

ACE Models with the NLSY

William Howard Beasley (Howard Live Oak LLC, Norman)
Joseph Lee Rodgers (University of Oklahoma, Norman)
Kelly Meredith (University of Oklahoma, Norman)
David Bard (University of Oklahoma Health Sciences Center, OKC)

February 23, 2012

Abstract

We describe how to use the `NlsyLinks` package to examine various biometric models, using the NLSY79.

[dss](#)

Not only am I hoping to capitalize on all researchers initially interested in NLSY BG studies, I'm also hoping for a small slice of attracting those that aren't initially interested in BG. If the vignette or articles are good, they should rank well among search engines for terms like "NLSY and R". Maybe some researchers using those terms (and not initially interested in BG) could stumble on the vignettes. If the examples are clear and are easy to run quickly, they may get interested in BG. Therefore, I'd like at least a few paragraphs explaining what BG is. This section is tentatively located in the appendix.

1 Terminology

This package considers both Gen1 and Gen2 subjects. **Gen1** refers to subjects in the original NLSY79 sample (<http://www.bls.gov/nls/nlsy79.htm>). **Gen2** subjects are the biological children of the Gen1 females -ie, those in the NLSY79 Children and Young Adults sample (<http://www.bls.gov/nls/nlsy79ch.htm>).

The **expected coefficient of relatedness** of a pair of subjects is typically represented by the variable `R`. Examples are: Monozygotic twins have $R=1$; dizygotic twins have $R=0.5$; full siblings (ie, those who share both biological parents) have $R=0.5$; half-siblings (ie, those who share exactly one biological parent) have $R=0.25$; adopted siblings have $R=0.0$. Other possibilities exist too. The font (and hopefully their context) should distinguish the variable `R` from the software `R`.

An **outcome variable** is directly relevant to the applied researcher; these might represent constructs like height, IQ, and income. A **plumbing variable** is necessary to manage BG datasets; examples are `R`, `SubjectID`.

2 Example: DF analysis with a simple DV for Gen2 subjects

The first example uses a simple statistical model and all available Gen2 subjects. The `CreatePairLinks` function will create a data frame where each represents one pair of siblings. This function examines the subjects' IDs and determines who is related to whom (and by how much). By default, each row it produces has at least six values/columns: (i) ID for the older subject `Subject1Tag`, (ii) ID for the younger subject `Subject2Tag`, (iii) ID for their extended family `ExtendedID`, (iv) their coefficient of expected relatedness `R`, (v *and beyond*) outcome values for the older subject; (vi *and beyond*) outcome values for the younger subject. The steps are:

1. Use the NLS investigator (<http://www.nlsinfo.org/investigator/>) to select and download Gen2 dataset. See this package's vignette dedicated to the NLS Investigator by typing `vignette("NlsInvestigator")` or by visiting <http://cran.r-project.org/web/packages/NlsyLinks/>.

2. Open R and create a new script (see Appendix A) and load the `NlsyLinks` package. If you haven't done so, first install the `NlsyLinks` package (see Appendix B).
3. Within the R script, identify the locations of the downloaded data file, and load it into a data frame.
4. Specify the outcome variable name and filter out all subjects who have a negative value in this variable. The NLSY typically uses negative values to indicate different types of missingness.
5. Call the `CreatePairLinks` function. Pass the raw data frame and the name of the outcome variable.
6. Use R's `lm` routine (i.e., general linear model) to estimate the coefficients of the DF model.
7. Execute the R script (see Appendix A). The interpretations of the DF analysis can be found in Rodgers, Buster and Rowe (2001).

```
### R Code for Example 1: ###  
# DF analysis with a simple DV and Gen1 subjects  
#Step 2: #Load the package containing the linking routines.  
require(NlsyLinks)
```

3 Appendix A: Creating and Saving R Scripts

bla bla

4 Appendix B: Installing the NlsyLinks Package

bla bla bla