ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Hlávka, Zdeněk

Working Paper

Electronic books for experts and users

SFB 373 Discussion Paper, No. 2003,22

Provided in Cooperation with:
Collaborative Research Center 373: Quantification and Simulation of Economic Processes, Humboldt University Berlin

Suggested Citation: Hlávka, Zdeněk (2003) : Electronic books for experts and users, SFB 373 Discussion Paper, No. 2003,22, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin, https://nbn-resolving.de/urn:nbn:de:kobv:11-10050213

This Version is available at:
https://hdl.handle.net/10419/22237

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Electronic books for experts and users

Zdeněk Hlávka *
C.A.S.E. & Institut für Statistik
Humboldt-Universität zu Berlin
Spandauer Str. 1, 10178 Berlin, Germany

email: hlavka@wiwi.hu-berlin.de

May 2003

Abstract

The paper describes our experience with the production of electronic textbooks. Electronic books are defined as supplements to usual printed textbooks. The first part of the paper describes the technical background of our electronic book system. In the second part we describe how this system can be used also by authors who do not want to spend too much time with setting up the system and learning all of the (un)necessary technical details. The last section discusses various forms of distribution of electronic books.

Contents

1 The technology of MD*Book **4**

2 New format: MD*Booklet **7**

3 Distribution and types of media **9**

What is e-book?

The purpose of an electronic book is to enhance the functionality of the ordinary printed book. The areas in which some improvement over the paper is possible are portability, interactivity and easy linking with other enhancing material. These aspects can be addressed in a different way depending on the **format** of the e-book and on the way in which it is **distributed**.

Using the standard (free) LaTeX tools, the following formats can be produced easily:

postscript for printed version of the book (LaTeX + `dvips`),

pdf for downloadable e-book (PDF LaTeX),

html for online e-book (`latex2html`).

Any computer system aiming at the production of electronic books should address also the following features:

*The support of the Deutsche Forschungsgemeninschaft, Sonderforschungsbereich 373 "Quantification and Simulation of Economic Processes" at Humboldt-Universität zu Berlin is gratefully acknowledged.

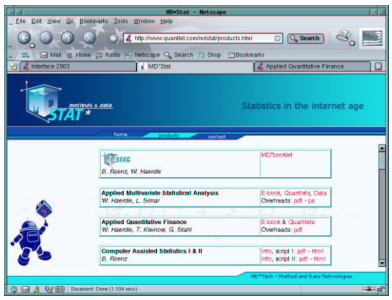
↪ automatic production of various output formats,

↪ interactive features,

↪ ability to create related documents with identical structure such as slides or exercises,

↪ extensions to other output formats.

The following text contains many screenshots of various electronic textbooks.
All examples of the electronic books presented in this paper can be downloaded at www.md-stat.com.
The necessary tools can be obtained at www.md-book.com.

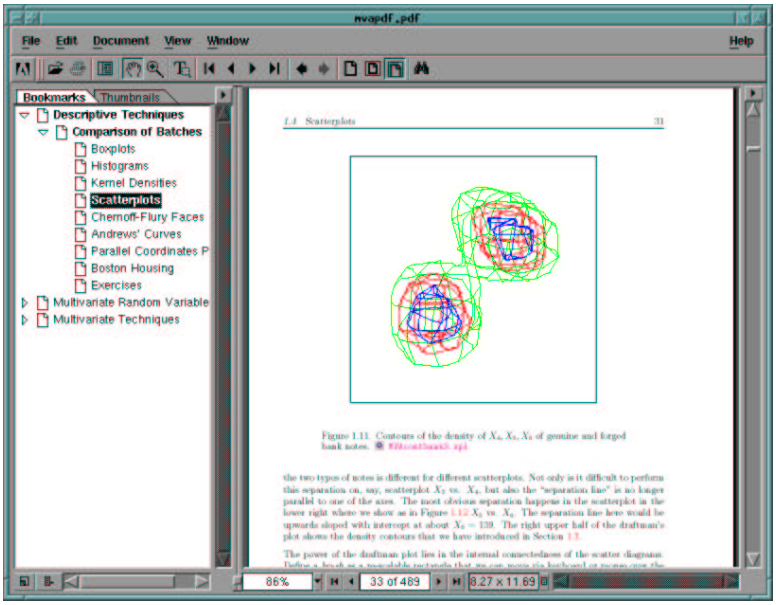


In the next subsections we give an overview of common e-book formats and their properties.

PDF format

The main advantage of the PDF format is that its layout and typesetting are identical to the printed book. In addition, it can contain links which make the navigation in the document easier. Links to other material on the internet are also possible. PDF file can be easily downloaded as its name, *portable document format*, suggests.

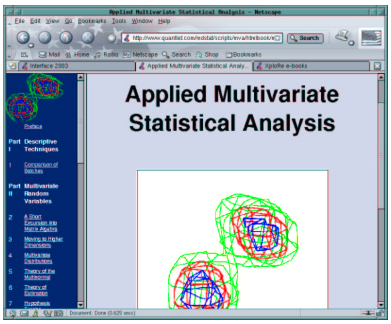
The main disadvantage of this format is that it inherits "paper related properties" of the printed book, e.g., the page size. Using a computer, one would certainly prefer variable page size which would be derived from the structure of the document. On the other hand, the page size allows to print the document or its part while keeping the nice layout. Unfortunatelly, printing is not the primary motivation here since it turns the e-book back into ordinary printed book.

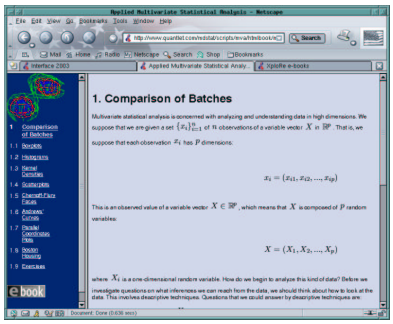


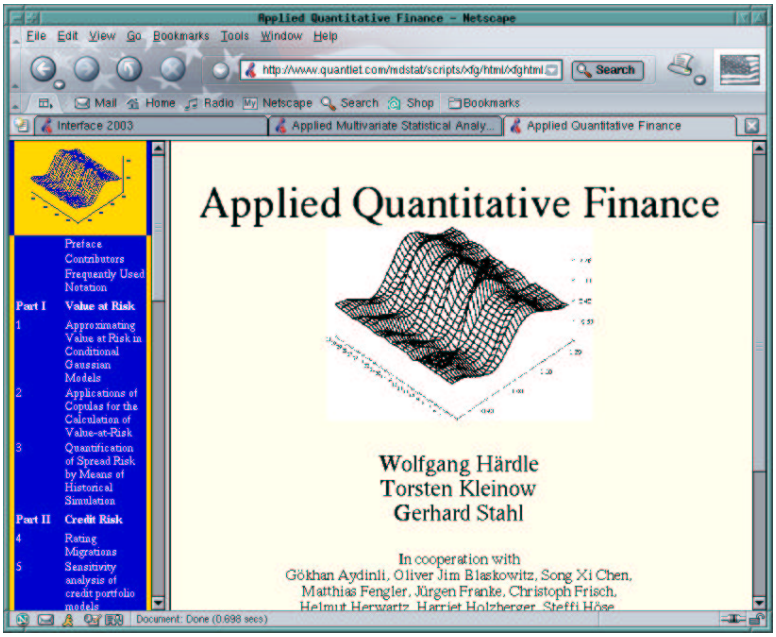
HTML format

In the current version of `latex2html`, the formulas are included as GIFs and thus the text in this format does not look as good as in PDF. Due to the nature of the HTML format, the producer does not control the layout of the page.

The advantages of HTML format are that the e-book structure follows the structure of the text, the look of the page can be controlled with CSS files and the possibility to include interactive components (e.g., Java).







Comparison of PDF and HTML

From our experience, the PDF version is suitable for download. It can be easily copied and it does not take much space. It is very well suited especially for displaying teaching material such as overheads. The formulas are displayed exactly as in the printed version (in the best LaTeX quality).

The HTML version of the document is very suitable as the online www version of the book. It stays on the web server and it can be used as a fast reference from anywhere in the world even though it would be quite time consuming to download all the necessary files to your computer. Due to the good structuring of the document, it is easier to find the formula you are

interested in. It is not necessary to download the complete book if you want to have a look only at one section.

It is hard to decide which e-book format is the best one. The final choice of the output format depends on other criterions. One should define the purpose of the document first.

The best way is probably to provide both formats and let the user (student, customer) choose which one suits his needs better. This means that there is a need for a system which handles more output formats (at least postscript, PDF and HTML) automatically.

1 The technology of MD*Book

MD*Book is an e-book preparation system based on LaTeX that manages many output formats automatically. In its simplest form, it has only Linux command-line interface which offers many possibilities of influencing the behavior of the system.

MD*Book expects as its input a **.sk** file which defines the structure of the document and LaTeX files which contain the text of the book. It is typically envoked by the following command which generates postscript and html version of the document from `inputfile.sk`:

```
mdbook2 -ps -html inputfile.sk
```

The file `inputfile.sk` can look, for example, like this:

```
USEPACKAGE      html
BEGIN ps pdf
USEPACKAGE      hyperref
END
FINALPS         psbook/mvatex.ps
FINALPDF        pdfbook/mvapdf.pdf
FINALHTML       htmlbook/mvahtml.html
INPUT ise
FORMAT          scrbook 12pt a4paper headsepline
START           title
PART
TITLE           Descriptive Techniques
CHAPTER         c1
TITLE           Comparison of Batches
LABEL           ch01
SECTION         c1boxin
TITLE           Boxplots
LABEL           c1boxin
SECTION         c1histo
TITLE           Histograms
INDEX           histograms
BIBLIOGRAPHY    mm
PRINTINDEX
END
```

According to the instructions given in the above file, **MD*Book** creates ordinary LaTeX file which can partially vary depending on the desired output format.

In the above example, the command `\usepackage{hyperref}` is used only if postscript or PDF version of the e-book is produced. Similar construction is available for the `INPUT` keyword. Thus `INPUT ise` is included as `\input{ise@html}` if the html version is produced and if the file `ise@html.tex`

exists. For more detailed description we refer to MD*Book homepage at www.md-book.com. After the LaTeX document is created, MD*Book runs

↪ LaTeX and `dvips` to obtain the postscript version,

↪ PDF LaTeXto obtain the PDF version,

↪ `latex2html` to obtain the HTML version.

The resulting output files are copied into the directories specified in the `FINALPS`, `FINALPDF` and `FINALHTML` keywords.

The use of MD*Book is necessary since many commands have to be defined differently for postscript, PDF and HTML version. MD*Book allows to generate all the output formats from the same LaTeX source file. This makes the file maintenance much easier.

Notice that we provide also `tex2sk` tool which allows automatic conversion of large LaTeX files into the MD*Book format. Here, the MD*Book keywords can be included using `\SKparam{}{}` command.

Keyword MODE

One of the appealing features of MD*Book is the `MODE` keyword. It allows to administer more documents which have identical structure but contain different text.

This feature is typically used if you want to generate both lecture script and overheads whose structure follows the time schedule of the course.

In the following example, the keyword `MODE slide` switches to the slide mode. The script and the overheads have common .sk file which contains the following instructions based on the combination of the `MODE` keyword and the `BEGIN END` condition.

```
BEGIN       slidemode
MODE        slide
END
```

This means that we switch to the slide mode if MD*Book is called in the following way:

```
mdbook2 -ps -pdf inputfile.sk slidemode
```

The slide mode implies that the following part of inputfile.sk

```
SECTION          c1boxin
TITLE            Boxplots
```

is then understood as

```
\section{Boxplots}
\input{c1boxinslide}
```

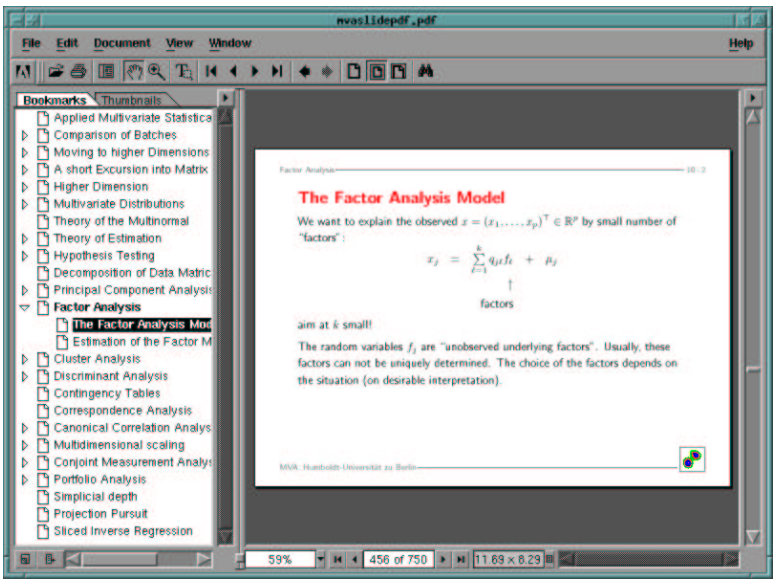
instead of

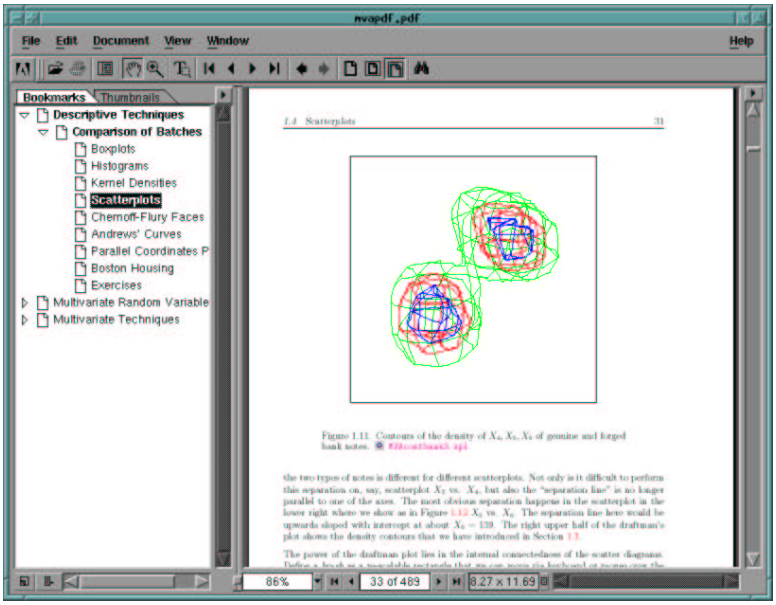
```
\section{Boxplots}
\input{c1boxin}
```

which would result from `mdbook2 -ps -pdf inputfile.sk`.

On the next page, you can compare two screenshots. The first one is taken from the e-book "Applied Multivariate Statistical Analysis", the other one from the corresponding overheads[1]. The navigation menus do not correspond to each other, but the structure of both documents (the order of sections) is the same.

[1] the overheads can be accessed using login "multi" and password "variate"

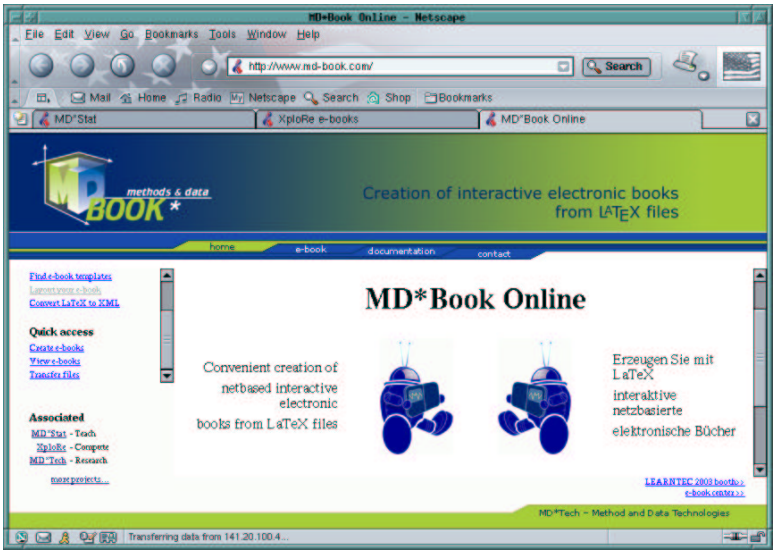




MD*Book online

The biggest problem of MD*Book is clearly its dependency on the specific version of the PDF LaTeX and `latex2html` commands. These packages keep changing and, at this moment, MD*Book works reliably only on SuSE Linux 7.1. On other systems, it is very fragile. The biggest problem is the `latex2html` which is still under development and which still changes a lot.

We can not guarantee that MD*Book works well with any version of LaTeX, PDF LaTeX and `latex2html`. Therefore, we offer the MD*Book online service (Klinke and Witzel 2002) at www.md-book.com which allows anybody to compile his LaTeX files on our server.



If you want to obtain nice electronic book, you should follow these steps:

1. create login,
2. create e-book project,
3. transfer your LaTeX files to our server,
4. create desired output format by clicking on the appropriate button,
5. download the resulting files.

In praxis, the LaTeX files usually need some tuning and running the system for the first time might take considerable amount of time. For larger projects, we strongly recommend the command line version of MD*Book.

2 New format: MD*Booklet

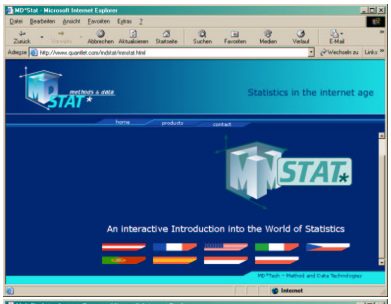
We have claimed that new e-book formats are desirable and should be possible. MD*Booklet is such a new format based on HTML (and `latex2html`) which uses a lot of Jscript in order to provide better and more user friendly navigation. There exist many .sk keywords which allow to personalize its graphical user interface.

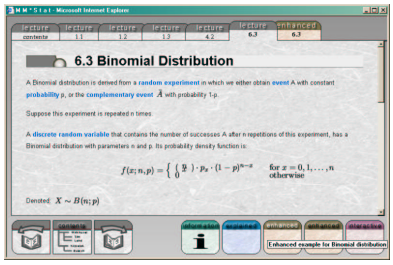
The MD*Booklet output format can be generated using the command

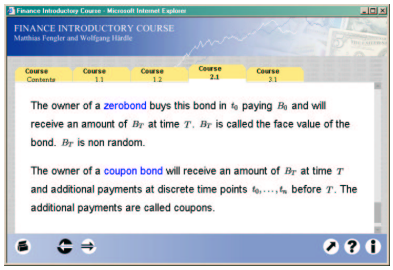
```
mdbook2 -java inputfile.sk.
```

This format is best suited for lecture notes. Its interactive features allow students to discover the explained topic by themselves. The most prominent (the first one) project using this output format is the German MM*Stat project (Härdle and Rönz 2001). Some other language versions can be accessed at www.quantlet.com/mdstat/mmstat.html.

The following screenshots show the starting page of the MM*Stat project and screenshots from MM*Stat, XploRe Introductory Course and Finance Introductory Course.







Computing components

The resulting documents can include also computing/interactive components. In MD*Book, this is currently solved by links from the document to the computing/interactive examples which are included in separate HTML pages.

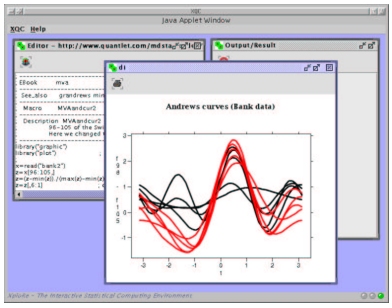
Such pages can be also automatically generated by MD*Book using the command

```
mdbook2 -xpl inputfile.sk
```

where `inputfile.sk` has to define the keyword `DIRXPL`, the directory containing all XploRe files. The automatically generated HTML pages allow to execure or edit the programs on the XploRe Quantlet Server (XQS). In principle, any other statistical package can be used, but only XploRe environment (Härdle, Klinke and Müller 2000) is embedded directly in MD*Book.

The client program (XploRe Quantlet Client) is implemented in Java and it runs under any common operating system (Linux, Windows, Mac OS). More information on the XploRe Client–Server technology can be found at www.xplore-stat.de.

After clicking on the edit button, you can modify the computer code which is downloaded from the server and execute it by clicking on the execute button in the upper left corner of the edit window. In this example, the graphical display containig Andrews' curves of Swiss bank notes date appears.



3 Distribution and types of media

After producing nicely using electronic book, one wishes to distribute it to as many people as possible. During last five years, we used these ways of distribution:

On-line HTML version,

Free download PDF version,

CD-ROM inserted in the printed version of the book,

Download version available using information in the printed book.

We have already discussed the advantages and disadvantages of the on-line (preferably HTML) and free download (preferably PDF) versions in the Introduction. It remains to discuss the CD-ROM and Download version of the electronic book.

CD-ROM

The electronic book on the CD contains basically the same components as the online versions. Some care has to be taken so that the links from the book don't go to internet but to the CD directories. This is done automatically by MD*Book. From the author's point of view, the keyword `TARGET cdrom_final` has to be included in `inputfile.sk`. In praxis, `inputfile.sk` contains the construction

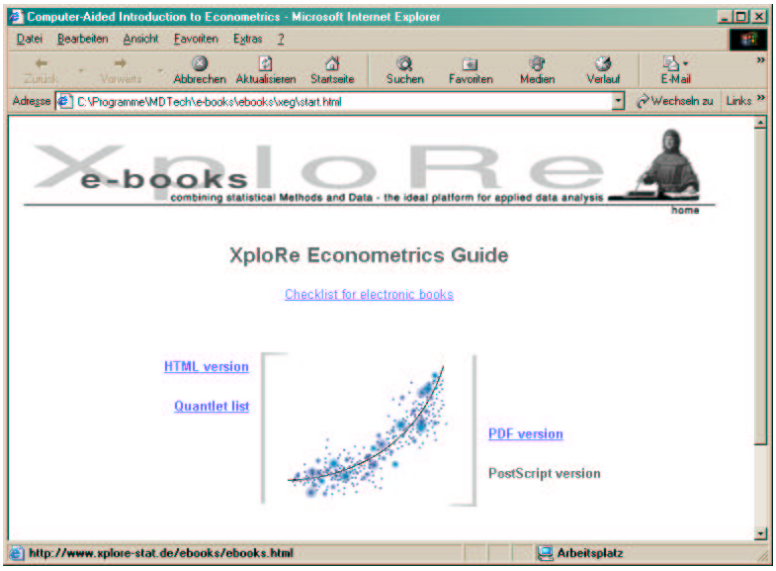
```
BEGIN cdrom
TARGET cdrom_final
END  cdrom
```

The CD-ROM is then generated by

```
mdbook2 -cdrom inputfile.sk
mdbook2 -ps -pdf -html -xpl inputfile.sk cdrom
mdbook2 -start inputfile.sk
```

The first command generates the required directory structure, the second command produces the e-book in postscript, PDF and HTML and the third command generates the CD-ROM starting page.

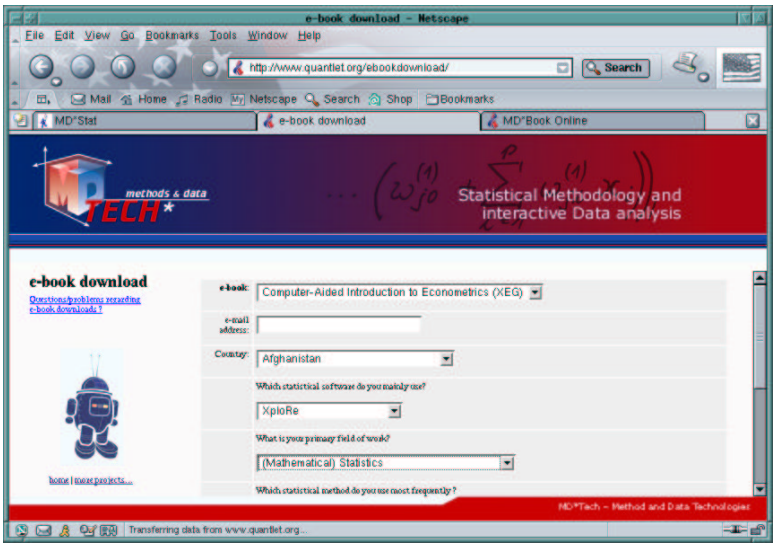
The next screenshot shows the starting page of the Computer Assisted Introduction into Econometrics (Rodriguez Poo 2003). You can see that the CD-ROM e-book contains both the HTML and the PDF versions of the text and a list of all included computational examples.



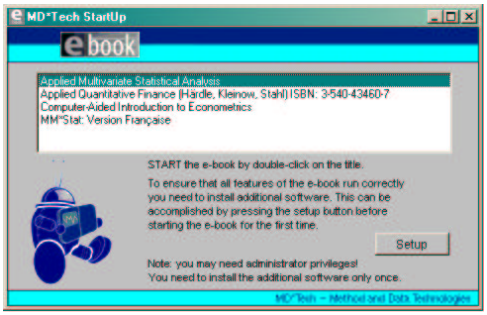
At first look, the CD-ROM version seemed to be very attractive. Unfortunatelly, we have found out that inserting the CD-ROM into the printed book cannot be automatized and thus it is very expensive. Also, the CD-ROM version cannot be updated. For these reasons, the download version is preferable.

Download

From technical point of view, the download version is identical to the CD-ROM version. The only difference is that it is installed on the harddisk. It is distributed as self-extracting executable file and it can be downloaded from our page using the serial number inserted in the printed book.



After downloading and installing the files, a shorcut icon appears on the desktop and in the Start menu. When starting the e-book by doubleclicking on the icon, the following menu containing all installed e-books appears.



After clicking on the book you are interested in, the local XploRe Quantlet Server is started so that the included computational examples can be run locally. The starting HTML page already looks exactly as in the CD-ROM version.

Conclusion

Springer-Verlag already offers some books containing CD-ROM (Härdle, Hlávka and Klinke 2000, Härdle, Klinke and Müller 1999) or downloadable e-books (Härdle, Kleinow and Stahl 2002, Rodriguez Poo 2003). In future, Springer will offer this technology also to more authors.

It is not clear which of the current formats (such as Open eBook or Adobe eBook) will become the standard in the next decade, but MD*Book is certainly a useful tool that can adapt and survive.

More information on MD*Book and many electronic books can be found at the following www pages:

↪ www.md-stat.com,

↪ www.md-book.com.

References

Härdle, W., Hlávka, Z. & Klinke, S. (2000). *XploRe - Application Guide*, Springer Verlag, Heidelberg.

Härdle, W., Kleinow, T. & Stahl, G. (2002). *Applied Quantitative Finance*, Springer Verlag, Heidelberg.

Härdle W., Klinke S. & Müller M. (2000). *XploRe - The Interactive Statistical Computing Environment, Academic Edition*, Springer Verlag, Heidelberg.

Härdle W., Klinke S. & Müller M. (1999). *XploRe Learning Guide*, Springer Verlag, Heidelberg.

Härdle W. & Rönz, B. (2001). *MM*Stat, an Interactive Introduction into the World of Statistics (in German)*, Springer Verlag, Heidelberg.

Klinke S. & Witzel R. (2002). *MD*Book online - a tool for creating interactive documents, DP 0221*, SFB 373, Humboldt-University of Berlin.

Rodriguez Poo, J. (Ed.) (2003). *Computer-Aided Introduction to Econometrics*, Springer Verlag, Heidelberg.