by the G7 software. Next, the routines demonstrate the creation of publication-ready spreadsheet tables, where these tables may be augmented by spreadsheet graphs. A variety of new spreadsheet capabilities are introduced. Finally, the table-making demonstration is continued to highlight the spreadsheet capabilities of the Compare software. In addition, progress toward a development of a new web site is summarized, particularly the online resources that relate to current and archived economic databases, software and documentation, and Inforum international partners.

Inforum research depends heavily on data and economic models. Specialized software is required to manage the data and to construct and assemble the models. Furthermore, additional software is required to present data and results in useful forms. Inforum has developed such software that is well-suited for use to build macroeconomic and macroeconomic-interindustry models and to develop and manage the associated time-series data.

Origins of this work can be traced to the 1960s. Currently-employed software has taken shape over the past decade or two, but it steadily evolves to meet new requirements. This report describes recent developments to the primary tools employed at Inforum and by others users around the world. In addition, development of the Inforum web site also is provided. The web site provides an important means of distributing economic data and software, and it often is the means by which research sponsors and other parties first learn about the

*Horst: Inforum - University of Maryland, Department of Economics, Tydings Hall, Ronald.Horst@gmail.com.

organization. The web site also evolves steadily, and important additions are noted below.

1 Software Documentation

Software documentation has been consolidated from a variety of sources. Some of that documentation has been maintained and distributed in recent years, and other material long had been forgotten. Previous work was in a variety of forms, with little standardization and no consistency.

All of this material has been converted for use in a new documentation system. So far, material for *G7*, *Compare*, *Build*, *IdBuild*, *Fixer*, *MacFixer*, and *Banker* have been incorporated. The most obvious benefit is the improved appearance of the materials. A less evident benefit is that maintenance and extension of the documentation now is easy, where maintenance of previous materials sometimes was difficult.

The documentation system allows several types of files to be created. We currently have produced three forms. First is a document in book layout that is stored in the PDF file format. This format is best for printing or when reading the complete document. The second form is HTML. These files are available on the Inforum web site and may be viewed with any web browser. Finally, the complete set of documentation also is available in a compiled Help system that is linked to the latest version of *G7*. This compiled Help file relies on a new system that fully is supported by Microsoft Vista, Windows 7, and Windows 8. Earlier versions of *G7* employed a Help system that no longer has built-in support by these operating systems.

All documentation types have full indexing and cross referencing, though work on references remains incomplete. Search is available in each document format.

New material has been added for recent changes to the software. Some particularly outdated material has been replaced or heavily revised, and minor editing has been completed. Because all documentation now is in one system, and because that system is easy to use, we expect that many more improvements will be completed in the next year. In particular, text may be borrowed from other sources to update or extend the current work.

2 www.inforum.umd.edu Web Site

The principal addition to the web site is the software documentation discussed earlier. Other portions of the site are updated maintained regularly. These include the news page, the EconData repository of economic databases, and the software distribution site. New pages have been added for the Nineteenth Inforum World Conference and the 2011 Inforum Outlook Conference. The collection of Inforum publications has expanded to include papers from the Nineteenth Inforum World Conference, sponsored research publications, and updates

of Clopper Almon's The Craft of Economic Modeling.

Several tools have been used to detect problems and to improve reliability of the site. These include Google's Webmaster Tools and <http://linkchecker.sourceforge.net/>. These tools enabled the detection and ultimately the correction of many broken and outdated links, both links within the site and links to external locations.

Features of Google Analytics have been incorporated in the Inforum site to monitor demand for each web page and file and document downloads. A variety of information can be analyzed to learn characteristics of site visitors and traffic patters throughout the site. All data are aggregated so that nothing can be identified for particular visitors. Aggregate statistics are available for geographic location, type of browser and operating system, language, mobile devices, referring site, how much time is spent on the site, and whether visitors are new or returning to the site. Citations for Google Internet searches are compiled and listed by keyword.

Work to incorporate the tools of Google Anlytics remains incomplete, and so current monitoring of the Inforum web site does not provide a complete count of downloaded documents. In particular, papers and other resources that are listed on conference pages or international partners pages are not monitored yet, though they are monitored on the main documents repository. Once this work has been completed, we will be able to compile counts of all file downloads and page visits.

Other software from Google and other parties has helped us to discover and to fix many broken links and other problems, though more work remains.

Updates have been made to the International Partners pages using information provided by each partner. The collection of economic data known as EconData continually is maintained and gradually is expanding. A number of items have been added to the News page in the past year, and some of these have been featured on the Home page. Finally, the software Downloads, Documentation, and Deomonstrations pages have been updated with the materials reported here.

3 *G7*

G7 is used to construct and analyze data, to estimate econometric equations, and to build large-scale structural econometric models and report the results in text, graphs, and spreadsheets. We summarize here some of the most significant changes completed in the past year. See the *G7: New for 2012* page in the Help files for additional details.

The most obvious development in *G7* in 2012 is the new Help files that were described earlier. These resources may be launched from the main menu at the top of the main window, or they may be launched by pressing F2 in the editor window. The same material also has been added to theweb site to allow online browsing. The previous Help system was not supported by these newer operating systems, though Help viewers were available for Windows Vista and 7 so that the older *G7* documentation still could be used. The new system has

native support in recent operating systems, including Windows 8, 7, and Vista.

An experimental set of commands has been introduced that eliminate some of the limitations of the *G7* workspace bank. The most severe restriction is the limited number of series that may be included in the workspace. While the number is not fixed, users typically find that no more than 4,000 to 7,000 series may be stored. A second limitation is slow data access. Access is slow because all data are stored on the hard drive, and little is retained in memory. New commands allow all data to be read from the existing workspace into memory. Data may be modified and new series may be added. Experiments show that well over one million series may be added. The new restriction is that 32-bit programs are limited to two gigabytes of memory. Once the workspace grows sufficiently large so that *G7* requires two gigabytes of RAM, the routine will fail. An added benefit is that execution speed is greatly increased for scripts that rely heavily on the workspace. Note that other bottlenecks remain, such as frequent access of Vam banks or writing to Excel files. This version of the workspace cannot be stored to file in binary form. Two options exist. The first is to employ a companion routine that will store data from the workspace in text form. The second is to shift data from memory into the standard workspace file located on a hard drive. However, this operation will fail if the capacity of the standard workspace bank is exceeded. Additional testing and development is needed, but the routine already should prove useful for assembling large data sets, printing them to a text file, and then using *Banker* to build compressed binary banks for subsequent distribution and use.

3.0.1 Data Display Improvements

Improvements to the graphical interface have been made in several areas. First is with the show command. This routine is used to display vector and matrix data in a spreadsheet form. Users may paste or type data into the form. Once changes have been made, the user now will be prompted to save the work, or the user may save the changes by selecting the appropriate menu item. The show window may be closed by clicking the escape key. In the past, this would stop the execution of a *G7* script because the same shortcut is employed for that purpose. Behavior now has changed so that the show window may be closed by pressing the escape key, but a running *G7* script will be interrupted only if the escape key is pressed a second time.

A second set of changes was made to the look command interface. Startup performance has been improved; this especially is important for large compressed banks that have stub files with many thousands of lines. When a row is selected and clicked, data for the specified series are printed to the output window and displayed in a graph. The frequency of the series now is detected and the dates for printing and graphing now are adjusted automatically. This especially is convenient for banks that contain series of various frequencies.

A third set of interface changes was made to *G7* graphs. The layout of graph legends has been improved so that the spacing is even across the bottom of the graph and longer series names may be displayed. Up to seven series now may

be graphed reliably. The results of algebraic expressions may be displayed by providing an expression surrounded by parentheses in place of one or more of the series names that ordinarily follow the graph command. Finally, display of line styles has been improved, but this feature remains imperfect.

Finally, a series name or algebraic expression may be selected in the editor. By clicking the right mouse button, a menu will appear. By selecting “Type Data,” the data for the series or the results of the expression may be printed in the output window.

3.0.2 Excel Interface Improvements

G7 can read data from Excel spreadsheets, and *G7* also can modify or create new spreadsheet documents and fill them with data and formulas. The reading techniques provide key tools for the development of databases for EconData and other uses. The set of features again has been extended to allow improved formatting and graphing.

G7 now can freeze a spreadsheet at a specified location. This allows convenient scrolling within published documents while keeping series labels, dates, and other key information within view. Background colors now may be set for cells, and two or more cells now may be merged. Text and numbers now may be given subscripts or superscripts.

An important new feature is a tool to create charts in a workbook. The charts currently are in the form of new sheets that are added to the workbook. To support these graphing features, but also for use elsewhere, *G7* now can write a series of dates to a worksheet with the dates recorded in the Excel dates format. Including the series for the horizontal axis, up to thirty data series may be displayed. While time often is displayed on the horizontal axis, any other data series may be employed instead. While only a few features have been introduced so far, graph titles, subtitles, and vertical axis titles may be set through use of the existing *G7* commands. Additional capabilities will be introduced in the near future so that formatted graphs may be created that are ready for publication.

3.0.3 Miscellaneous Improvements

Strings, keywords, and a new set of functions were introduced about two years ago. Since then, this list of features has grown steadily and the tools have become more powerful and more reliable. Several new keywords have been added to allow the user to recover the settings of the format command. The function that is used most often to recover the definitions of named strings has been improved. A new operator, +=, was introduced to allow existing string definitions to be extended easily.

Finally, an “if” function has been added to allow the conditional execution of sections of text. If the condition is determined to be true, then one section of script will be employed, and an alternative section will be employed otherwise. For example, a particular data series may be employed under one condition, but

an alternative series may be used otherwise. While similar capability existed earlier with the *if* command, this routine allows routines to be built with fewer lines of code and less repetition.

Last year, the capability was introduced to record, display, and execute the commands that are typed into the *G7* command box. Those original controls were placed in the *G7* menu. Corresponding commands now are available in the *G7* scripting language.

Several of the @-functions have been improved. These include improvements to @mean, @stdev, and @ggr that allow greater control and better reliability. In addition, problems were fixed in the chain weighting routines, and now up to 1500 data series may be employed in the weighting process.

A variety of bug fixes and other minor improvements are listed in the “*G7*: New for 2012” page of the Help files. Notes with some additional information are available in the logs on the Download page of the Inforum web site. Other new features include the ability to clear user-defined functions, a new timer command.

The current version of *G7* may be found on the conference CD and on the Inforum web site. Full documentation may be found in the same locations, and the web site now provides online help.

4 Compare

Compare is a data presentation tool that principally is used to build tables and spreadsheets. Recent versions of *Compare* can create formatted spreadsheets in Excel format. Additional features gradually are being added to provide greater control over these documents.

Compare now can freeze a spreadsheet at a specified location. This allows convenient scrolling within published documents while keeping series labels, dates, and other key information within view.

Compare also can create charts in a workbook. The charts currently are in the form of new sheets that are added to the workbook. *Compare* now can write a series of dates to the worksheet with the dates recorded in the Excel dates format. These dates are used for the horizontal axis. Up to 29 other series may be graphed, including the alternative series that are displayed when multiple banks are loaded. Graph titles, subtitles, and vertical axis titles may be specified. A list of available graph types is provided in the *Compare* manual. Work so far establishes proof of concept, but more work is needed to provide tools suitable for professional use. Additional capabilities will be introduced in the near future so that formatted graphs may be created that are ready for publication.

The current version of *Compare* may be found on the conference CD and on the Inforum web site. Full documentation may be found in the same locations, and the web site now provides online help.

References

- [1] Ronald Horst. (2013) *Inforum Software Development in 2013*.
www.inforum.umd.edu/organization/conferences/iwcxxi/iwcxxi.html.
- [2] Ronald Horst. (2012) *Inforum in 2012: Information and Technology*. (Paper and slides)
www.inforum.umd.edu/organization/conferences/iwcxx/iwcxx.html.
- [3] Ronald Horst. (2012) *Inforum Software Development in 2011*
www.inforum.umd.edu/papers/conferences/2011/usa_horst.pdf.
- [4] Troy Wittek and Ronald Horst. (2010) *New Features of G7, Compare, and the Inforum Website* (Slides)
www.inforum.umd.edu/organization/conferences/iwcxviii/iwcxviii.html.