



NLSY79 data: drinking habits of youth

Contents

1. Introduction.....	1
2. Latent variables for alcohol use.....	3
3. Drinking and smoking habits of children.....	5

1. Introduction

The data for the current study were obtained from the National Longitudinal Survey of Youth (NLSY) of Labor Market Experience in Youth, a study initiated in 1979 by the United States Department of Labor in order to study the transition of young people into the labor force. The NLSY is a multistage, clustered probability sample of households.

In 1986, a separate survey of all children born to NLSY79 female respondents began, greatly expanding the breadth of child-specific information collected. In addition to all the mother's additional demographic and development information collected from either the mother or child. For children aged 10 or older, information has been collected from the children biennially since 1988 on a variety of factors including attitudes towards schooling, dating and friendship patterns, health, substance use, and home responsibilities.

In this example we use a subset of these data (`nlsy79habits.lsf`) and look specifically at questions relating to smoking. The first few lines of the data are shown below. This data set contains information on smoking, drinking and the use of marijuana.

LISREL for Windows - [nlsy79habits.lsf]

File Edit Data Transformation Statistics Graphs Multilevel SurveyGLIM View Window Help

	SMKEY90	SMK3M90	DNKEY90	DNK3M90	MRJEV90	MRJ3M90	SMKEY92	SMK3M92
1	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	0.00	0.00
2	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
3	0.00	0.00	0.00	0.00	0.00	0.00	-9.00	-9.00
4	0.00	0.00	1.00	2.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00
6	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
7	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
8	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
9	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	0.00	0.00
10	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
11	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
12	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
13	0.00	0.00	1.00	0.00	0.00	0.00	1.00	2.00
14	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00
	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00	-9.00

SCRL

A short description of the variables is as follows:

- MOMID: Mother's identification code
- MOMWT79: 1979 sampling weight for mother
- CHILDID: Child identification code
- CHAGE: Child's age in years at the first measurement occasion (1988)
- CRACE: Race of child (1 = white, 2 = black, 3 = non-black, non-hispanic)
- CSEX: Gender of child (1 = male, 2 = female)
- REGION: region of residence of the mother of the child (1 = North East, 2 = North Central, 3 = South, 4 = West)

The substance abuse variables SMKEV88, SMK3M88, ..., MRJPRS96 listed above are distinguished by three letters:

- INCO88 – INCO96: Family income for the years 1988, 1990, ..., 1996
- SMK = cigarette use
- DNK = alcohol use
- MRJ = marijuana use

For the substance use variables, the first three letters are followed by the subscripts

- EV : Ever used (0 = no, 1 = yes)
- 3M: used during last three months (0 = never, 1 = no, 2 = yes)
- PRS: pressured to use tobacco or other substances (0 = no, 1 = yes)
- 30D: used during last 30 days (0 = no, 1 = yes)
- REC: how recently used. Scale from 0 to 7 indicating never to very recently
- RACE1, RACE2 and RACE3 are dummy coded variables created from CRACE.

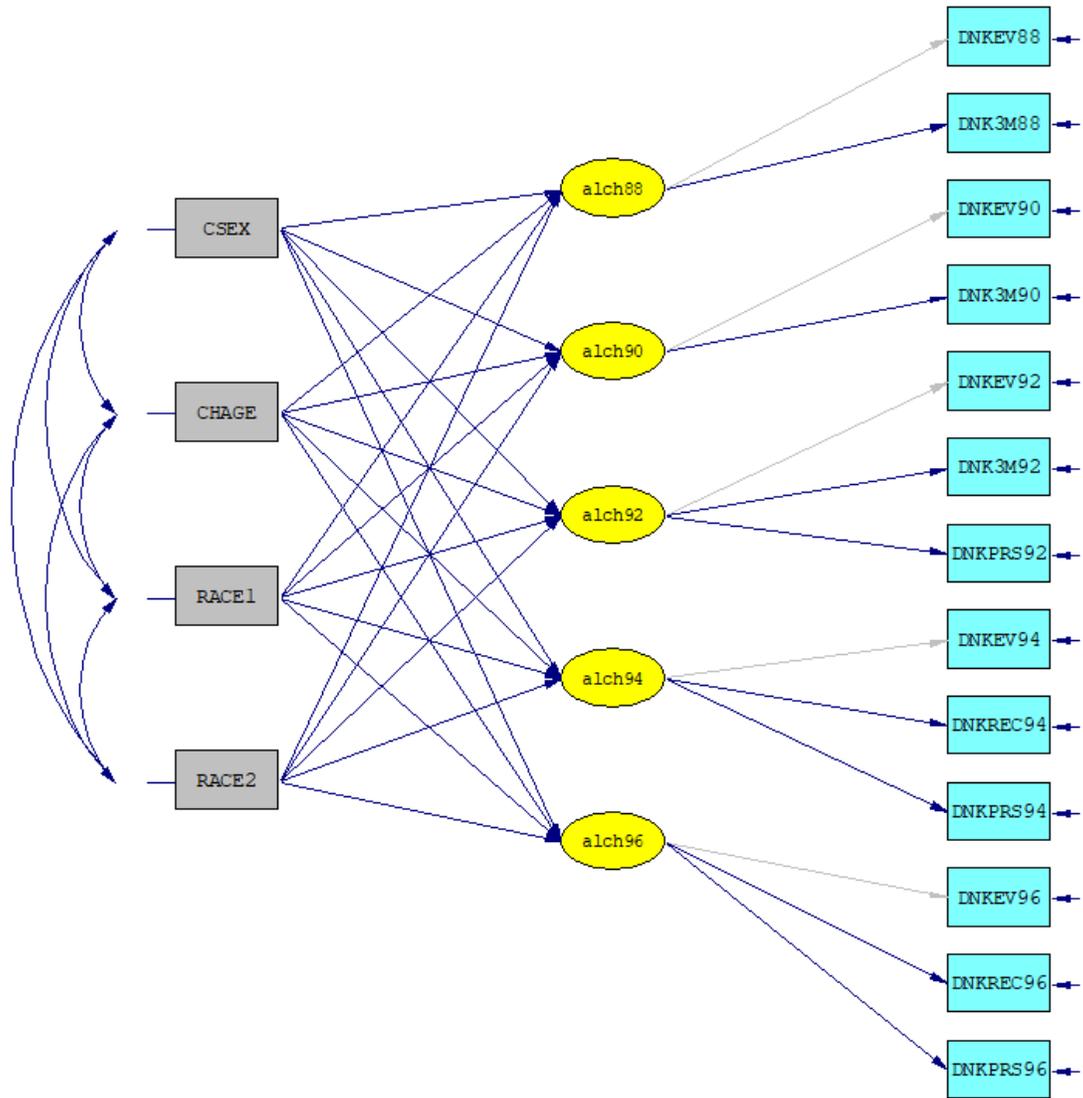
Of interest here is the smoking habits of youth over time. In 1988, 1990 and 1992 two questions on the use of tobacco were asked. In the last two completed questionnaires, 1994 and 1996, five questions were answered. These variables are assumed to measure the same underlying concept, namely cigarette use. For each of the 5 occasions, we assume that the use of the tobacco related questions are the dependent variables. In LISREL context, these are Y-variables. We assume that the five latent variables (cigr88, cigr90, cigr92, cigr94, and cigr96) are indicators of cigarette use on the five measurement occasions. When latent variables are associated with dependent variables, the latent variables are ETA variables in the LISREL framework.

We also have demographic information on gender, age and race that can be used as predictors of the Y-variables. These predictor variables are X-variables in the LISREL framework, and we include them in the model as potential predictors of the latent variables. While it is not necessary for the user to memorize the eight matrices traditionally used in LISREL, the above definitions are important as they play a role in correctly constructing a model by building a path diagram.

In Section 2 a latent variable model for alcohol use is discussed, and in Section 3 we consider a latent variable model for drinking and smoking habits. The results from the analysis in the latter section are used to fit a vector time series model in Section 4, using the current version of LISREL.

2. Latent variables for alcohol use

In the conceptual path diagram shown below, the variables available for measuring alcohol use are used as indicators of the exogenous latent variables alch88, alch90, ..., alch96. It is further assumed that the gender, race and age of a child are predictors of the propensity to use alcohol.



SIMPLIS syntax to fit the model to the data is given next. The default option in LISREL is to assume that the covariance matrix of the measurement errors of the indicators of the endogenous latent variables is a diagonal matrix. To estimate covariances, the Set the Error Covariances of var i and var j Free commands are added.

```

LISREL for Windows - [alcohol.SPI]
File Edit Options Window Help
Drinking habits of adult youth NLSY79 study
Raw Data from file 'NLSY79HABITS.LSF'
Latent Variables  alch88 alch90 alch92 alch94 alch96
Relationships
DNKEV88=1.00*alch88
DNK3M88=alch88
DNKEV90=1.00*alch90
DNK3M90=alch90
DNKEV92=1.00*alch92
DNK3M92=alch92
DNKPRS92=alch92
DNKEV94=1.00*alch94
DNKREC94=alch94
DNKPRS94=alch94
DNKEV96=1.00*alch96
DNKREC96=alch96
DNKPRS96=alch96
alch88=CSEX CHAGE RACE1 RACE2
alch90=CSEX CHAGE RACE1 RACE2
alch92=CSEX CHAGE RACE1 RACE2
alch94=CSEX CHAGE RACE1 RACE2
alch96=CSEX CHAGE RACE1 RACE2
Set the Error Covariance of alch90 and alch88 Free
Set the Error Covariance of alch92 and alch88 Free
Set the Error Covariance of alch92 and alch90 Free
Set the Error Covariance of alch94 and alch88 Free
Set the Error Covariance of alch94 and alch90 Free
Set the Error Covariance of alch94 and alch92 Free
Set the Error Covariance of alch96 and alch88 Free
Set the Error Covariance of alch96 and alch90 Free
Set the Error Covariance of alch96 and alch92 Free
Set the Error Covariance of alch96 and alch94 Free
Path Diagram
End of Problem
Ready

```

3. Drinking and smoking habits of children

In the SIMPLIS syntax (**habits.spl**) shown below, we extend the model for alcohol use to include information on smoking habits. The first section of the syntax file indicates that there are ten latent variables. Each pair of latent variables `alch(yr)` and `cigr(yr)` corresponds to a specific year: 1988, 1990, 1992, 1994, and 1996. Indicators of the latent variables for 1988 to 1992 are listed in the Relationships: section of the syntax file.

```

TI Drinking and Smoking habits of adult youth
Raw Data from file 'nlsy79habits.LSF'
Latent Variables  alch88 cigr88 alch90 cigr90 alch92 cigr92
                  alch94 cigr94 alch96 cigr96
Relationships:
DNKEV88 = 1.00*alch88
DNK3M88 = alch88
SMKEV88 = 1.00*cigr88
SMK3M88 = cigr88

```

DNKEV90 = 1.00*alch90
DNK3M90 = alch90
SMKEV90 = 1.00*cigr90
SMK3M90 = cigr90

DNKEV92 = 1.00*alch92
DNK3M92 DNKPRS92 = alch92
SMKEV92 = 1.00*cigr92
SMK3M92 = cigr92

DNKEV94 = 1.00*alch94
DNKREC94 DNKPRS94 = alch94
SMKEV94 = 1.00*cigr94
SMKPRS94 SMKREC94 SMK30D94 = cigr94

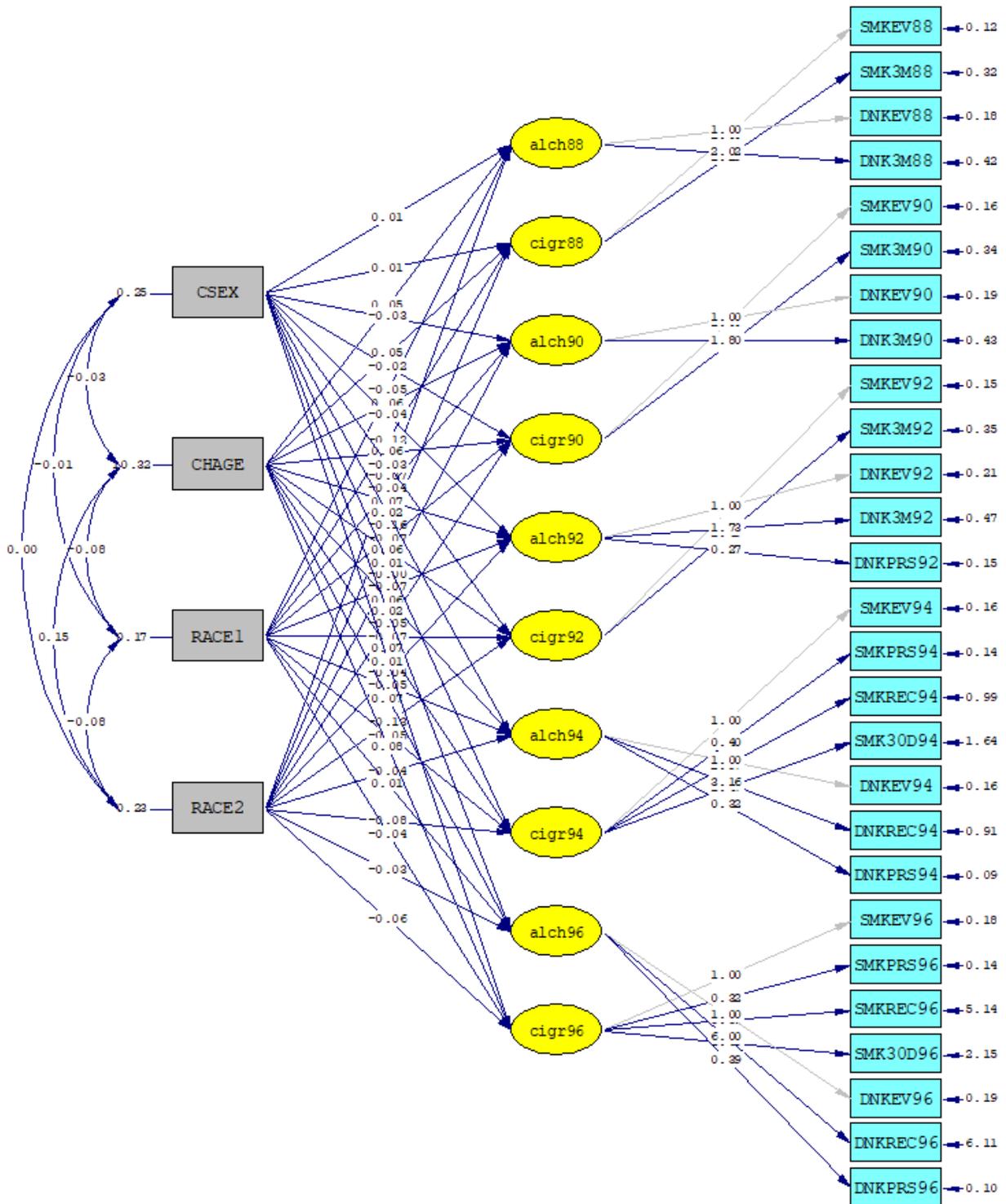
DNKEV96 = 1.00*alch96
DNKREC96 DNKPRS96 = alch96
SMKEV96 = 1.00*cigr96
SMKREC96 SMKPRS96 SMK30D96 = cigr96
alch88 alch90 alch92 alch94 alch96 = CSEX CHAGE RACE1 RACE2
cigr88 cigr90 cigr92 cigr94 cigr96 = CSEX CHAGE RACE1 RACE2
Set the Error Covariance of cigr88 and alch88 Free
Set the Error Covariance of alch90 and alch88 Free
Set the Error Covariance of alch90 and cigr88 Free
Set the Error Covariance of cigr90 and alch88 Free
Set the Error Covariance of cigr90 and cigr88 Free
Set the Error Covariance of cigr90 and alch90 Free
Set the Error Covariance of alch92 and alch88 Free
Set the Error Covariance of alch92 and cigr88 Free
Set the Error Covariance of alch92 and alch90 Free
Set the Error Covariance of alch92 and cigr90 Free
Set the Error Covariance of cigr92 and alch88 Free
Set the Error Covariance of cigr92 and cigr88 Free
Set the Error Covariance of cigr92 and alch90 Free
Set the Error Covariance of cigr92 and cigr90 Free
Set the Error Covariance of cigr92 and alch92 Free
Set the Error Covariance of alch94 and alch88 Free
Set the Error Covariance of alch94 and cigr88 Free
Set the Error Covariance of alch94 and alch90 Free
Set the Error Covariance of alch94 and cigr90 Free
Set the Error Covariance of alch94 and alch92 Free
Set the Error Covariance of alch94 and cigr92 Free
Set the Error Covariance of cigr94 and alch88 Free
Set the Error Covariance of cigr94 and cigr88 Free
Set the Error Covariance of cigr94 and alch90 Free
Set the Error Covariance of cigr94 and cigr90 Free
Set the Error Covariance of cigr94 and alch92 Free
Set the Error Covariance of cigr94 and cigr92 Free
Set the Error Covariance of cigr94 and alch94 Free
Set the Error Covariance of alch96 and alch88 Free
Set the Error Covariance of alch96 and cigr88 Free

```
Set the Error Covariance of alch96 and alch90 Free
Set the Error Covariance of alch96 and cigr90 Free
Set the Error Covariance of alch96 and alch92 Free
Set the Error Covariance of alch96 and cigr92 Free
Set the Error Covariance of alch96 and alch94 Free
Set the Error Covariance of alch96 and cigr94 Free
Set the Error Covariance of cigr96 and alch88 Free
Set the Error Covariance of cigr96 and cigr88 Free
Set the Error Covariance of cigr96 and alch90 Free
Set the Error Covariance of cigr96 and cigr90 Free
Set the Error Covariance of cigr96 and alch92 Free
Set the Error Covariance of cigr96 and cigr92 Free
Set the Error Covariance of cigr96 and alch94 Free
Set the Error Covariance of cigr96 and cigr94 Free
Set the Error Covariance of cigr96 and alch96 Free
LISREL OUTPUT: AD=OFF ND=3 SC PS=TSERIES.COV
Path Diagram
End of Problem
```

The second part of the syntax file lists the indicators for the 1994 and 1996 latent variables. This is followed by two commands which indicate that these latent variables are linearly dependent on the gender, age, and racial group of a child.

The LISREL OUTPUT: command is used to specify a completely standardized solution and to save the covariance matrix of the ten latent variables to the text file **tseries.cov**.

The path diagram obtained for this analysis is shown below.



Chi-Square=3698.57, df=347, P-value=0.00000, RMSEA=0.050