Scientific Software International

Estimating residuals using the Kenny-Judd data

1. Introduction

Here we use the Kenny-Judd data and estimate residuals for these data. The Kenny–Judd model is

$$y = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta, \quad (1)$$

where ξ_1 and ξ_2 are latent variables and ζ is a random error term assumed to be uncorrelated with ξ_1 and ξ_2. Kenny & Judd (1984) considered the case when there are two observable indicators x_1 and x_2 of ξ_1 and two observable indicators x_3 and x_4 of ξ_2, such that

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \lambda_2 & 0 \\ 0 & 1 \\ 0 & \lambda_4 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}. \quad (2)$$

Kenny & Judd (1984) did not include the constant intercept terms α and τ_i in (1) and (2), but as argued by Jöreskog & Yang (1996), these are necessary for a correct analysis of the model.

Solving for ξ_1 and ξ_2 in the first and third equation in (2) and substituting into (1) gives

$$y_1 = \alpha + \gamma_1 x_1 + \gamma_2 x_3 + \gamma_3 x_1 x_3 + u, \quad (3)$$

where

$$u = -\gamma_1\delta_1 - \gamma_2\delta_3 - \gamma_3\left(x_1\delta_3 + x_3\delta_1 - \delta_1\delta_3\right) + \zeta. \quad (4)$$

Note that the γ's in (3) are the same as in (1) but the error term u is not the same as ζ. The error term u in (3) is correlated with the variables on the right side in (3) so that ordinary least squares cannot be used to estimate this equation. However, as shown by Bollen & Paxton (1998), x_2, x_4, and x_2x_4 can be used as instrumental variables. If the model holds, these instrumental variables are uncorrelated with u.

As a first step, we estimate the residuals using the data in the LSF file **kjudd.lsf**. The first few lines of this file are shown below.

kjudd.lsf

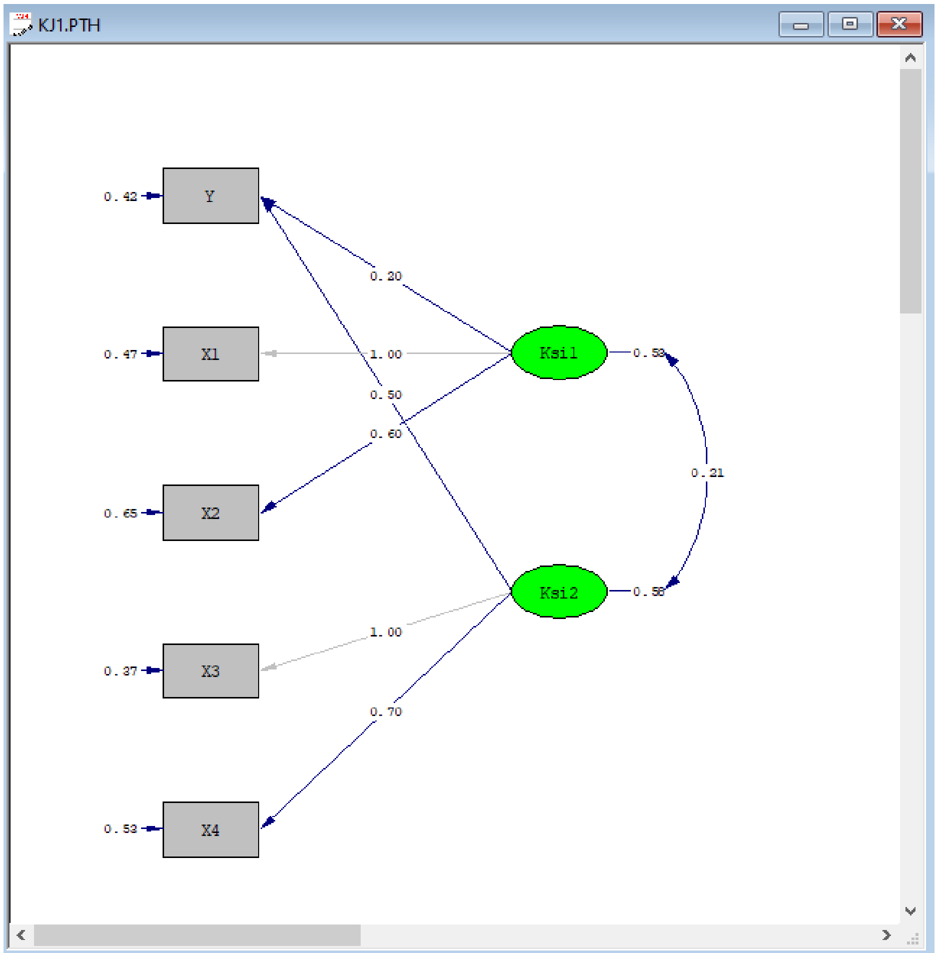
	Y	X1	X2	X3	X4
1	1.70	2.13	-0.37	0.04	0.15
2	1.76	0.65	-1.05	0.10	1.32
3	1.94	1.61	1.25	0.73	-0.08
4	0.86	-0.44	-0.32	0.37	-0.69
5	0.75	1.06	-0.46	-0.88	0.40
6	0.90	-0.51	-0.30	-0.12	1.21
7	1.33	-0.83	-1.33	1.01	0.56
8	1.00	-1.21	-0.19	0.71	-0.22
9	0.63	-0.61	-0.97	0.23	-0.37
10	2.05	-0.51	0.77	1.12	-0.83
11	0.67	-0.28	-0.26	-0.16	-0.67
12	0.92	-0.33	1.24	-0.38	-0.14
13	1.64	0.24	1.39	0.05	0.09
14	1.09	-0.87	-1.30	0.19	-1.79
15	-0.02	-0.43	0.24	-1.71	-0.31
16	0.24	-0.72	-0.92	0.81	0.19
17	1.07	0.60	0.02	1.82	-1.03
18	0.68	-0.68	0.65	-1.77	-0.09
19	2.01	0.50	0.52	2.19	0.07
20	1.61	-0.78	-2.29	-1.24	0.68

We use the syntax file **KJ1.SPL** to do so.

KJ1.SPL

```
Raw data from file kjudd.lsf
Latent Variables: Ksi1 Ksi2
Relationships
Y=Ksi1 Ksi2
X1=1*Ksi1
X2=Ksi1
X3=1*Ksi2
X4=Ksi2
Path Diagram
LSFfile KJUDD.LSF
Estimate Residuals
End of Problem
```

The path diagram shows the estimates obtained under this model.



The maximum likelihood estimates for this model are given below.

```
LISREL Estimates (Maximum Likelihood)

         Measurement Equations


        Y = 0.204*Ksi1 + 0.500*Ksi2, Errorvar.= 0.424  , R² = 0.331
 Standerr  (0.0560)     (0.0552)                (0.0237)
 Z-values   3.645        9.053                   17.853
 P-values   0.000        0.000                   0.000

       X1 = 1.000*Ksi1, Errorvar.= 0.474  , R² = 0.526
 Standerr                         (0.0863)
 Z-values                          5.496
 P-values                          0.000

       X2 = 0.599*Ksi1, Errorvar.= 0.646  , R² = 0.226
 Standerr  (0.0974)               (0.0417)
 Z-values   6.147                  15.499
 P-values   0.000                  0.000

       X3 = 1.000*Ksi2, Errorvar.= 0.372  , R² = 0.609
 Standerr                         (0.0494)
 Z-values                          7.523
 P-values                          0.000
```

```
      X4 = 0.703*Ksi2, Errorvar.= 0.532  , R² = 0.350
 Standerr  (0.0621)               (0.0331)
 Z-values   11.331                16.055
 P-values   0.000                 0.000
```

```
        Covariance Matrix of Independent Variables

                  Ksi1       Ksi2
              --------   --------
     Ksi1        0.527
               (0.092)
                 5.698

     Ksi2        0.211      0.578
               (0.030)    (0.061)
                 7.067      9.514
```

The LSF file **KJUDDnew.lsf** shown below is created as part of the output from this run. The variables R_X1 to R_X4 are the estimates of the measurement errors δ_1 to δ_4, while R_Y represents the estimate of the measurement error ε_1 .

KJUDDnew.lsf

	Y	X1	X2	X3	X4	Ksi1	Ksi2	R_Y	R_X1	R_X2	R_X3	R_X4
1	1.70	2.13	-0.37	0.04	0.15	1.17	0.19	0.19	0.92	-1.07	-0.17	0.02
2	1.76	0.65	-1.05	0.10	1.32	0.17	0.52	0.28	0.44	-1.15	-0.44	0.96
3	1.94	1.61	1.25	0.73	-0.08	1.32	0.55	0.22	0.25	0.46	0.16	-0.46
4	0.86	-0.44	-0.32	0.37	-0.69	-0.40	-0.07	-0.20	-0.08	-0.08	0.42	-0.63
5	0.75	1.06	-0.46	-0.88	0.40	0.40	-0.46	-0.28	0.62	-0.70	-0.44	0.73
6	0.90	-0.51	-0.30	-0.12	1.21	-0.42	0.17	-0.28	-0.13	-0.04	-0.31	1.09
7	1.33	-0.83	-1.33	1.01	0.56	-0.80	0.69	-0.03	-0.07	-0.85	0.30	0.08
8	1.00	-1.21	-0.19	0.71	-0.22	-0.78	0.26	-0.15	-0.47	0.28	0.44	-0.39
9	0.63	-0.61	-0.97	0.23	-0.37	-0.69	-0.12	-0.35	0.04	-0.55	0.33	-0.28
10	2.05	-0.51	0.77	1.12	-0.83	0.00	0.57	0.58	-0.56	0.77	0.53	-1.22
11	0.67	-0.28	-0.26	-0.16	-0.67	-0.32	-0.39	-0.25	-0.00	-0.06	0.21	-0.39
12	0.92	-0.33	1.24	-0.38	-0.14	0.06	-0.30	-0.12	-0.43	1.21	-0.09	0.08
13	1.64	0.24	1.39	0.05	0.09	0.53	0.16	0.27	-0.33	1.08	-0.12	-0.01
14	1.09	-0.87	-1.30	0.19	-1.79	-0.86	-0.40	0.29	-0.05	-0.78	0.58	-1.50
15	-0.02	-0.43	0.24	-1.71	-0.31	-0.39	-1.27	-0.49	-0.08	0.48	-0.46	0.59
16	0.24	-0.72	-0.92	0.81	0.19	-0.78	0.24	-0.90	0.02	-0.45	0.55	0.03
17	1.07	0.60	0.02	1.82	-1.03	0.32	0.66	-0.51	0.24	-0.17	1.14	-1.49
18	0.68	-0.68	0.65	-1.77	-0.09	-0.33	-1.08	0.11	-0.39	0.85	-0.70	0.68
19	2.01	0.50	0.52	2.19	0.07	0.53	1.36	0.04	-0.07	0.21	0.81	-0.88
20	1.61	-0.78	-2.29	-1.24	0.68	-0.99	-0.40	0.83	0.17	-1.69	-0.86	0.97

Looking at the univariate summary statistics for these variables, obtained by using the **Data Screening option**, we see that the range of the measurement error R_X3 is the largest (-2.697; 2.310).

```
Univariate Summary Statistics for Continuous Variables

 Variable      Mean  St. Dev.  Skewness  Kurtosis  Minimum Freq.  Maximum Freq.
 --------      ----  --------  --------  --------  ------- -----  ------- -----
        Y     1.181     0.796     1.208     3.612   -1.290     1    5.680     1
       X1     0.042     1.001     0.010    -0.097   -3.170     1    3.440     1
       X2    -0.003     0.914    -0.137     0.074   -3.960     1    2.440     2
       X3     0.018     0.975     0.014     0.019   -3.420     1    3.080     1
       X4    -0.006     0.904     0.054    -0.011   -2.760     1    3.080     1
     Ksi1    -0.000     0.726     0.074     0.064   -2.414     1    2.716     1
     Ksi2    -0.000     0.760     0.419     0.286   -1.926     1    2.634     1
      R_Y    -0.000     0.599     0.772     1.799   -1.738     1    2.828     1
```

```
R_X1    0.000     0.406    -0.004     0.194   -1.321     1    1.512     1
R_X2    0.000     0.731    -0.058    -0.044   -2.697     1    2.310     1
R_X3    0.000     0.399    -0.234     0.122   -1.653     1    1.139     1
R_X4    0.000     0.650    -0.033     0.010   -2.073     1    2.015     1
```

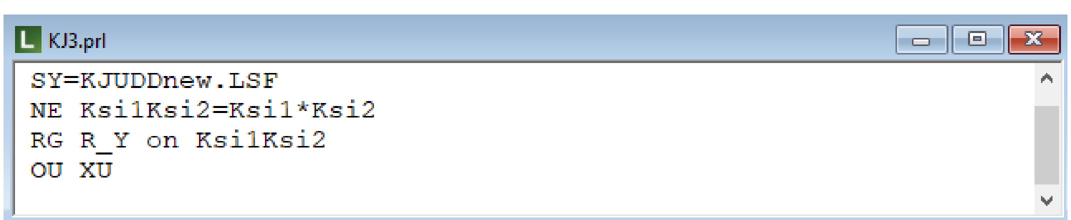
Next, we estimate the regression of Y on Ksi1, Ksi2 and Ksi1*Ksi2 using the syntax file **KJ2.PRL** and the newly created LSF file **KJUDDnew.lsf**.

```
Estimating the regression of Y on Ksi1, Ksi2 and Ksi1*Ksi2
SY=KJUDDnew.LSF
NE Ksi1Ksi2 = Ksi1*Ksi2
CO ALL
RG Y on Ksi1 Ksi2 Ksi1Ksi2
OU XU
```

The output for this analysis is as follows:

```
Estimated Equations

        Y = 1.085 + 0.198*Ksi1 + 0.482*Ksi2 + 0.458*Ksi1Ksi2
 Standerr  (0.0174) (0.0246)     (0.0242)     (0.0263)
 t-values   62.380   8.032        19.918       17.410
 P-values   0.000    0.000        0.000        0.000

             + Error, R² = 0.572

  Error Variance = 0.272


 The following chi-squares test the hypothesis that all
 regression coefficients are zero except the intercept.

 Variable       -2lnL  Chi-square   df  Covariates
 --------  ----------  ----------   --  ----------
        Y    1531.087     849.125    3   Ksi1 Ksi2 Ksi1Ksi2

Analysis of Variance Table

   Regression d.f.    Residual d.f.          F  Covariates
   ---------------    -------------          -  ----------
      362.072    3     270.687  996     444.083  Ksi1 Ksi2 Ksi1Ksi2
```

We are also interested in estimating the regression of R_Y, the estimate of the measurement error ε_1, on Ksi1*Ks2. The file **KJ3.PRL** is used for this purpose.



```
SY=KJUDDnew.LSF
NE Ksi1Ksi2=Ksi1*Ksi2
RG R_Y on Ksi1Ksi2
OU XU
```

For this model we find that the estimated equations are

Estimated Equations

```
     R_Y =  - 0.0949 + 0.451*Ksi1Ksi2 + Error, R² = 0.244
Standerr     (0.0173) (0.0251)
t-values     -5.481    17.931
P-values      0.000    0.000

Error Variance = 0.271


The following chi-squares test the hypothesis that all
regression coefficients are zero except the intercept.
Variable       -2lnL  Chi-square   df  Covariates
--------  ----------  ----------   --  ----------
     R_Y    1531.959     279.265    1   Ksi1Ksi2
```

Analysis of Variance Table

```
   Regression d.f.    Residual d.f.           F  Covariates
   ---------------    -------------           -  ----------
       87.280    1     270.924  998     321.513  Ksi1Ksi2
```

From the estimated equation, we conclude that both the intercept and estimated coefficient of the interaction term is not significant.