

Notes for the **R** code accompanying the book
*Statistical Studies of Income Poverty
and Inequality in Europe*

Nicholas T. Longford*

Version May 2014

This document describes the organisation of the R-code files. The code can be downloaded from www.sntl.co.uk/EUsilc.html as a zip-file in which nine zip-files and this document are packed. The nine zip-files, named **Chapt x .zip**, $x = 1, 2, \dots, 9$, contain the code for the respective chapters x .

The files associated with chapter x have names

EU x AB.an N ,

where $x = 1, 2, \dots, 9$, AB indicates the type of the code, and N is the order number. AB is one of the following:

- **ch** — check (on the data)
- **e1** — elementary (basic and infrastructural) matters and items
- **ev** — evaluation (often application of a function)
- **ew** — evaluation (often following up on **ev**)
- **ex** — evaluation (only one file, **EU7ex.an4**)
- **fn** — function
- **fp** — function for additional computing (only one file, **EU5fp.anD**)

*Email: sntlnick@sntl.co.uk. SNTL and Universitat Pompeu Fabra, c/Ramon Trias Fargas 25–27, 08005 Barcelona, Spain.

- **gl** — illustration (graphics)
- **gr** — graphics for the output of **ev**, **ew**
- **gs** — graphics for the output of **ev**, **ew**
- **gt** — graphics for the output of **ev**, **ew**
- **gu** — graphics for the output of **ev**, **ew**
- **gv** — graphics for the output of **ev**, **ew** (only one file, **EU1gv.an1**)
- **io** — input-output

Exceptions are in Chapter 9, where there are code files **EU9evCO.an1** and **EU9evCO.an1**; *CO* is the two-letter acronym of a country (e.g. AT, BE, CZ, DK). Files **EU9grCO.an2** and **EU9grCO.an2** are used similarly for graphics related to country *CO*.

Altogether there are 265 files; 60 of them are for principal functions (with names **EU?fn.an?**), 96 for evaluations (**EU?ev*.an?** and **EU?ew*.an?**) and 88 are for graphics (**EU?g*.an?**). For Chapters 1–9 there are $26 + 8 + 38 + 28 + 46 + 37 + 26 + 25 + 31$ files.

Files **EUxAB.anN** for fixed numerals *x* (Chapter) and *N* form a unit with related code for **AB** equal to **fn**, **fp**, **ev**, **ew**, **gr**, **gs**, **gt** and **gu**. In most units, not all the files are present. The unit is referred to as **EUx-N**. For example, unit **EU3-2** contains files **EU3fn.an2**, **EU3ev.an2**, **EU3ew.an2**, **EU3gr.an2**, **EU3gs.an2** and **EU3gt.an2**. Files **EUxel.anN**, **EUxgl.anN** **EUxio.anN** do not belong to the unit **EUx-N**, except for **EU4gl.an7** and **EU6gl.an6**. Some chapters contain more than nine units; the counter for *N* continues after 9 with A, B, C, etc. In a few instances, there are gaps; for example, there is no unit **EU5-9**, but there is unit **EU5-A**.

EUxfn.anN contains the principal function of the unit. **EU5fp.anD** contains a function for some additional computing. **EUxev.anN** contains execution code, using the function(s) declared in **EUxfn.anN**, in the same unit. Further executions may be in file **EUxew.anN**. Graphics based on the output of **EUxev.anN** and **EUxew.anN** are in files **EUxgr.anN**, **EUxgs.anN**, etc.

The function-files `EU?fn.an?` contain sometimes an example execution. (Note that the symbols `?` and `*` are used in the same way as in the operating system, to indicate wildcards, of unit length or any length, respectively.) The function should be introduced into the workspace before using the execution-files in the same unit. I recommend the expression

```
source("~/Splus/EUsilc12/EU3fn.an2")
```

for the example above. Of course, you have to change the directory path. Next, ‘source’ the execution files `EU3ev.an2`, `EU3ew.an2`, followed by the graphics files `EU3g?.an2`. The graphics in `EUxgr.anN` is not always associated (solely) with `EUxev.anN`; sometimes it is based (also) on `EUxew.anN`. Chapter 9 is an exception, see below.

Chapters can be skipped, except for Chapters 1 and 2. In any case, they contain simpler code that is useful for ‘warm-up’, and to get accustomed to some of the authors’ conventions. The functions declared in `EU1el.an1` and the vectors and list in `EU1el.an2` are used throughout the chapters.

The files `EUxio.anN` should be introduced at the beginning of Chapter x . There are four such files, in chapters $x = 1, 5, 7$ and 9. The files `EUxch.anN` are for data checking and orientation. They can be skipped if you are confident that you have the right data and all the procedures that generate the datasets are in good order.

The files `EUxel.anN` contain mostly simple (auxiliary) functions and other code of ‘infrastructural’ nature and `EUxgl.anN` contain code (function and execution) for illustrations that are reproduced in the book. They can be skipped by experienced R users, but they are of value to beginners in R or at the first reading of the book.

The files `EU1el.an1` and `EU1el.an2` contain functions used throughout the chapters. The functions declared in `EUxio.anN` in Chapter x may be useful (and some are used) in later chapters. There is some near-duplication in the functions across the chapters, but that is to make the units within a chapter (nearly) self-contained. Certainly, there is compatibility from earlier chapters to later chapters.

If a function (or some other user-defined object) has not been introduced in

the workspace, and is required in an expression that you want to execute, you obtain the error message

```
Error: could not find function "xyz"
```

If it is a function (object) generated by some earlier code, find the file (in MAC/OS or Linux) by the command

```
grep xyz EU???.an?
```

in the appropriate directory. You can sharpen the search for function `xyz` by the command

```
grep "xyz " EU?f?.an?
```

because it will search only function-files (with mid-part `fn` or `fp`), and will report only instances in which `xyz` is followed by a space, as in

```
xyz <- function(...)
```

In Chapter 9, there is a principal function (`POTF`) declared in `EU9fn.an1`, secondary functions called by `POTF` are declared in `EU9fn.an2`, and further simple (auxiliary) functions are declared in `EU9fn.an3`. There are many of these functions, and it is not practical to organise them like functions in the other chapters.

Each file has a primary description in its first line, e.g.,

```
## Principal function 3, Chapter 3
```

in `EU3fn.an3`, followed by a line giving the filename

```
## Filename EU3fn.an3
```

and the next line the topic

```
## Relative poverty gaps
```

In files which contain several functions, these functions are listed, e.g.,

```
## Functions LogitW, KHsel1, KHbin, KHmtc, KHbal, KHdiat
```

in `EU9fn.an2`. In each function, the first line describes its purpose, e.g.,

```
### Logistic regression with sampling weights
```

in function `LogitW`. It is followed by listing the arguments and their description, reproduced in many instances in the chapter appendices; e.g.,

```
## X          The regression matrix (vector if only one variable)
## y          The outcome variable (vector of zeros and ones)
## w          The sampling weights
...

```

All loops that span more than just a few lines are indicated by comments at the beginning and end of the loop, e.g.,

```
} ## End of the iteration loop (while)
```

Each file with expressions for execution that takes more than a few seconds has a comment at the bottom stating how long the execution takes; e.g.,

```
## Takes 78 sec.
```

in `U5ev.an6`. The times, in seconds of CPU time, are on a vintage MacBook Pro; newer equipment is much faster. But the quoted time gives an indication of what to expect.

The file `merge.do`, included in the zip-file, contains an example of extracting and reshaping the panel data in `Stata`. The extracted datasets are used in Chapters 6 and 7.

Please report all bugs, errors, glitches, poor descriptions, ambiguities, and ways of improving the code and documentation to me at `sntl@nick@sntl.co.uk`. I may be embarrassed and annoyed at first (not with you, but about my lack of perfection), but I will appreciate your interest, the pursuit of correctness, perfection and elegance, and will be keen to fix any problem.

The following two pages contain an index of all the 204 user-defined functions and their location in the files.

R is for eveR

Nicholas T. Longford
Barcelona, May 2014.

Index of user-defined R functions

Function	File	Function	File	Function	File
CRange	EU7el.an1	EUmixF	EU4el.an1	EUpres	EU6gr.an2
CVcor	EU7el.an1	EUmixG	EU4gr.an6	EUpvAR	EU8fn.an3
Dround	EU1el.an1	EUmixG2	EU4gr.an2	EUpvARG	EU8gr.an3
DRW	EU6gl.an6	EUmixG3	EU4gs.an3	EUpvSFG	EU8gs.an4
EUboot	EU2gl.an2	EUmixG4	EU4gr.an4	EUqua2	EU5fn.an3
EUBoots	EU2fn.an1	EUmixH2	EU4gt.an2	EUqua3	EU5fn.an8
EUbts	EU2fn.an1	EUmixT	EU4gr.an3	EUquaF	EU3gr.an9
EUcheck	EU1ch.an2	EUMmx0	EU7el.an1	EUquaG	EU5gr.an3
EUcnfG	EU4gr.an7	EUMmx1	EU7fn.an1	EUquaH	EU5gr.an8
EUcnfH	EU4gl.an7	EUMmx2	EU7fn.an4	EUread	EU1io.an1
EUcnfM	EU4fn.an8	EUMmxD	EU7gs.an4	EUreadC	EU6ch.an4
EUconf	EU4fn.an7	EUMmxF	EU7fn.an2	EUreadMx	EU7io.an1
EUconM	EU7fn.an7	EUMmxG	EU7gr.an1	EUreadP	EU6ch.an1
EUcouD	EU5el.an1	EUMmxH	EU7gr.an4	EUreadP2	EU6ch.an3
EUdns	EU4el.an1	EUMmxJ	EU7gt.an4	EUreadQ	EU6fn.an1
EUeff	EU1el.an1	EUMmxU	EU7gr.an6	EUreadR	EU5io.an1
EUfnr	EU3ev.an2	EUMtb	EU7fn.an3	EUregA	EU5fn.an5
EUhist	EU1gl.an2	EUneg	EU3ch.an1	EUregB	EU5fn.an6
EUloc	EU1el.an1	EUnewton	EU3fn.an2	EUregC	EU5fn.anC
EUlogn	EU3fn.an8	EUnorm	EU1gl.an1	EUregD	EU5fn.anD
EULrnz	EU3gl.an1	EUoddsR	EU1el.an1	EUregE	EU5fn.anB
EULrzA	EU3fn.an5	EUpair	EU7gs.an3	EUregF	EU5gr.an1
EULrzB	EU3fn.an6	EUpat4Q	EU6fn.an2	EUregG	EU5gr.anD
EULrzD	EU3fn.an7	EUperv	EU3fn.anA	EUregH	EU5gr.anC
EULrzE	EU5fn.an4	EUpia	EU1el.an1	EUregV	EU5fp.anD
EULrzF	EU3gr.an5	EUpoef	EU8fn.an5	EURoPal	EU6gl.an1
EULrzG	EU3gt.an5	EUpoefG	EU8gr.an5	EUsca	EU1el.an1
EULrzH	EU3gr.an6	EUpov1	EU1fn.an1	EUscaR	EU1el.an1
EULrzJ	EU3gr.an7	EUpov1D	EU5fn.an2	EUscaM	EU1el.an1
EULrzK	EU5gr.an4	EUpov2	EU1gr.an1	EUshapQ	EU6ew.an1
EULrzL	EU3gr.an8	EUpov2C	EU1gt.an1	EUSQua	EU3fn.an9
EULrzW	EU3fn.an5	EUpov2Y	EU1gs.an1	EUSReF	EU8gr.an6
EULvar	EU3fn.an8	EUpovA	EU3fn.an1	EUSReF2	EU8gs.an6
EUmed	EU2gl.an1	EUpovB	EU3fn.an3	EUSReF3	EU8gt.an6
EUMgr	EU7gr.an3	EUpovC	EU3fn.an4	EUSReg	EU8fn.an6
EUmix0	EU4el.an2	EUpovF	EU3gr.an1	EUssz	EU5fn.anA
EUmix1	EU4fn.an1	EUpovFD	EU5gr.an2	EUsszF	EU5gr.anA
EUmix2	EU4fn.an2	EUpovG	EU3gr.an3	EUSTA	EU7fn.an6
EUmixA	EU4gl.an1	EUpovL	EU3el.an1	EUSTpf	EU8fn.an5
EUmixB	EU4fn.an5	EUpovS	EU3el.an1	EUSTra0	EU8fn.an1
EUmixB2	EU4fn.an6	EUpovX	EU1fn.an1	EUSTra0G	EU8gr.an1

Function	File	Function	File	Function	File
EUSTra1	EU8fn.an2	EUYear	EU1el.an1	KHtab	EU9fn.an4
EUSTra1G	EU8gr.an2	expl	EU1el.an1	LeUni	EU1el.an1
EUSTra2G	EU8gs.an2	ExtrC	EU1el.an1	Lgt	EU1el.an1
EUSTra4	EU8fn.an4	ExtrE	EU1el.an1	lgtP	EU1el.an1
EUSvar	EU1fn.an4	ExtrL	EU1el.an1	LogitW	EU9fn.an2
EUtab	EU4fn.an3	ExtrR	EU1el.an1	LTable	EU1io.an1
EUtraMI	EU6fn.an9	Figure	EU1io.an1	MaxS	EU9fn.an2
EUtransA	EU6fn.an3	HDeckG	EU9gl.an1	MeVa	EU1el.an1
EUtransD	EU6fn.an4	HTable	EU1io.an1	MeVaR	EU9fn.an3
EUtransE	EU6gr.an4	InflX	EU9gl.an2	NiNh	EU3ch.an1
EUtransF	EU6fn.an5	InitMx	EU7el.an1	Perc	EU1el.an1
EUtransG	EU6gr.an3	Kern	EU8el.an1	Pos10	EU6fn.an6
EUtransH	EU6gs.an3	KernSq	EU1el.an1	POTF	EU9fn.an1
EUtransK	EU6fn.an6	KHbal	EU9fn.an2	POTm	EU9fn.an5
EUtransL	EU6gr.an5	KHbin	EU9fn.an2	Prob	EU1el.an1
EUtransM	EU6fn.an7	KHblf	EU9gs.an1	PwTra	EU6fn.an6
EUtransN	EU6fn.an8	KHblg	EU9gr.an1	QuantB	EU2fn.an2
EUtransP	EU6gl.an6	KHcheck	EU9fn.an3	QuantC	EU1fn.an2
EUtransT	EU6gt.an3	KHcnt	EU9gsAT.an2	QuantG	EU1gr.an2
EUtransW	EU6fn.anB	KHdiat	EU9fn.an2	QuantH	EU2gr.an2
EUtraR	EU6ew.anA	KHist	EU9grAT.an2	QuantW	EU1fn.an2
EUtrn	EU7gr.an2	KHmtc	EU9fn.an2	REGR	EU9fn.an5
EUwei	EU1fn.an3	KHpat	EU9fn.an4	RegrW	EU9fn.an5
EUweiF	EU1gr.an3	KHread	EU9io.an1	SMOO	EU1el.an1
EUweiG	EU1gs.an3	KHsell	EU9fn.an2	SNdns	EU7el.an1
EUweiH	EU1gt.an3	KHsmr	EU9fn.an3	TRL	EU6gl.an6
EUweiJ	EU1gu.an3	KHsmt	EU9fn.an3	VarIn	EU9fn.an3