# A brief history of Stata on its 20th anniversary

Nicholas J. Cox
University of Durham, UK
n.j.cox@durham.ac.uk

## 1 Introduction

Notwithstanding good advice against gilding the lily, a capsule history of Stata might provide a useful perspective on the recollections and reflections given by those who were there at the beginning or in Stata's early years. The account here is based mostly on talks given by Bill Gould at Users Group meetings in London and Boston in 2004 and on conversations with him, other developers of Stata, and other users over the last decade.

The timeline of key events is given in the appendix as table 1; you might want to read through that now. Table 2 gives a compact list of Stata's release dates, the main stations on its journey.

## 2 Californian creation

Stata got started in California, in the mid-1980s. In 1984, a UCLA graduate named William Gould was part of a small but successful business selling computer time on mainframes called Computing Resource Center (CRC) and based in Santa Monica. At the time, personal computers (PCs) that could be taken seriously were being launched on the market, and it was clear that they were no longer just toys for hobbyists. Further, several people were writing statistical programs directly intended for PCs, such as Systat or MicroTSP. Stata was based on Bill's bold assumption, the notion that (with help) he could write a program that could stand comparison with any on the market.

Stata started life within the company as DiAL, a name taken from a never-released project; a logo existed and several spare folders were lying around the office. At one time, the name S was considered, until Bill Rogers pointed out that a group at Bell Labs had got there first. The name was changed to Stata just before release. In some early documentation, it was shouted out, all capitals, as "STATA", but the presently used form emerged quickly.

Bill Gould and Sean Becketti spent a year writing the first version of Stata. It was written, then as now, in the C programming language. Anyone who knows C will recognize that the syntax of Stata is influenced by the syntax of C, although many details (including preferences for lowercase names and terseness of output) were also generic to the Unix operating system. In that first year, much of the time was spent not writing code, but arguing about the design and trying very hard to get it right, thus establishing a tradition of discussing even many simple things like command names, until, ideally, a solution emerges by consensus among the developers.

Stata 1.0 was a small program that could not claim to cover all of even mainstream statistics, any more than its competitors did. It could more fairly be described as a regression package with data management features. Stata 1.0 existed only as a program for PCs running the DOS operating system. Most of its 44 commands (table 3) will look familiar to present users. Recently, there have been half-serious suggestions that Stata 1.0 should be reissued (apart from its bugs, naturally). At least the manual could be lifted without difficulty.

Stata 1.1 was a release to fix bugs, but 1.2 brought a new menu system and better online help. Then Stata 1.3 brought real graphics on IBM CGA and Hercules Monochrome cards, through an add-on called Stata/Graphics. In that version, a `program` command was implemented, although its existence was not revealed until later. Making Stata extensible, so that users could in effect add their own commands, and the software could begin growing from a package into a language, was undoubtedly one of the most important steps in its history.

For the first few years of Stata's history, there were releases of various kinds, both new versions and extra sets of Stata programs (sometimes sold as "kits", or later as Support Disks), every few months. Among the more notable additions were `anova`, `logit`, and `probit` in Stata 1.5. Some users had proved oddly resistant to an initial bias that ANOVA was just a roundabout way of doing regression.

A move beyond DOS to an extra operating system, Sun/Unix, came in 1988. 386/ix and HP/9000, DEC RISC, IBM RS/6000, DEC Alpha, and other variants of Unix (later of Linux) have followed. The first Stata for Macintosh (written by Bill Rogers) was released in 1992. Over the last decade, other versions for the Macintosh have been developed, overseen by Chinh Nguyen, most recently for OS X. Maximizing platform-independence has long been a major principle for Stata.

Stata 2.0 was a major new release, with revised graphics, now integrated into the Stata executable and not implemented through add-ons, and with string variables, survival analysis (Cox and Kaplan–Meier), and stepwise regression. Stata 2.05 saw the documentation rebound in a new format. Shortly afterwards was released Stage 1.0, an interactive graphics editor, written in one month and never updated. Much later, Stage was ported to Linux, but in other respects it was not modified. Many users would want to underline that StataCorp has yet to produce an interactive graphics editor for Stata 8.

The first Stata text was *Statistics with Stata*, by Lawrence Hamilton (1990) (see his contribution to this issue). This was another sign of maturity, and itself fed awareness and sales of Stata to mutual benefit. Teaching of statistics with Stata, or of Stata itself, has never been the first concern of Stata's developers, who see Stata as primarily software for research, but there has always been strong and positive support for people who write texts linked to Stata. Among many subsequent texts and teaching ventures, one that must be mentioned is StataQuest, a joint project with Duxbury Press. The accompanying text was written by Ted Anagnoson and Richard DeLeon; see also the contribution by Ted to this issue.

Later copies of Stata were sold in various countries outside the United States through distributors. Peter Hedstrom of Metrika Consulting in Sweden was the first Stata distributor, and Ana Timberlake of Timberlake Consultants in Britain was the second. The number now is much larger, as can be seen from the StataCorp web site.

In April 1990, CRC exited the mainframe service bureau business. The sideline product had become the mainline product.

Stata 2.1 saw other major additions: byte variables, factor analysis, and ado-files. `reshape`, previously distributed in a kit, also become integrated into Stata. The support for ado-files has been the most important of these additions over the longer term. That is, a program could be not only defined by a file but also automatically loaded into memory when the user needed it. Its addition at this time was in a way an accident, as Bill Gould explains in his interview.

Another venture at this time was the introduction of the *Stata Technical Bulletin* (STB) as a publication issued 6 times per year, initially under the editorship of Joe Hilbe (see his contribution to this issue). This continued publication (later under the editorships of Sean Becketti and Joe Newton) for just over 10 years (61 issues) and, in a strong sense, continues in its new guise as the *Stata Journal* you are reading. Table 5 gives a list of people who have served as members of the editorial boards of the two journals.

The STB, as printed journal plus diskette, served as a medium for publication of extra-official programs between releases, superseding the irregularly issued kits and Support Disks. Increasingly, and more importantly, it served as a medium for users to make public their own programs and thus to formalize the existence of a genuine user community. Researchers could not only benefit from others' programming labor, but also gain some credit from the publication of their own work. It is striking how many of the additions to "official Stata" (that which the company sells and supports) from the early 1990s had roots, directly or indirectly, in user-written programs. `glm`, to which both Joe Hilbe and Patrick Royston (who also contributes to this issue) made outstanding contributions, is just one example among several.

## 3    Texan transpose

In August 1993, the company moved to College Station, Texas, and became Stata Corporation. More recently, its name has been tweaked for legal reasons to StataCorp LP. In 2000, the company moved out of what had become very cramped quarters near Texas A&M University to a new building on its own campus. In 2005, a second building will be completed.

For a few years from Stata 3.0 on, the main development thrust was to extend the corpus of regression-like modeling commands. Among many others were added flavors of regression for limited, categorical, bounded, and multiple dependent variables, robust and quantile regression, and support for non-normal distributions. Central to this set of commands was an engine for maximum-likelihood-based modeling that emerged as `ml`

and thus permitted users to program their own modeling commands. Later came much more for panel data, time series, survey estimation, and cluster analysis (and much more yet).

A development with which the company had little to do, at least initially, was the Statalist listserver, started by David Wormuth in August 1994. It was based on hardware and software running at Harvard University. After a few years, Marcello Pagano took over moderation of the list, assisted by Nicholas Cox, who maintains the FAQ. Statalist quickly established its own traditions. It has been a most valuable forum for discussions about Stata, and statistical science generally, and has encouraged the writing and dissemination of many hundreds of Stata programs. Some were transient— but likely the solution to somebody's problem—but many have proved of widespread and lasting value.

Many emails included listings of programs, allowing interested users to copy and paste into their own files. Some of the longer programs needed to be split into multiple postings! Note that Statalist has never carried attachments. Installing programs in this way was bound to be onerous and error-prone. A key shift occurred in 1997 when Christopher (Kit) Baum volunteered to put Stata programs sent to him by users within the Statistical Software Components (SSC) archive he was maintaining. A central searchable archive of user-written programs has proved very popular. Although some users continue to make their programs available through personal web sites, the majority of user-written Stata programs in the public domain are to be found within the SSC archive. This was a development with which the company had initially nothing to do. Later, however, they added an `ssc` command to Stata based on programs by Baum and Cox, thus giving users a handle to install programs directly without the intermediary of a browser.

Stata 4.0 saw the first port to Microsoft Windows (3.1 to be precise). At the time, PC users were moving in large numbers from DOS to Windows, and there was a strong case that Stata should follow suit. What was more difficult for the developers to see was exactly how a windowed interface could be consistent with the command language that was so utterly central to Stata's design.

Various possibilities were entertained in an internal project, but the company would not market until they were happy with a new design. One day, Bill Sribney, then working as a developer, came into work with a vision of Command, Results, Review, and Variables windows. This immediately appealed, and it remains at the center of all Stata's windowed interfaces, even for Macintosh and Unix. Later have followed Statas for other flavors of Windows (95, NT, XP); here Alan Riley has overseen development.

Stata Users Group meetings began in 1995, the first being held in London and organized by Ana Timberlake. Table 4 gives a list of the meetings held as of the end of 2004. Quickly there emerged a pattern followed by all subsequent Users Group meetings. The logistics of times, places, registration, food and drink, and so forth are always handled by the distributor (in the United States, by StataCorp). The program is always handled by users. User presentations are typically combined with a "Report to users" from a senior Stata developer. The formal meeting always ends with a "Wishes

and grumbles" session (this was Ana Timberlake's expression, since copied across the world) at which users shout out requests and complaints, met with "Yes", "No", or "Maybe". In some recent meetings, there have been longer review talks and training sessions, which have proved popular.

It is difficult to encapsulate what Users Group meetings mean for those concerned. Evidently, Stata's work is highly computer-based, so that much interaction between (and among) users, developers, and technical support can be based on email. Equally evidently, there can be no real substitute for face-to-face interaction, and many important developments have grown from these meetings. Sophia Rabe-Hesketh spoke about `gllamm` at a Users Group meeting, thus launching one of the most versatile and widely used user-written programs. Users too report on exchanges, collaborations, and friendships that would not have arisen from disciplinary conferences, and they often remark that the meetings are more pleasant, and even more useful, than such conferences. Which disciplines? I could not possibly comment.

By industry standards, Stata was a bit slow to respond to the Internet and all the possibilities it created, but it can fairly claim to have caught up. The company's first net-based venture was into what were called NetCourses. James Hardin, then a developer, now at the University of South Carolina and an Associate Editor, was most active in supporting the NetCourse idea, despite a widespread view in the company that it would not work. Some NetCourses have led to books put out by Stata Press, StataCorp's publishing arm, the first being a text on maximum likelihood by Bill Gould and Bill Sribney and now extended in its second edition by new coauthor Jeff Pitblado.

The web site *www.stata.com* was launched in 1996 and has since become central to the Stata enterprise. It was somewhat later, in Stata 6, that Stata itself became web-aware, allowing reading of datasets (and other files) across the net and `update`s of Stata without the need for distributing diskettes or CDs.

The Internet also had long-term consequences for the STB. When it started, distribution of diskettes by mail was for most users the most practical way to acquire copies of new programs. But the rise of email and of web sites made that role increasingly redundant. As Statalist and the SSC and other archives exemplified, users could publish their programs on the Internet as soon as they liked and make fixes and enhancements as soon as they liked. Even with frequent publication rate and short queues, the STB could not possibly match that. Diskettes were phased out to the extent that what was published in the STB could be installed in Stata by users across the net (regardless of whether they were paid subscribers), but this did not affect the major issue. For that and other reasons, the STB was relaunched in 2001 as the *Stata Journal*. Joe Newton maintained continuity as Editor of both and was joined from the launch by Nicholas Cox. The story has been told in greater length elsewhere (Newton and Cox 2003).

2001 also saw the first 64-bit version, for Solaris. The second was for 64-bit Linux, and others will follow for Windows and OS X when the manufacturers release the operating systems. Another adjustment to bigger sizes was the release of Stata/SE, with support for larger datasets.

Stata 8 was the biggest release in the history of Stata, and one of the most painful internally, because of the very large mass of code and documentation put out all at once. This featured a completely new graphics system. Its principal architect was Vince Wiggins, who oversaw the development team. The new graphics was the biggest single project in the history of the company. Although for many users, it came some years later than they had hoped, the delay was not for want of much long and hard work. A new graphics had been developed some years before, but that project was pulled as not good enough. StataCorp struggled for some while to maintain backward compatibility with previous graphics, but reluctantly had to sacrifice that goal. Those who have worked with it appreciate the flexibility and versatility of the new design, and many users have adopted it, and even built upon it, with enthusiasm.

Another project of almost equal size was the introduction of dialogs in Stata 8. To be sure, Stata had previously had some dialog functionality, but by the standards of many windowed products it was rather limited. Naturally, many users, and many more potential users, are accustomed to dialogs in much of their software experience. One of several crucial considerations was the need to make dialogs compatible with the command language; any design that in effect bifurcated Stata into "command language Stata" and "dialog-based Stata" would have violated Stata's design principles. Key to the way that dialogs work is echoing of the equivalent command to the Command window, so that beginning users can, it is hoped, see how natural and consistent the language is once you get used to it. Not a few advanced users reported, at least privately, that playing with dialogs was one of the main ways they came to grips with the new and much more complicated graphics syntax.

The experience of releasing Stata 8 prompted a change in release policy towards dribbling out more code between major releases. The web-aware character of Stata and the ease with which `update`s are possible for users (except those behind very thick firewalls) encouraged a return to the idea of more frequent releases. Stata 8.1 was sparked by an updated `ml` and Stata 8.2 by some moderately substantial changes to graphics. Other substantial features have been added recently, such as the time-series additions (VECs and structural VECs) released in 2004 and marked by a new edition of the *Time-Series Reference Manual*.

## 4  Pervasive principles

What principles have guided Stata? Let me attempt to distill some of the most pervasive.

Despite modern user interfaces, the heart of Stata remains the command language. Whatever is done via menus or dialogs is, ideally, echoed as a command. The overwhelming emphasis on a command language follows from a firm belief that statistical analysis cannot be reduced to a small series of standard tasks. In particular, smarter statistical users, especially those near the cutting edge of research in many fields, do not want the statistical equivalent of a burger bar, with choice from a fixed menu, however appealing the individual dishes may seem to some tastes. They do want to go beyond

the menu and go inside the kitchen, not to peel the potatoes or fillet the fish, but to order something not on the menu and discuss it with the chef.

One key to good Stata practice is combining do-files and log files to produce reproducible research. A do-file documents a series of analytical steps. A log file documents a series of results. Any statistical project can go through many iterations, digressions, and dead ends, but one aim should remain the production of the do-file and the log file, documenting what you would have done at the outset had you been smart enough and well-informed enough to know what you now know. It hardly needs underlining that this combination depends on a command language.

The executable provides a kernel for the statistical operating system that is Stata. It gives the functionality that users would usually find too difficult or time-consuming to write themselves. The executable in 2005 is much bigger than the first executable compiled in 1984; no surprise there. But a more important measure is the proportion of Stata written in Stata, as ado-files or other human-readable program files. Asymptotically, it will approach 1!

Maintaining platform-independence is important for Stata. Why does StataCorp do this when almost all users use just one of the platforms supported (and most of those use one particular platform)? The answer lies in a mix of reasons that are technical, business, and aesthetic. It seems neither prudent nor tasteful to yoke Stata to any particular platform.

Maintaining backwards compatibility as Stata is extended is also vital. StataCorp's marketing style reflects a hope that users will stay loyal over the long run, evidenced by the high fraction who upgrade on each new release. To encourage and reward that, StataCorp tries to minimize changes that will break users' programs, do-files, or habits. This is not always easy, and mechanisms have been developed to allow this, most notably version control.

Statistical software differs from most other software in one key respect: the way that users need, and indeed contribute to, an extensible language. Such software, however good, still leaves many users with statistical needs they want to satisfy, but many of them do have the ability to write programs to meet those needs. Stata users are often aware of methods, even whole fields, not yet covered by Stata, as is inevitable given the size and complexity of statistics and the large number of sciences using statistics. Stata has gained considerably from what users themselves have done. For this to be true, it is not essential that all Stata users become Stata programmers. Just a few hundred who have put their work into the public domain have been sufficient to give other users extra leverage and to have impacted Stata itself. A user community contributing to Stata directly and indirectly drives a mighty feedback loop, propelling Stata's continued growth.

# 5   Acknowledgments

# 6   Appendix

Table 1: Timeline of Stata events

| | |
|---|---|
| March 1983 | IBM/XT announced |
| | Compaq luggable (without hard disk) announced |
| January 1984 | Development of Stata begins |
| | • DiAL (later Stata) by Bill Gould and Sean Becketti |
| August 1984 | IBM/AT announced |
| January 1985 | Stata 1.0 released |
| | • regression package with data management |
| | • 44 commands |
| February 1985 | Stata 1.1 |
| | • bug fixes |
| May 1985 | Stata 1.2 |
| | • new menu system |
| | • better online help |
| | • `keep` command |
| August 1985 | Stata 1.3 |
| | • real graphics on IBM CGA and Hercules Monochrome card! |
| | • Stata/Graphics, extra-cost item |
| January 1986 | revealed existence of `program` command in 1.3 |
| February 1986 | Stat.Kit and Graph.Kit |
| May 1986 | Stata 1.4 |
| | • new documentation |
| | • formatted `infile` |
| | • first to "officially" contain programming |
| | • still in a ring binder |
| February 1987 | Stata 1.5 (official, preofficial versions existed) |
| | • `anova` |
| | • `logit`, `probit` |
| | • Data.Kit |
| | • expanded memory support |

| | |
|---|---|
| August 1987 | BiTurbo Stata<br>• this was the day of "Turbo Pascal" |
| May 1988 | Stata for Sun/Unix<br>• first Stata for Unix version |
| June 1988 | Stata 2.0<br>• new graphics; graphics now part of Stata<br>• string variables<br>• survival analysis: Cox and Kaplan–Meier<br>• stepwise regression<br>• all new Kits<br>• still in a ring binder |
| January 1989 | Stata for 386/ix<br>• second Stata for Unix version |
| May 1989 | Stata 2.05<br>• same as 2.0, but documentation rebound in new format |
| December 1989 | Stage 1.0<br>• written in one month, never updated |
| December 1989 | *Statistics with Stata* by Lawrence Hamilton published by Duxbury Press<br>• first Stata book |
| March 1990 | Intercooled Stata<br>• for the 80386 (flat memory) |
| March 1990 | Stata for HP/9000<br>• third Stata for Unix version |
| April 1990 | CRC moves from West Los Angeles to Santa Monica, CA<br>CRC exits mainframe service bureau business |
| May 1990 | Reshape.Kit |
| May 1990 | Metrika Consulting (Peter Hedstrom, Sweden) first invoice<br>• first Stata distributor |
| September 1990 | Stata 2.1<br>• byte variables<br>• factor analysis<br>• ado-files<br>• `reshape` part of Stata in ado |
| January 1991 | bootstrap, etc., on support disk, written in ado<br>• Stata needs vehicle to distribute ado-files |
| May 1991 | STB issue 1, Joe Hilbe, Editor |

| November 1991 | Timberlake (Ana Timberlake, UK) first invoice |
| | • 2nd distributor |

| February 1992 | Stata for DEC RISC |
| | • fourth Stata for Unix version |
| | • first request from manufacturer |

| March 1992 | Stata 3.0 |
| | • `logistic`, `ologit`, `oprobit`, `clogit`, `mlogit`, ... |
| | • `tobit`, `cnreg`, `rreg`, `qreg`, `weibull`, `ereg`, ... |
| | • `epitab` |
| | • weights reconsidered: `pweights` |

| June 1992 | Stata for IBM RS/6000 |
| | • fifth Stata for Unix version |

| July 1992 | Stata for Macintosh |
| | • first Stata for Macintosh version |
| | author: Bill Rogers |

| May 1993 | STB begins 3rd year with Sean Becketti, editor |

| August 1993 | CRC moves to College Station, TX, becomes StataCorp |

| August 1993 | Stata 3.1 |
| | • `mvreg`, `sureg`, `heckman`, `nlreg`, `areg`, `canon` |
| | • `nbreg` |
| | • constrained linear regression |
| | • `ml` (first version) |
| | • `codebook` |

| January 1994 | Stata for DEC Alpha |
| | • sixth Stata for Unix version |

| March 1994 | StataQuest |

| August 1994 | Statalist begins |
| | • moderated by David Wormuth |

| January 1995 | Stata 4.0 |
| | • Stata for Windows 3.1 |
| | • `xtreg`, etc. |
| | • `glm` |

| February 1995 | Stata for Convex |
| | • seventh Stata for Unix version |

| March 1995 | First UK meeting |
| | • two-day meeting held at London School of Economics (LSE) |
| | organizer: A. Timberlake |

| April 1995 | Stata for Linux |
| | • eighth Stata for Unix version |

| June 1995 | NC-151 given for first time |
| | • Hardin supported NetCourse idea; most thought it would not work |
| | • course by Gould and Hardin |
| August 1995 | Stata for Windows 95 and Windows/NT |
| September 1995 | Stata 4.0+ for Mac |
| | • second Macintosh version |
| September 1995 | NC-101 given for first time |
| | • second NetCourse |
| January 1996 | First session of NC-152 |
| | • third NetCourse |
| | authors: Gould and Sribney |
| March 1996 | *www.stata.com* web site announced |
| May 1996 | STB editorship change, good-bye to Sean Becketti and welcome to Joe Newton |
| October 1996 | Stata 5.0 |
| | • `xtgee`, `xtprobit`, . . . |
| | • `prais`, `newey`, `intreg`, . . . |
| | • survey estimation commands |
| | • `fracpoly` |
| | • `st` extended |
| February 1997 | NC-200 given for first time |
| | • course on ML estimation, by Gould and Sribney |
| | • Gould and Sribney rewrite `ml`, then write course |
| | • Gould and Sribney writing book |
| November 1997 | First ado-file stored in SSC archive |
| | • Seed, Paul. `strip.ado`. |
| | "Stata module to strip unwanted characters" |
| | • note that SSC archive *predates* Stata's web capabilities |
| April 1998 | Stage for Linux |
| January 1999 | Stata 6 |
| | • web aware (`use http://`, updating, etc.) |
| | • new `ml` |
| | • time-series operators, etc. |
| | • `arima`, `arch`, . . . |
| | • `st` rewritten |
| January 1999 | *Maximum Likelihood Estimation with Stata* book by Gould and Sribney |
| | • first Stata Press book |
| February 2000 | Stata on CD-ROM |

| | |
|---|---|
| April 2000 | first session of NC-631 |
| | • survival analysis |
| | • authors: Cleves and Gould |
| July 2000 | STB-56 becomes first edition available electronically |
| December 2000 | Stata 7 |
| | • frailty |
| | • `xtabond`, etc. |
| | • cluster analysis |
| | • `nlogit` |
| | • `roc` |
| | • SMCL |
| | • windowed interface for Unix |
| | • ado-files byable |
| December 2000 | Moves to StataCorp campus in College Station |
| May 2001 | *Generalized Linear Models and Extensions* book by Hardin and Hilbe |
| | • second Stata Press book |
| May 2001 | STB-61, last issue of STB |
| August 2001 | Stata for Macintosh OS X |
| September 2001 | First round GLLAMM speedup code added |
| | • included in free update |
| November 2001 | Stata for SGI Irix |
| | • ninth Stata for Unix version |
| November 2001 | Stata for 64-bit Solaris (first 64 bit) |
| | • tenth Stata for Unix version |
| | • first 64-bit port |
| November 2001 | Stata Journal, Volume 1, Number 1, published |
| | • Joe Newton and Nicholas Cox, editors |
| February 2002 | Stata/SE released |
| | • support for large datasets |
| May 2002 | *An Introduction to Survival Analysis using Stata* book by Cleves and Gould |
| | • third Stata Press book |
| January 2003 | Stata 8 |
| | • graphics |
| | • extended GUI, dialog boxes available for all commands |
| | • `manova`, . . . |
| | • more survey |
| | • more time series (VARs, SVARs) |
| | • more GLLAMM internalization |

| April 2003 | *www.stata.com* adds searchable Statalist archive |
|---|---|
| July 2003 | Stata 8.1 |
| | • free upgrade |
| | • updated `ml` |
| | • *Maximum Likelihood with Estimation*, 2nd edition |
| | by Gould, Pitblado, Sribney |
| October 2003 | Stata 8.2 |
| | • free upgrade |
| | • number of changes required for graphics |
| September 2003 | Measurement error software available |
| | • authors: Carroll, Hardin, Schmiediche |
| December 2003 | Stata for 64-bit Linux |
| | • eleventh Stata for Unix version |
| | • second 64-bit port |
| May 2004 | *A Visual Guide to Stata Graphics* book by Michael Mitchell |
| | • Stata Press book |
| July 2004 | Time-series additions to Stata: |
| | • free upgrade |
| | • VECs and structural VECs |
| | • New time-series manual |

Table 2: Releases of Stata

| 1.0  | January 1985   | 3.1 | August 1993   |
|------|----------------|-----|---------------|
| 1.1  | February 1985  | 4.0 | January 1995  |
| 1.2  | March 1985     | 5.0 | October 1996  |
| 1.4  | August 1986    | 6.0 | January 1999  |
| 1.5  | February 1987  | 7.0 | December 2000 |
| 2.0  | June 1988      | 8.0 | January 2003  |
| 2.05 | June 1989      | 8.1 | July 2003     |
| 2.1  | September 1990 | 8.2 | October 2003  |
| 3.0  | March 1992     |     |               |

Table 3: Stata 1.0 and Stata 1.1

```
append      dir         infile      plot        spool
beep        do          input       query       summarize
by          drop        label       regress     tabulate
capture     erase       list        rename      test
confirm     exit        macro       replace     type
convert     expand      merge       run         use
correlate   format      modify      save
count       generate    more        set
describe    help        outfile     sort
```

Stata 1.0 had `int`, `long`, `float`, and `double` variables; it did not have `byte` or `str#`.

Table 4: Stata Users Group meetings

|  |  | organizers of program | StataCorp attendees |
| --- | --- | --- | --- |
| London | 21–22 Mar 1995 | A. Timberlake | W. Gould, W. Sribney |
| London | 5 Jul 1996 | P. Royston |  |
| London | 5–6 Jun 1997 | N. J. Cox | W. Gould, J. Hardin |
| London | 15 May 1998 | M. Hills | V. Wiggins |
| London | 17–18 May 1999 | S. P. Jenkins, P. Sasieni | W. Gould, A. Riley |
| Córdoba | 20–21 May 1999 | A. Tobias | W. Gould, M. Cleves |
| Madrid | 12 May 2000 | A. Tobias | M. Cleves |
| London | 15 May 2000 | N. J. Cox, P. Royston | W. Gould |
| Utrecht | 17 May 2000 | J. Weesie, A. Volovics | W. Gould |
| Boston | 12–13 Mar 2001 | C. F. Baum, N. J. Cox, M. Pagano | W. Gould, D. Drukker |
| London | 14–15 May 2001 | B. L. De Stavola, S. P. Jenkins | W. Gould, R. Gutierrez |
| London | 20–21 May 2002 | N. J. Cox, P. Royston | W. Gould, R. Gutierrez |
| Maastricht | 23 May 2002 | W. van Putten, A. Volovics | W. Gould, R. Gutierrez |
| Boston | 18–19 Mar 2003 | E. Allred, C. F. Baum, N. J. Cox, M. Pagano | W. Gould, C. Nguyen, L. Gilmore |
| London | 19–20 May 2003 | S. Rabe-Hesketh, S. P. Jenkins | W. Gould, R. Gutierrez, C. Nguyen |
| Dublin | 22 May 2003 | R. Conroy, A. Kelly | W. Gould, R. Gutierrez |
| Berlin | 12 Aug 2003 | U. Kohler | W. Gould, V. Wiggins |
| Berlin | 5 Apr 2004 | J. Giesecke, U. Kohler | A. Riley, R. Gutierrez |
| London | 28–29 Jun 2004 | N. J. Cox, P. Royston | W. Gould, V. Wiggins |
| Boston | 23–24 Aug 2004 | E. Allred, C. F. Baum, N. J. Cox, R. Goldstein, P. A. Lachenbruch, M. Pagano | W. Gould, V. Wiggins, C. Farrar, G. Farrar |
| Adelaide | 10 Oct 2004 | J. Carlin, A. Esterman, P. Hakendorf, K. Keesman, C. Rickards, M. Rosier, P. Ryan, S. Stillman | A. Riley, R. Gutierrez |
| Rome | 25 Oct 2004 | U-L. Bell, R. Bellocco, G. Capelli, M. Pagano | V. Wiggins |

Table 5: Editors and Associate Editors, *Stata Technical Bulletin* and *Stata Journal*
(Associate Editors unless otherwise stated)

| | |
|---|---|
| Joseph Hilbe | STB Editor 1991–1993, 1993–1994 |
| J. Theodore Anagnoson | STB 1991–1994 |
| Richard DeLeon | STB 1991–1994 |
| Paul Geiger | STB 1991–1994 |
| Richard Goldstein | STB 1991 |
| Stewart West | STB 1991–1994 |
| Lawrence C. Hamilton | STB 1991–1994 |
| Sean Becketti | STB Editor 1993–1996 |
| Francis X. Diebold | STB 1994–2000 |
| Joanne M. Garrett | STB 1994–2001, SJ 2001– |
| Marcello Pagano | STB 1994–2001, SJ 2001– |
| James L. Powell | STB 1994–1998 |
| J. Patrick Royston | STB 1994–2001, SJ 2001– |
| H. Joseph Newton | STB Editor 1996–2001, SJ Editor 2001– |
| Nicholas J. Cox | STB 1998–2001, SJ Executive Editor 2001–2004, SJ Editor 2005– |
| Jeroen Weesie | STB 2000–2001, SJ 2001– |
| Christopher F. Baum | STB 2001, SJ 2001– |
| Rino Bellocco | SJ 2001– |
| David Clayton | SJ 2001– |
| Charles Franklin | SJ 2001– |
| James Hardin | SJ 2001– |
| Stephen Jenkins | SJ 2001– |
| Jens Lauritsen | SJ 2001– |
| Stanley Lemeshow | SJ 2001– |
| J. Scott Long | SJ 2001– |
| Thomas Lumley | SJ 2001– |
| Sophia Rabe-Hesketh | SJ 2001– |
| Philip Ryan | SJ 2001– |
| Allan Gregory | SJ 2002– |
| Jeffrey Wooldridge | SJ 2002– |
| Mario A. Cleves | SJ 2003– |
| Roger Newson | SJ 2003– |
| Mark E. Schaffer | SJ 2003– |
| William D. Dupont | SJ 2004– |
| Ulrich Kohler | SJ 2004– |
| Nicholas J.G. Winter | SJ 2004– |

# 7   References

Newton, H. J. and N. J. Cox. 2003. The Stata Journal so far: Editor's report. *Stata Journal* 3(2): 105–108.

**About the Author**

Nicholas Cox is a statistically minded geographer at the University of Durham. He contributes talks, postings, FAQs, and programs to the Stata user community. He has also co-authored fifteen commands in official Stata. He was an author of several inserts in the *Stata Technical Bulletin* and is Executive Editor of the *Stata Journal*.