Guide to

EconData

A Source of Economic Time Series Data
from the University of Maryland
1993 April 23

This Guide tells how to use the ***EconData*** service of the University of Maryland. Several hundred thousand economic time series, produced by the U.S. Government and distributed by the government in a variety of formats and media, have been put into a standard, highly efficient, easy-to-use form for personal computers and made publicly available through this service. These series include various national accounts, labor information, price indices, current business indicators and industrial production, information on states and regions, and international data. A more complete list of the series covered can be found in the file CONTENTS.DOC. These series are all in the form of banks for the G regression and model building program and its public domain version, PDG. The DOS version of PDG and the banks can be downloaded from the University's computer systems by anyone connected to Internet or, more slowly, by telephone. The data have been put in this form by the INFORUM research group in the Economics Department. The dissemination by computer network is made possible by the Computer Science Center.

HOW TO GET THE DATA

EconData may be accessed in three ways:

1. Through Internet using an ftp (file transfer protocol) for "anonymous" login
2. Through Internet using a Telnet package to use the "info" interface (intended to make the downloading process easier).
3. By modem and telephone at 1200 or 2400 baud.

The first two are far faster than the third, so we begin by describing them.

<u>Internet with the anonymous login.</u>

Internet, is a computer network linking most of the major research universities in the United States and many other countries. If you have

obtained this document by way of Internet, you must know what is in this section. If, however, you obtained it by mail or by dial-up modem or through the University of Maryland's interactive INFO system, you will find the method described here vastly faster.

Exactly how one connects with Internet is different at each site and must be learned from the local network administrator. All the rest of how to use ***EconData*** is the same everywhere. Generally, a user at a PC connected to a local network that connects to Internet has a command to make the connection with "ftp" -- a File Transfer Protocol. Let us say that this command has the name "ftpcall", though it will almost certainly be something else. You would then type:

```
ftpcall info.umd.edu
```

That command connects you to the "info" machine at the University of Maryland. You will see on your screen:

```
user name:
ftp>
```

You then type:

```
user anonymous
```

You are then asked for a password. It will be your identification while connected, so your name would be a good choice. After giving the password, you again get the "ftp>" prompt. Reply with

```
binary
cd /info/EconData/Instructions
```

Do not forget the capital letters in "EconData" and "Instructions". Then, the first time you are connected, do:

```
get gbanks.doc
get guide.doc
get contents.doc
cd ../Tools
get pkz110.exe   (if you do not already have PKZip version 1.1 or higher)
get pkzip2.exe   (A newer version of PKZIP)
get pkunzip2.exe (A newer version of PKUNZIP)
get unzip.exe
get unzip.doc
get dezip20.com
get pdgarc.exe
```

```
get pdgsup.exe
bye
```

Those instructions will obtain the essential tools you need and a demonstration bank. After studying the "contents.doc" file, you can come back for particular banks. Let us suppose you decide that you want nipaq.zip, the quarterly national accounts. It is in the Data/USNational/Accounts directory, so you again repeat the above procedure as through the line "binary" and then do:

```
cd /info/EconData/Data
get nipaq.zip
bye
```

You can then follow the example below for using nipaq.zip

Internet via Telnet with the INFO interface.

A telnet program is one which provides terminal emulation over the Internet to a user, who can either be connected directly with a PC or via a mini or mainframe computer. As with FTP there are many versions of Telnet software and specific questions about its use should be directed to your system administrator. While the speed of transfer may of may not be affected by using Telnet rather than FTP, a connection made with Telnet has an advantage in that it offers the users full screen menus. These menus significantly improve the user friendliness of the INFO bulletin board and ***EconData***, especially for first time users.

For the purposes of this explanation, we will assume that your Telnet program is simply called telnet (as it often is). The command to issue Telnet should be something like this:

```
telnet info.umd.edu
```

When connected you should be given a login prompt, such as:

```
login:
```

you should reply:

```
info
```

This login connects you to a more user-friendly interface which allows you to view the documentation files before the downloading. Learn what the

various commands available to you do by using the arrow keys to select the command and then tapping the '?' key. You first want to "View" ***EconData***. Then you want to "X-fer" the files listed above. When you choose "X-fer", you will be asked to choose the type of transfer protocol. You will probably want to choose either the FTP or TFTP protocols, the choice of which will depend upon on your particular Telnet program.

Telephone and modem connection.

If not connected to internet, you can use your modem for a dial-up call to download some of the files. (Some data files are so large that this method of downloading may require more than the maximum 90 minute connection allowed by the system.) Follow these steps:

1. Start your communications program. Set the terminal emulation as vt100 if possible, otherwise note how it is set. Dial 301-403-4333 with the following communications parameters: baud 2400 if possible, otherwise 1200 or 300; 8 data bits, no parity, 1 stop bit.

2. After the terminal is connected, tap the 'Enter' or 'Return' key every second until the annex prompt appears.

3. At the "annex" prompt, type
 telnet info.umd.edu

4. At the "login" prompt, type
 info

5. To the question about terminal type, answer "vt100" or the other type which you set your communication program to emulate.

6. Learn what the various commands available to you do by using the arrow keys to select the command and then tapping the '?' key. You first want to "View" EconData. Then you want to "X-fer" the files listed above. You will want to use the Kermit protocol.

7. To leave the system, select Quit on the menu. That will take you back to the "annex" prompt. Reply "ha" to this prompt to "hangup".

THE FIRST FILES YOU NEED

The very first time you connect, you should transfer the following files:

In the **Instructions** subdirectory:

README.DOC Brief notes on recent changes in ***EconData***

CONTENTS.DOC A description of the subdirectories of ***EconData*** and the files contained in those subdirectories

GUIDE.DOC The current version of this document. If the date at the top is the same as the date of the file, you presumably do not need the file, since you are looking at it. This document, explaining how to use these economic time series. The same material is available as a file for WordPerfect 5.1 as GUIDE.WP.

In the **Tools** subdirectory:

PDGARC.EXE The public domain version of the G regression package, necessary for using the data files.

PDGSUP.EXE The supplementary programs for the G regression package, used for making tables and manipulating data.

PKZ110.EXE The PKZip archival package. If the files you want end in ".zip", you need this program or some other compatible program to de-archive them. PKZip is shareware, and a modest registration is required for legal use beyond an initial trial period. We started off with DWC because it was public domain and out-performed the well-known PKArc. However, PKZip gives significantly greater compression, so we have gradually switched to it. As a postscript to this, we have now begun to use a newer version PKZIP2.EXE and PKUNZIP2.EXE. You may need this newer verion to unzip some of the Zipped files available on ***EconData***.)

PKZip and PKUNZip are the packages that we used and tested to create the archive you see here. However, we have also made available two public domain programs that were designed to strictly de-archive Zipped files. We have made these available to you if you want to save on license fee, but to you, but we have not fully tested them. If you are interested, you might be able to use the following programs instead of PKZip:

UNZIP.EXE
UNZIP.DOC

DE-ARCHIVING

Except for a few ASCII files ending in the ".doc" extension, all files on EconData have been archived with PKZIP and have the ".zip" extension. They must be de-archived before they can be used.

For PKZIP, the process is slightly different. At the DOS prompt, just type

PKZ110

and the one file will explode into a number of files, one of which is PKUNZIP.EXE and others are documentation and registration information. To de-archive the file xxxx.zip, just type

```
pkunzip xxxx
```

The commands for the public domain alternatives to pkunzip are very similar. You should refer to the accompanying documentation.

PUBLIC DOMAIN G

All of the time-series are in the form of banks for G and PDG. PDG is contained in the file PDGARC.EXE, and important supplementary programs are in PDGSUP.EXE. These are "must" files for the initial download.

Once you have downloaded these files and you are back on your own computer, create a pdg subdirectory by typing:

```
c:
cd \
md pdg
cd pdg
```

Now move or copy to this directory the PDGARC.EXE, and PDGSUP.EXE files. De-archive PDG by typing:

```
pdgarc
pdgsup
```

You should now backup the PDGARC.EXE, PDGSUP.EXE files to diskette. You may then delete PDGARC.EXE and PDGSUP.EXE from your hard disk to save space.

Now, edit your autoexec.bat file to include c:\pdg in the path. Also, your config.sys file should have the lines

```
FILES = 20
BUFFERS = 20
```

The "20's" may be replaced with larger numbers. Probably your config.sys file is already fine; if not, modify it and reboot. Now it is time to try out PDG. Do:

```
cd \pdg
pdg demo
```

You will get a demonstration of PDG's functions.

USING PDG WITH THE ECONOMIC TIME-SERIES DATA

Once PDG is in the path, you can access it from any subdirectory. Now you simply need to download a time-series of interest from the Data subdirectory in ***Econdata***.

For illustration purposes, we will assume that you have downloaded the quarterly National Income and Product Accounts time-series, NIPAQ.ZIP. On you own computer, you need to set up a subdirectory for this data. This is done by typing:

```
cd \
md nipaq
cd nipaq
```

Now copy or move the NIPAQ.ZIP file to this directory, and type:

```
pkunzip nipaq
```

You may then delete NIPAQ.ZIP. Note that if you are using one of the public domain alternatives to PKUNZIP, then you would replace that programs name for "pkunzip" in the above example.

Each time-series has its own G configuration file, G.CFG, specific to that series. Provided that each series is kept in a different subdirectory, then when you start PDG, the G.CFG file will automatically assign this bank. You start PDG by typing:

```
pdg
```

Answer the opening question with a 'Enter', and then, at the : prompt, do:

```
look
```

and you may now explore the quarterly NIPA accounts.

Of course, it is possible to have all of the time-series data in one subdirectory; but then you must specify which bank you would like to look at once you are in PDG. You do this at any : prompt by typing:

```
hbk nipaq
```

If you are using PDG in another directory, you can assign this bank by the G or PDG command:

```
hbk \nipaq\nipaq
```

These steps should get you started. The "demo" will introduce you to the basic functions of PDG; the "help" files are extensive and may be printed out. There are also interactive lessons in the material you have downloaded. The book "*The Craft of Economic Modeling*" by Clopper Almon (Ginn Press) gives many examples of the use of G.

AVAILABLE DATA

Read the file CONTENTS.DOC for a list of currently available data.

COMMON QUESTIONS

Q. Do I have to download a whole huge file every time it is updated?
A. No, just download a small file with the updates and run a simple program to update the old large file from the small new one. Suppose you have downloaded nipaq, the quarterly national accounts, and a new quarter rolls around. You will then soon find a file in the Data directory named newnipaq.zip. Download it to the same directory with your nipaq bank. Then do

```
hsplice nipaq newnipaq revnipaq
```

SPLICE is a program in the pdgsup.dwc file which you have installed into the pdg directory. This command will create a new compressed bank named revnipaq. Start PDG and assign it with the command

```
hbk revnipaq
```

Then examine a few series to insure that everything has worked properly. Then exit PDG, delete nipaq.hin and nipaq.hbk and do

```
ren revnipaq.* nipaq.*
```

Your nipaq bank is now updated.

Q. Can I download an individual series?
A. No. Only whole banks can be moved. But whole banks now move faster than individual series did a few years ago.

Q. How large are the banks and how long does it take to download one?
A. A typical bank is 200 - 400K bytes, though some reach 2 or 3 megabytes. We experience a file transfer rate of about 5 or 6 K bytes per second in downloading via Internet from systems at other universities. (The local rate is some 45 K per second). Thus, the complete annual national accounts, 5,600 series, each back to 1929 or whenever it begins, occupies 425 K bytes and would require about 75 to 90 seconds. Since downloading a 1.44 Megabyte diskette of data requires only about 3 minutes, there is little reason to mail diskettes among sites connected to Internet. Download by dial-up telephone is much slower. The theoretical maximum rate is .24 K per second, so a file will take 20 to 25 times as long by dial-up as it does via Internet.

Q. How current are the data?
A. Banks which we can update from the Commerce Department's Economic Bulletin Board (EBB) are updated within a few days of release of new data. These include the quarterly National Income and Product Accounts, the Business Conditions Indicators, the Blue Pages of the Survey of Current Business, and the Indexes of Industrial Production. The Federal Reserve cooperates by lending us a Flow of Funds tape when it appears, so this bank is also very up-to-date. Other data we have to purchase. In particular, the Bureau of Labor Statistics has explicitly refused any cooperation other than selling us the tapes. Consequently, these files are updated only when we have a particular need to do so. That is at least once a year but seldom monthly.

Q. What is the advantage of ***EconData*** over EBB?
A. The Commerce EBB is primarily composed of "documents," often press releases, not banks. Almost every file is in a different format. There is no standard way to get the data into an analytical program. Nor is there any standard way to update a previous file with new information. On EconData, every bank is in the format of a G data bank. Updating is easy. To update our banks from the EBB, we have had to write C-programs -- a different one for every EBB file -- to put the EBB material into a standard form. Writing these programs was a tiresome bother. We have put them in the "Tools" directory of ***EconData***, in case we should not be quick enough for you in getting the file ready some month. We have urged the EBB staff to at least offer our banks to their users. They have declined to do so because they accept only files from U.S. Government agencies. We then urged agencies to use these banks to submit data to the EBB. The Federal Reserve submits the industrial production indexes in the form of an ASCII file for input into PDG. The staff of other agencies, though expressing personal admiration of the programs, have been apprehensive that making the data so easy to use might constitute competition with private industry.

Q. How hard are the data to use, once I get them?
A. If you are using PDG or G as your main analysis program, you have only to type a three-letter command plus the bank name to assign the bank. Then the "look" command allows you to search for the series you want, graph it or display it numerically. A brief command, followed by the names of the variables you want, will write the series to ascii files or 1-2-3 worksheet files. Nearly any program can read them in this form.

Q. What else do PDG and G do?
A. PDG will do algebraic and functional (such as logarithmic) transformations of series and convert a series from monthly to quarterly or from quarterly to annual periodicity -- and vice versa, approximately. You can bring in your own data and make banks that have series selected from your own data and one or more of the EconData banks. PDG will graph data in a variety of ways -- lines (up to seven) with user-controlled line size, style, and marking, multiple-scale graphs, bar graphs, stacked bars, parallel bars, scatter graphs, scatter graphs with connected points, semi-logarithmic graphs, and so on. It allows on-screen annotation of graphs and can "remember" the annotations that were made on screen so that when the same graph is repeated with updated data, the annotation will automatically appear. It will do screen dump prints to most printers. (Full G also supports PostScript printers, TIFF files, and HP- compatible printers with downloaded fonts.) A companion program makes tables that can include growth rates and annual averages of monthly or quarterly data. PDG does ordinary least squares regression, regression with "soft" or "stochastic" constraints, and distributed lags via soft constraints. Full G, for $95, adds Hildreth-Lu correction for autocorrelation, ARIMA methods, seemingly unrelated regression and stacked regression with constraints across equations, non-linear regression, automatic Chow tests, and pooled time-series and cross-section regression. With its sister program Build, it can build simultaneous equation models with hundreds of equations. Both G and PDG have extensive on-line documentation. There are versions of both programs in French, Italian, Spanish, Russian, and Chinese. The Italian version is being officially used in data dissemination.

Q. What is the form of the data banks? How do they pack so much data into so little space? Can they be used without PDG or G?
A. The precise form is described, byte-by-byte, in the gbanks.doc file in the Instructions directory of EconData. While one cannot simply look at the banks with the DOS "type" command or a text editor, there would be no problem for authors of other data handling programs to read the G banks. Basically, for each series, the frequency, starting date, and number of decimal places are recorded. The decimal point is then shifted to the right to give a sequence of integers, and the first differences of

these integers are taken. If the differences are all between -32K and +32K, then the series can be stored with perfect accuracy as an initial value and then these differences, stored as 2-byte integers. Otherwise, the storage defaults to 4-byte floating point. G and PDG can use the series in this form; they reconstitute each series as needed. For transmission, the resulting banks have been further compressed by putting them through PKZIP, data compression and archiving programs. PKZIP are available in the "Tools" directory of ***EconData***, along with some public domain programs which seem to be very good at unzipping ZIP files.

Q. How is ***EconData*** supported?
A. The INFORUM research group has developed many of these data banks for its own use in building the most comprehensive interindustry models of the U.S. economy. It has a PC with a 9-track tape drive, the essential piece of hardware to bridge the gap between the PC world and the 9-track tape world which most government data suppliers still live in. Since INFORUM is not in the business of selling data, it has been willing to make these banks available to other academic researchers through Internet. Several other groups have asked INFORUM to put particular bodies of data into this form and have agreed to let the result be placed on ***EconData***. Other groups at the University of Maryland have also contributed interest and support. Further contributions -- particularly if already in G-bank form -- would be most welcome.