

Applied Econometrics with R: Package Vignette and Errata

Christian Kleiber
Universität Basel

Achim Zeileis
Wirtschaftsuniversität Wien

Abstract

“Applied Econometrics with R” (Kleiber and Zeileis 2008, Springer-Verlag, ISBN 978-0-387-77316-2, pp. vii+222) is the first book on applied econometrics using the R system for statistical computing and graphics (R Development Core Team 2008). It presents hands-on examples for a wide range of econometric models, from classical linear regression models for cross-section, time series or panel data and the common non-linear models of microeconometrics, such as logit, probit, tobit models as well as regression models for count data, to recent semiparametric extensions. In addition, it provides a chapter on programming, including simulations, optimization and an introduction to R tools enabling reproducible econometric research. The methods are presented by illustrating, among other things, the fitting of wage equations, growth regressions, hedonic regressions, dynamic regressions and time series models as well as models of labor force participation or the demand for health care.

The book is accompanied by the R package **AER** (Zeileis and Kleiber 2008) which contains some new R functionality, some 100 data sets taken from a wide variety of sources, the full source code for all examples used in the book, as well as further worked examples, e.g., from popular textbooks.

This vignette provides an overview of the package contents and contains a list of errata for the book.

Keywords: econometrics, statistical software, R.

1. Package overview

1.1. R code from the book

The full R code from the book is provided in the demos for **AER** the package. The source scripts can be found in the `demo` directory of the package and executed interactively by calling `demo()`, e.g.,

```
R> demo("Ch-Intro", package = "AER")
```

One demo per chapter is provided:

- Ch-Intro (Chapter 1: Introduction),
- Ch-Basics (Chapter 2: Basics),

- `Ch-LinearRegression` (Chapter 3: Linear Regression),
- `Ch-Validation` (Chapter 4: Diagnostics and Alternative Methods of Regression),
- `Ch-Microeconometrics` (Chapter 5: Models of Microeconometrics),
- `Ch-TimeSeries` (Chapter 6: Time Series),
- `Ch-Programming` (Chapter 7: Programming Your Own Analysis).

This list of demos is also shown by `demo(package = "AER")`.

The same scripts are contained in the `tests` directory of the package so that they are automatically checked and compared with the desired output provided in `.Rout.save` files.

To make the code fully reproducible and to avoid some lengthy computations in the daily checks, a few selected code chunks are commented in the scripts. Also, for technical reasons, some graphics code chunks are repeated, once commented and once without comments.

1.2. Data sets

The **AER** package includes some 100 data sets from leading applied econometrics journals and popular econometrics textbooks. Many data sets have been obtained from the data archives of the *Journal of Applied Econometrics* and the *Journal of Business & Economic Statistics*. Some of these have been used in recent textbooks, among them Baltagi (2002), Davidson and MacKinnon (2004), Greene (2003), Stock and Watson (2007), and Verbeek (2004). In addition, we provide all further data sets from Baltagi (2002), Franses (1998), Greene (2003), and Stock and Watson (2007).

Detailed information about the source of each data set, descriptions of the variables included, and usually also examples for typical analyses are provided on the respective manual pages. A full list of all data sets in **AER** can be obtained via:

```
R> data(package = "AER")
```

Additionally, overview manual pages for some textbooks are provided. They list all data sets from the respective book and provide extensive code for replicating many of the empirical examples. See, e.g.

```
R> help("Greene2003", package = "AER")
```

for data sets and code for Greene (2003). The overview manual pages currently available in **AER** are:

- `Baltagi2002` for Baltagi (2002),
- `CameronTrivedi1998` for Cameron and Trivedi (1998),
- `Franses1998` for Franses (1998),
- `Greene2003` for Greene (2003),
- `StockWatson2007` for Stock and Watson (2007).

1.3. New R functions

AER provides a few new R functions extending or complementing methods previously available in R:

- `tobit()` is a convenience interface to `survreg()` from package **survival** for fitting tobit regressions to censored data. In addition to the fitting function itself, the usual set of accessor and extractor functions is provided, e.g., `print()`, `summary()`, `logLik()`, etc. For more details see `?tobit`.
- `ivreg()` fits instrumental-variable regressions via two-stage least squares. It provides a formula interface and calls the workhorse function `ivreg.fit()` which in turn calls `lm.fit()` twice. In addition to the fitting functions, the usual set of accessor and extractor functions is provided, e.g., `print()`, `summary()`, `anova()`, etc. For more details see `?ivreg`, `?ivreg.fit`, and `?summary.ivreg`, respectively.
- `dispersiontest()` tests the null hypothesis of equidispersion in Poisson regressions against the alternative of overdispersion and/or underdispersion. For more details see `?dispersiontest`.

2. Errata

No errata, yet. Please report any errors you find to us.

References

- Baltagi BH (2002). *Econometrics*. 3rd edition. Springer-Verlag, New York. URL <http://www.springeronline.com/sgw/cda/frontpage/0,10735,4-165-2-107420-0,00.html>.
- Cameron AC, Trivedi PK (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge.
- Davidson R, MacKinnon JG (2004). *Econometric Theory and Methods*. Oxford University Press, Oxford.
- Franses PH (1998). *Time Series Models for Business and Economic Forecasting*. Cambridge University Press, Cambridge. URL <http://www.few.eur.nl/few/people/franses/research/book2.htm>.
- Greene WH (2003). *Econometric Analysis*. 5th edition. Prentice Hall, Upper Saddle River, NJ. URL <http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm>.
- Kleiber C, Zeileis A (2008). *Applied Econometrics with R*. Springer-Verlag, New York. ISBN 978-0-387-77316-2.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Stock JH, Watson MW (2007). *Introduction to Econometrics*. 2nd edition. Addison-Wesley, Reading, MA.

Verbeek M (2004). *A Guide to Modern Econometrics*. 2nd edition. John Wiley & Sons, Hoboken, NJ.

Zeileis A, Kleiber C (2008). *AER: Applied Econometrics with R*. R package version 1.0-0, URL <http://CRAN.R-project.org/package=AER>.

Affiliation:

Christian Kleiber

Wirtschaftswissenschaftliches Zentrum

Universität Basel

Petersgraben 51

CH-4051 Basel, Switzerland

E-mail: Christian.Kleiber@unibas.ch

URL: <http://www.wvz.unibas.ch/stat/team/kleiber/kleiber.htm>

Achim Zeileis

Department of Statistics and Mathematics

Wirtschaftsuniversität Wien

Augasse 2-6

AT-1090 Wien, Austria

E-mail: Achim.Zeileis@R-project.org

URL: <http://statmath.wu-wien.ac.at/~zeileis/>