The Linux Operating System:
Debian GNU/Linux

James G. MacKinnon

Department of Economics
Queen's University
Kingston, Ontario, Canada
K7L 3N6

E-mail: jgm@qed.econ.queensu.ca
Home page: http://www.econ.queensu.ca/pub/faculty/mackinnon/
Phone: 613 533-2293   Fax: 613 533-2257

I am grateful to Dirk Eddelbüttel, John Maheu, Bruce McCullough, Danny Quah, Colin Telmer, and several Debian users and developers for comments on earlier drafts.

April, 1999

1. Introduction

The Linux operating system has become the leading alternative to the operating systems from Microsoft Corporation that are installed on almost all new personal computers which use x86-compatible microprocessors. It provides an excellent environment for doing econometric research. In this review, I will explain what Linux is, discuss its advantages and disadvantages from the point of view of an econometrician, and provide a number of references. Many of the references will take the form of "universal resource locators," or URLs, which can be accessed using a web browser. To save space, I will omit the prefix http:// from these URLs.

Linux is probably the best-known example of "open-source" software, programs for which the source code is freely available on the Internet. Open-source software is free in the sense that it can be obtained without payment, and it is free in the sense that users are allowed to modify it, but it is not free in the sense that anyone can do whatever they want with it. Almost all open-source software is copyrighted by its authors. It is released under a variety of different licenses. The most commonly used is the "GNU General Public License," or GPL, of the Free Software Foundation (www.fsf.org). The key feature of the GPL is that software released under it must always remain open. Anyone is free to modify GPL software, but if they distribute the modified software, they must make the source code available under the terms of the GPL. The Linux kernel and other key components of the operating system are released under the GPL, but many components are released under other, generally similar, licenses.

Linux is named after Linus Torvalds, who started working on it in 1991 when he was a graduate student in Helsinki. It has thousands of authors, who collaborate via the Internet. Technically, Linux just refers to the core of the operating system, the so-called "kernel," which interacts directly with the hardware and supervises the operation of other programs. However, a working Linux system includes many other components, without which it would not be very useful. Many of these are GNU programs from the Free Software Foundation, while others were developed from work originally done at universities such as MIT and Berkeley. The GNU programs include the **gcc** compiler, for C and C++ programs, the **bash** shell, the **emacs** editor, the **gzip** compression utility, and the GNU versions of standard Unix utilities such as **tar** and **make**. In most cases, these GNU programs have more features and better performance than commercial equivalents.

A working Linux system consists of a great deal of software, of which the Linux kernel is only a small part. All of this software is normally packaged as what is called a "distribution." There are several popular distributions. This review concentrates on Debian GNU/Linux (www.debian.org), the one with which I am most familiar. In part because it comes with more useful packages than most other Linux distributions, this distribution is particularly suitable for econometricians. Debian is entirely non-commercial; it is maintained by hundreds of volunteers from around the world. The current version at the time this review was written is Version 2.1. Other popular distributions are commercial. They include Red Hat (www.redhat.com), Caldera (www.caldera.com), Slackware (www.slackware.com), and SuSE (www.suse.com); a more complete list may be found at www.linux.org. Each of these distributions may be particularly suitable for some users. The prices of commercial distributions vary greatly, from less than US $5 to more than US $100. The money pays for one or more CD-ROMs, probably a book of some sort, sometimes some commercial software, and, for the most expensive distributions, a modest amount of technical support. It is not necessary to pay anything at all, since anyone with a fast Internet connection can download Debian, and several other distributions, for free. Most software for Linux runs on all modern distributions, and most of this review, with some exceptions that are noted, applies to all of them.

As its name suggests, Linux is, for all practical purposes, a version of Unix, and modern versions of Linux are designed to be POSIX compliant. From the point of view of an experienced Unix user, therefore, it is no more different from most proprietary versions of Unix, such as Sun's Solaris and IBM's AIX, than the various proprietary versions are from each other. However, Linux is, in my experience, far easier to use "out of the box" than any proprietary version of Unix, partly because it comes with so many useful programs already installed. Most of these, including all the GNU tools, can be installed on any version of Unix, but doing so generally takes time and effort.

For those who are only familiar with contemporary, commercial, desktop operating systems, Linux may at first seem somewhat strange. The primary user interface for many users is the command line of the **bash** shell, multiple copies of which may be running at once in different virtual screens and different windows under the graphical X Window System, or X for short. The **bash** shell is extremely powerful. It offers command completion, filename completion, and numerous ways to recall and edit previous commands, along with powerful programming capabilities.[1] Several other shells, notably **tcsh**, **zsh**, and a free version of **ksh**, are also available as part of Debian GNU/Linux.

Not all Linux users like to use the command line as their primary user interface. Several "point and click" interfaces are also available. It appears that two of the most sophisticated ones, which are under active development, are likely to become *de facto* standards The first of these, **KDE** (www.kde.org), is starting to come into widespread use. The second, **Gnome** (www.gnome.org), has recently been released in a first version that is reputed to be quite buggy. It is possible to run many DOS programs using the **dosemu** package (www.dosemu.org) and some, generally older, Windows programs using the **wine** package (www.wine.org), and it is also possible to run many programs for the Macintosh using the commercial package **Executor** (www.ardi.com). However, commercial software for current Microsoft operating systems generally will not work on Linux. Therefore, moving from another operating system to Linux normally requires installing new software (much of it probably free) and learning new ways of doing things.

---

[1] Command completion and filename completion are useful features of modern Unix shells. The first word on a line is assumed to be a command, and subsequent words are assumed to be filenames. For example, if one types "tex" and then hit the [TAB] key once, **bash** beeps. If one hits [TAB] again, it lists the names of 11 commands that start with "tex". If one then types enough letters to uniquely identify any one of these commands, hitting [TAB] will cause the remaining letters to be filled in. This is an example of command completion. Filename completion works in roughly the same way. It is particularly useful when moving around in the filesystem. For example, one can type the command "`cd /usr/lib/texmf/web2c/`" by typing about nine characters and hitting the tab key several times.

2. Hardware Requirements and Installation

Linux is most commonly run on personal computers with x86-compatible processors (i386 or later) from Intel Corp. and other manufacturers. Versions are also available for several other instruction-set architectures, including Alpha, SPARC and UltraSPARC, later versions of the Motorola 680x0 line, PowerPC, and StrongARM. Linux generally uses available hardware more efficiently than modern, commercial operating systems. A 133 MHz. Pentium with 32 MB of RAM and a 1 GB hard disk should perform in a very satisfactory manner when running Linux, provided that it has an accelerated video card if it is running X. Even an old 486 with 16 MB of RAM can often run Linux well enough to be useful. Of course, with faster processors and more memory, Linux performs even better. In particular, adding more memory makes many disk-intensive operations much faster. Linux may be expected to perform extremely well on any currently available personal computer with supported hardware. Current versions can even utilize two or more processors, and the latest kernels (not yet the default in Debian) are reported to provide excellent support for multiprocessor machines.

Linux supports most personal computer hardware, but there can be problems with very recent hardware and with proprietary devices that use nonstandard protocols. Before one attempts to install Debian GNU/Linux on a particular computer, it is important to know what modem, network card, video card, sound card, and printer are part of the system. Unfortunately, it can sometimes be difficult to know whether a particular piece of hardware is supported by a particular version of Linux. One place to start is the Hardware-HOWTO, which is included with many Linux systems and can be found at www.metalab.unc.edu/mdw/HOWTO/.[2] Any device with "win" in its name, such as winmodems and winprinters, almost certainly will not work with Linux, because these lobotomized devices depend on special software drivers using secret protocols that work only with one popular operating system. Since it is sometimes hard to tell whether an internal model is really a winmodem, the safest approach is to use an external modem. In most cases, if a printer or other device is advertised as working with several operating systems, it will work with Linux.

A free version of the X Window System, which displays graphics on the screen, is included with all major Linux distributions. The program that actually interacts with the video card is called an "X server." There are several of these and, collectively, they support most, but not all, popular video cards. Most of the free X servers come directly or indirectly from the XFree86 project (www.xfree86.org), which frequently issues new releases that add new features and support additional hardware. In addition, some of the producers of commercial Linux distributions have released free X servers for particular video cards. Unfortunately, because the makers of video cards are constantly changing their specifications, recently introduced cards are often not supported by a free X server. Commercial X servers from Xi Graphics (www.xigraphics.com) and Metro Link (www.metrolink.com) may support recent cards for which no free servers are yet available, and they frequently offer performance and features which are not available in the free servers.

Linux often works better with "generic" PC hardware than with systems from big-name computer makers, because the latter sometimes use nonstandard hardware that can

[2] A HOWTO is a document that explains how to install or configure some aspect of Linux. The HOWTOs evolve over time, and some of them can be very useful.

cause problems. A growing number of relatively small companies sells computers with Linux preinstalled. Some of the bigger ones advertise in the print magazine *Linux Journal* (www.linuxjournal.com). Others can be found via links on web sites such as www.linux.org. Recently, IBM, Dell, and several other large computer makers have announced that they will offer support for Linux on selected servers and workstations, and some of the announcements suggest that they also plan to support it on selected desktop and laptop machines.

Installation of Debian GNU/Linux is quite easy if all hardware is properly identified and supported, although it is wise to obtain assistance from someone who has gone through the process before. Installation of future versions is likely to be easier, as a good deal of effort is currently being devoted to simplifying the process. There are several methods of installation. For modern computers that can boot from a CD-ROM, the entire installation can be done using the CD-ROM drive. Otherwise, one boot diskette will be needed to start the process. After the initial installation of a very basic system, the rest of the system can be installed over the network. Since Debian GNU/Linux is open-source software, anyone is free to make CD-ROMs, and many vendors do so; see the Debian web site for a long list of vendors in numerous countries. Detailed installation instructions are available on the web site and on every CD-ROM.

The various distributions of Linux differ substantially in the way they deal with installation. The Red Hat, Caldera, and SuSE distributions are reported to have somewhat easier, but arguably less robust, installation procedures than does Debian. At the other extreme, the Slackware distribution has the reputation of being especially suitable for people who really want to understand how everything works, and not very suitable for anyone else. Sometimes, one distribution will fail to install correctly on a particular machine, while other distributions will install with no difficulty. The most common reason for this is that the drivers for a nonexistent device interfere with the operation of a device that is actually present. If this happens, one can either try another distribution or use another Linux machine to compile a custom kernel with support only for the hardware of the target machine.[3]

It is quite common to run Linux and another operating system (OS) on the same computer. The easiest and safest way to install Linux on a computer that already has an OS on it is to put Linux on a second hard disk. The **lilo** boot loader, which comes with every Linux distribution, is one of several programs that can be used to boot whichever OS is desired. Linux is also perfectly happy to share a hard disk with another OS, but it can be a little tricky to install two or more operating systems on the same disk. New users are advised to read the relevant HOWTOs, or other documentation, before attempting to do this. Since Linux can read and write to most of the different filesystems that have been used by Microsoft operating systems over the years (FAT16, VFAT, FAT32, and so on), it is easy to copy files back and forth. Normally, only one operating system can be active at once. However, a recently released commercial product called **VMware** (available from www.vmware.com) apparently makes it possible to run other operating systems as guests

[3] The kernels that come with Debian and other Linux distributions typically contain support for a great variety of hardware devices. A custom kernel is compiled to support only the hardware that is actually present. It therefore uses less memory and will not cause driver conflicts.

under Linux, thereby allowing one to use programs running under two or more operating systems at the same time.

Debian GNU/Linux makes it very easy to install, upgrade, and remove software packages. Version 2.1 includes over 2250 packages. In most cases, to install a new package, one simply becomes root (the system administrator), starts a program called **dselect**, points it at an ftp site or CD-ROM (in practice, this step can often be omitted, since it only needs to be done once), chooses what package(s) to install, and hits the [ENTER] key a few times. If the chosen package requires other packages, **dselect** will point this out and offer to install them too. Sometimes, the package installation routines will require the user to answer a few questions. Although **dselect** generally works well, it does take a while to get used to it. A replacement called **apt** that will be easier to use and even more powerful is currently slated to be available in Version 2.2. Packages can also be installed directly from the command line by using the **dpkg** command and the new **apt-get** command. The latter can automatically download requested packages and their prerequisites.

The ability to upgrade easily from one version of Debian GNU/Linux to the next is an important part of the system's design. In principle, one simply starts **dselect** as root, points it at an ftp site or CD-ROM with the new version, and lets it upgrade everything. Alternatively, one can issue two **apt-get** commands to perform a complete upgrade. In practice, however, upgrades tend to require a bit more effort than this, and it is highly advisable to read the instructions. If one wishes to upgrade the kernel and is using a custom kernel, as I normally do, the kernel will have to be recompiled and then installed separately. This is made quite easy by a Debian program called **kernel-package**, but a certain amount of understanding is required, and it helps to have done it before.

The **dpkg** program has many useful features. For example, it is easy to obtain a list of all the packages that have been installed, to find out what package owns a particular file, and to find out what files are owned by a particular package. It is also easy to create list of packages that can be fed into **dselect** so as to "clone" an existing system on another machine. Removing packages is just as easy as installing them, and the package manager makes it difficult (but not impossible) to remove packages that are essential for the operation of other packages.

3. Advantages and Disadvantages of Linux

Like several other modern versions of Unix, Linux is very stable and multitasks extremely well. In the absence of power failures (an uninterruptible power supply that will allow the system to run for at least fifteen minutes without external power is a very good idea), hardware failures, or extremely ill-behaved programs, a Linux system should never crash. Uptimes of many months are common. I routinely run computationally-intensive jobs for days or weeks at a time in the background on modern computers running Linux. Provided these jobs are given a low priority (using the **nice** command) and do not use so much memory that the system begins to swap, they have no impact on the responsiveness of the system. One can use the system for hours without noticing that a big job is running.

In addition to being a true multi-user operating system, Linux works well across a network. This means that Linux users can access their office machine from home, or from 5000 kilometers away, almost as easily as they can when sitting in front of it. For text-mode applications, the speed of the network is not very important. However, for graphical applications, which can also be used remotely, a high-speed network is essential. One big advantage of Linux, and other Unix systems, is that it is both practical and safe to give accounts on an office machine running Linux to graduate students and coauthors. This makes it easy to share data, programs, drafts of papers, and even CPU time. For example, if one person needs to run several big jobs and has colleagues whose machines are not being heavily used, each of their machines can be used to run one of the big jobs. This can be done manually or automatically. It is also possible to put together clusters of Linux machines using various types of networking hardware and the free **Beowulf** (www.beowulf.org) software. Some of these clusters appear on lists of the world's fastest supercomputers (for example, see www.top500.org).

Because it is free, and because it comes with so much free software, some of which will be discussed later in this review, Linux can be very much less costly than commercial alternatives. Moreover, the hassles associated with software licensing are completely avoided if one uses only free software. These hassles include the operational difficulties associated with copy protection, license managers, and license files, and the problems imposed by trying to abide by inflexible software licenses. In practice, of course, most people will use commercial software as well as free software, and so some of these hassles will be encountered with Linux, but generally to a much more limited extent than on most operating systems.

Linux makes an ideal operating system for servers. Even a machine that would be considered obsolete today can easily handle the electronic mail of a large University department and function as a web server. Setting up the popular **Apache** web server program on a Debian system that is already connected to the Internet takes about two minutes; of course, providing the content for the web site will take a great deal longer. A Linux machine running the **Samba** software package can easily and inexpensively provide file and print services to machines running Microsoft operating systems. Linux is also an extremely cost-effective way to provide students with access to computing services. One or more powerful machines, with modern processors and at least 64 MB of memory, can be configured as application servers and connected to a larger number of low-powered machines, with older processors and no more than 32 MB of memory, that are configured as "X terminals." Students can then access the Internet and make use of applications running on the servers, either graphically or in text mode, from the X terminals. This is a particularly attractive approach if some students need to run large jobs that take days or weeks.

Unlike commercial operating systems, Linux does not attempt to hide how it operates or limit what its users can do. It is almost infinitely customizable. A knowledgeable user who does not like the way something works can probably change it. In almost every case, the way Linux does things is logical, but learning to appreciate the logic can take time. Tools like **find**, **sort**, and **grep** are very powerful and have many options. Scripting languages like **Perl**, **AWK**, and **Python** provide the ability to perform extremely complex tasks. Even though detailed on-line documentation is available through **man** and/or **info**, few Linux users employ more than a small fraction of the system's capabilities. It is always fun to discover more of these as time goes by.

Many computers that run Linux are connected to the Internet, and this is certainly likely to be the case for computers in Economics departments. Any computer that is connected to the Internet is potentially in danger of being broken into. The default setup for a Debian system is reasonably secure, something that cannot be said for all Linux distributions. However, because Linux is so widely used, and because the source code is available, ways of compromising its security are found from time to time. Security holes are normally fixed very quickly, but it is absolutely essential to install security-related fixes as they become available. It is a very good idea to subscribe to the Debian security announcement mailing list.

In my view, Linux has only one significant disadvantage: Not nearly as much commercial software is available for it as for Microsoft's popular operating systems. In particular, not surprisingly, none of the popular applications from Microsoft is available. However, as I discuss in the following sections, quite a lot of excellent software for econometrics, statistics, document processing, personal productivity, and many other areas is available for Linux. Some of this software is free, and some of it is commercial. Moreover, both the quality of free software and the availability of commercial software is improving at a very rapid pace, and it seems likely that the recent surge of interest in Linux will cause a lot more commercial developers to port their products to it.

4. Econometric and Statistical Software for Linux

The best place to find out about technically-oriented software for Linux is the web site SAL.KachinaTech.com. It provides information about the vast majority of programs that might be of interest to econometricians. One of the strengths of the Debian distribution is that an extraordinarily large number of useful free programs is included as part of the distribution itself. As a consequence, these programs, a few of which will be mentioned below, are extremely easy to install and keep up to date. Installing and updating commercial programs and other free programs generally requires somewhat more effort.

Most of the classic, commercial econometric packages that have their roots in batch processing on mainframes are available for Linux. These packages include **Limdep** (www.limdep.com), **RATS** (www.estima.com), **Shazam** (shazam.econ.ubc.ca), and **TSP** (www.tspintl.com). The Linux versions of these packages generally perform exactly like the versions for other Unix operating systems, but they may not be as graphically-oriented as some other versions of the same packages.

Commercial programs of more recent origin, which tend to be more graphically oriented, are often not available for Linux. One important exception is **XploRe** (www.xplore-stat.de), the Linux version of which may, at present, be downloaded without charge. This sophisticated package, which is particularly strong in the area of nonparametric methods, has recently been reviewed by Teyssière (1998).

There are currently very few open-source packages for doing econometrics. Most programs that are freely available are very specialized. One exception is the program **tsrf** (http://econ.lse.ac.uk/staff/dquah/softw.html). Although it is not really a substitute for classic programs like TSP, it does have some of the features of such programs, it offers certain unique capabilities, and it is readily extensible in a way that commercial programs can never be.

There are several general-purpose statistical packages for Linux. Two notable commercial ones are **Stata** (www.stata.com) and **S-Plus** (www.mathsoft.com). There is also an impressive open-source package called **R** which is quite similar to **S-Plus** (both were derived from **S**) and is included with the Debian distribution.

Many econometricians use matrix programming languages instead of, or in addition to, conventional statistics packages. Here Linux offers several choices in addition to **R** and **S-Plus**, which have many of the capabilities of matrix languages. **GAUSS** (www.aptech.com), which is widely used in economics, is available, as is **MATLAB** (www.mathworks.com), which is widely used in a number of disciplines. Rust (1993) provides an excellent, albeit now somewhat dated, comparison of these two packages. Linux also offers free alternatives to both of these commercial packages, which are quite expensive. The package **Ox** (www.nuff.ox.ac.uk/users/doornik) is not open-source, but it is currently available at no cost, for educational and research purposes only, for Linux and several other platforms. Cribari-Neto (1997) compares the features and performance of **Ox** and **GAUSS**. The programs **Octave**, **Yorick**, and **Scilab**, which are available as Debian packages, provide no-cost alternatives to **MATLAB**.

Many researchers prefer to use conventional high-level languages instead of, or in addition to, matrix programming languages. For many languages, among them C and C++, Linux is an ideal platform. Debian GNU/Linux includes two versions of the **gcc** compiler, along with a number of valuable development tools, including the GNU **make** utility, the **ddd** and **gdb** debuggers, and the **gprof** profiler. When combined with freely available libraries for scientific computing and statistics, many of which are included in Debian, these tools provide an excellent environment for developing advanced programs in econometrics and statistics; see Eddelbüttel (1996). There are also several commercial C and C++ compilers for Linux.

Much of the free software for numerical computation that is available on the net is written in Fortran, and many econometricians (including the author) are more familiar with Fortran than with newer languages. Several Fortran compilers are available for Linux. Debian GNU/Linux includes two open-source ones, the **f2c** translator, which works transparently with **gcc** when it is called by using the **fort77** script, and the **g77** compiler. These both handle Fortran 77 code, with some extensions. The **g77** compiler, which is not yet finished, typically produces faster code than **f2c**, but it could be more reliable, although it is improving in this respect. At present, there are no open-source Fortran 90 or Fortran 95 compilers, but there are several commercial ones. A good source of information about Fortran compilers for Linux is http://studbolt.physast.uga.edu/templon/fortran.html.

Another type of software that is widely used by econometricians is computer algebra systems. A large number of these systems, including all of the major ones, is available for Linux. Among these are **Mathematica** (www.mathematica.com), **Maple** (www.maplesoft.com), and **MuPAD** (www.mupad.de). See Hutton and Hutton (1995) for a review of **Maple**.

In summary, even though the choice is somewhat more limited than it is for the most popular proprietary operating systems, a large and growing collection of econometric and statistical packages and programming tools is available for Linux, and many of them are available at no cost. When coupled with the many advantages of the underlying operating system, this makes Linux an excellent environment for theoretical and applied econometrics.

5. Other Software for Linux

Since most econometricians use the same computers for many purposes, availability of software for a wide range of applications is an important consideration when choosing an operating system. In this section, I briefly discuss several important categories of software for Linux.

For most researchers, software for document processing is absolutely essential. The best software for preparing technical documents is the open-source TEX system. Debian GNU/Linux comes with the **tetex** distribution of TEX, which has become the standard in the Unix world and includes a great many useful programs in addition to Plain TEX and LaTEX. Installing **tetex** is very easy. There are two essential packages and six optional ones that can be installed in the usual way, and it is advisable to run the **texconfig** program afterwards. When combined with a suitable editor (many of the editors that come with Linux support syntax highlighting for TEX), **tetex** provides a superb environment for writing technical documents. The **tetex** system has many useful features, including automatic generation of fonts at any magnification and automatic updating of the document display when the source is run through TEX. With most window managers, it is easy to arrange for several documents to be displayed in different virtual screens, a feature that I find invaluable when working on multiple versions of the same document.

Like a well-known, commercial, scientific word processor for Windows, the open-source program **Lyx** (www.lyx.org) provides a WYSIWYG front-end to LaTEX. This program has been developing rapidly, and recent versions are reputed to work well. The version included in Debian 2.1 is not entirely up-to-date, but a more recent version is included in the "unstable" distribution that will eventually become Debian 2.2. It can easily be installed on a 2.1 system.

Even though TEX produces much more attractive technical documents than any WYSIWYG word processor, some people prefer to use programs of the latter type, especially for short, non-technical documents. Three major commercial word processors are available. A Linux version of WordPerfect 8 has recently been released by Corel (linux.corel.com), in both a downloadable version that is free for personal use and a full-featured, non-free version. The Applixware office suite (www.applix.com) includes a WYSIWYG word processor with many features, a spreadsheet, and several other programs. The StarOffice office suite (www.stardivision.com) has a great many features and is reported to resemble Microsoft Office. Even for Linux users who rely on TEX, these commercial programs can be useful, because they are able to read documents in the proprietary formats of popular commercial word processors.

Although there are several free spreadsheets for Linux, none of them currently approaches the capabilities of the latest commercial ones. I have been very happy with **NExS** (www.xess.com), a modern, graphically-oriented, commercial spreadsheet that runs under X on Linux and on several other versions of Unix. A rather similar program, which has a common origin, is **XESS** (www.ais.com). These two programs have become much less expensive over the course of the past two years. The Applixware and StarOffice suites also include modern spreadsheets.

Several versions of the well-known web browser from Netscape (www.netscape.com) are available for Linux. Like their counterparts for other operating systems, these provide many features, tend to use a lot of memory, and occasionally crash (but without causing

Linux to crash when they do). In the spring of 1998, Netscape released the source code of their unfinished next-generation browser under a license similar to the GNU General Public License. It is expected that several new browsers based on this source code, both free ones and ones from Netscape, will become available during 1999.

There are many ways to send and receive electronic mail under Linux. These include Netscape Communicator, the **emacs** editor, and a wide variety of specialized mail clients such as **pine**, **elm**, and **mutt**. The choice is largely a matter of personal taste, although some mail clients may be more suitable for a machine that is permanently connected to a network and others may be more suitable for a machine that connects only intermittently. For advanced users, Linux offers many capabilities. For example, incoming mail can be filtered and grouped by using the **procmail** package.

6. Concluding Remarks

The Linux operating system provides an excellent platform for doing theoretical and applied econometrics. Debian GNU/Linux is a particularly good distribution of Linux for this purpose, because its package manager makes upgrading easy and because many programs of interest to econometricians are included as part of the distribution. However, it does take a while to become used to Linux, especially for those who have no previous experience with Unix. New users may wish to begin by installing it on an older computer, or as a second operating system on a desktop machine that will be rebooted from time to time anyway. Some econometricians will undoubtedly find that Linux suits their needs extremely well, while others may find that Linux as a desktop operating system is not to their taste. However, even those who are in the latter group will probably find that Linux is an ideal operating system for running web servers, mail servers, and compute servers.

Those who wish to learn more about Linux (and there is a lot to learn!) should look at some of the many books on the subject that can be found in any good bookstore. A search on www.amazon.com turns up well over 100 titles, not including general Unix books, many of which are relevant. Welch and Kaufman (1996) is a very good, although now somewhat dated, introduction to Linux, while Hekman (1998) is a useful reference. Many books come with a Linux distribution on a CD-ROM. However, it is not advisable to install Linux from such a CD-ROM, because these distributions tend to be old. It is much better to acquire an up-to-date set of CD-ROMs from one of the vendors listed on the Debian web site.

One very important feature of Linux has not yet been mentioned. Running Linux is fun! It brings back much of the excitement that used to characterize computing in days gone by. One reason for this is that there is a strong sense of community among Linux users. Another reason is that Linux users experience a sense of rapid progress: New software is constantly appearing, older software is constantly improving, and even the most experienced can always find new and better ways to use the enormously powerful tools that come with the system.

References

Cribari-Neto, F. (1997). "Econometric programming environments: GAUSS, Ox and S-PLUS," *Journal of Applied Econometrics*, **12**, 77–89

Eddelbüttel, D. (1996). "Object-oriented econometrics: matrix programming in C++ using GCC and Newmat," *Journal of Applied Econometrics*, **11**, 199-209.

Hekman, J. (1998). "Linux in a Nutshell: A Desktop Quick Reference," Sebastopol, California: O'Reilly and Associates.

Hutton, J., and J. Hutton (1995). "The Maple computer algebra system: A review," *Journal of Applied Econometrics*, **10**, 329–337.

Rust, J. (1993). "GAUSS and MATLAB: A comparison," *Journal of Applied Econometrics*, **8**, 307–324.

Teyssière, G. (1998). "XploRe 4.0, An interactive statistical computing environment," *Journal of Applied Econometrics*, **13**, 673–679.

Welch, M. and L. Kaufman (1996). *Running Linux*, Second Edition. Sebastopol, California: O'Reilly and Associates.