Implementation of sampling weights in a linear growth curve model

Contents

1 Describing the data 1

2. Implementation of sampling weights in a linear growth curve model 3

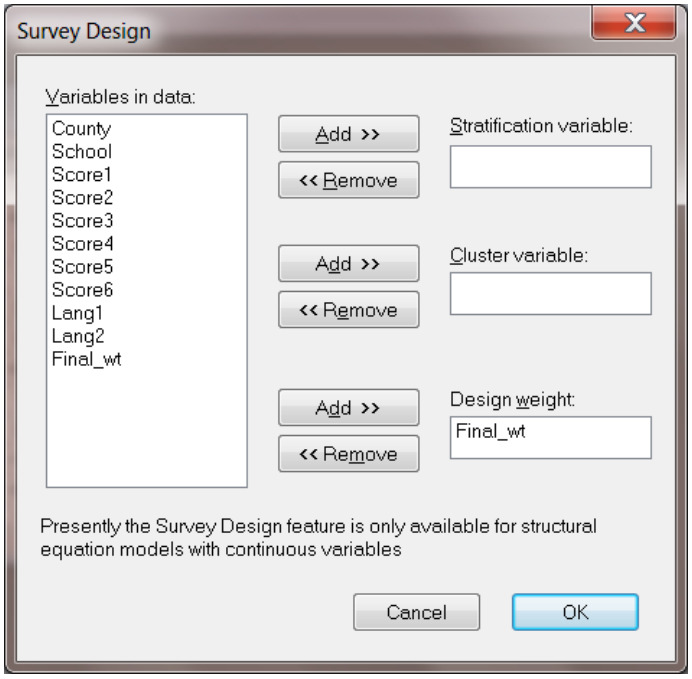
2.1 Discussion of results 10

3. Incorporating stratum and cluster variables in the model 12

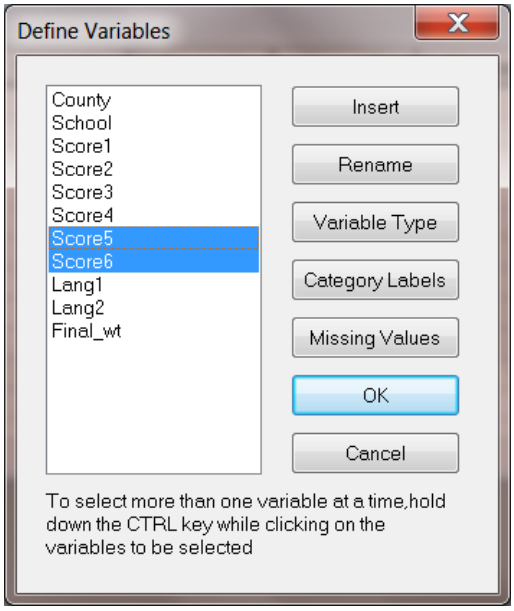
3.1 Discussion of results 12

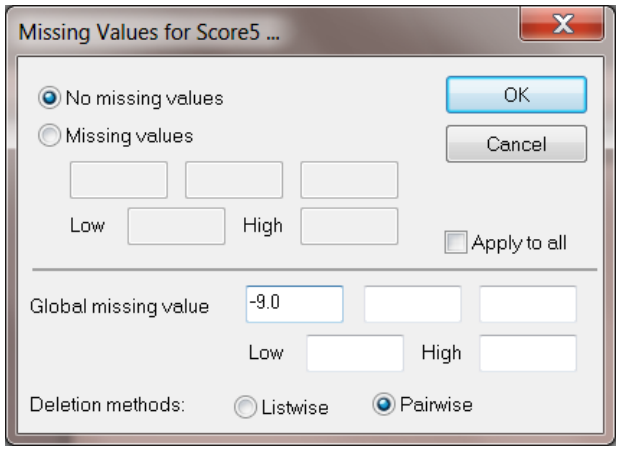
1 Describing the data

We use the data **surveysem.lsf** located in the **Complex survey sample examples** folder to illustrate the implementation of sampling weights and the use of of stratum/cluster variables. We start by opening this file. From the main menu bar, select the **Data**, **Survey Design** option and add the variable Final_wt in the **Design weight*:*** box. Click **OK** when done.



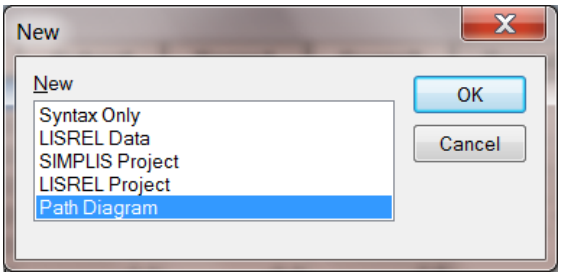
Next use the **Data**, **Define Variables** option to select the variables Score5 and Score6, and then click the **Missing Values** button to invoke the **Missing Values** dialog box. Enter -9.0 as shown below. Click **OK**, and then use the **File**, **Save** option to ensure that these changes are contained in the LSF file.



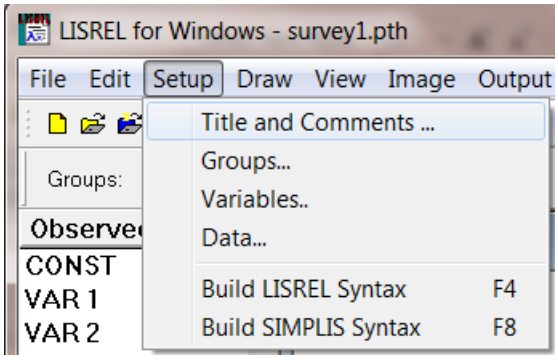


2. Implementation of sampling weights in a linear growth curve model

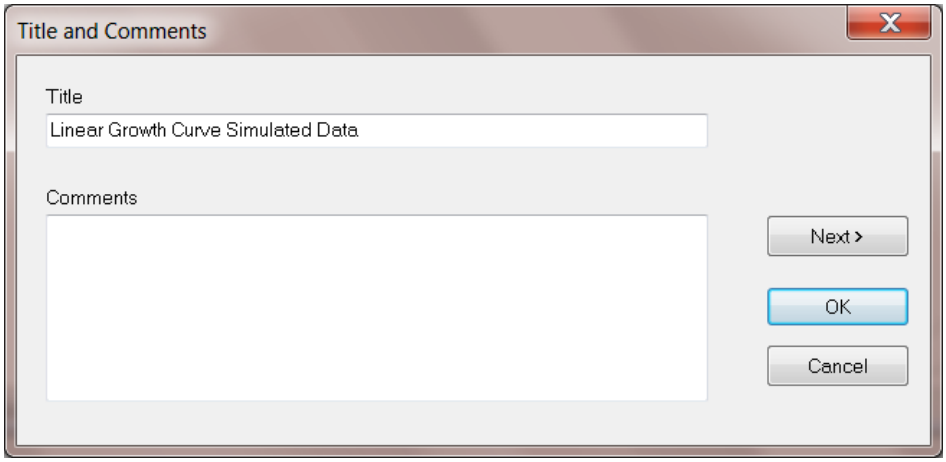
To generate the SIMPLIS commands interactively, we proceed as follows. Using the **File**, **New** option, select the **Path Diagram** option from the **New** dialog box.



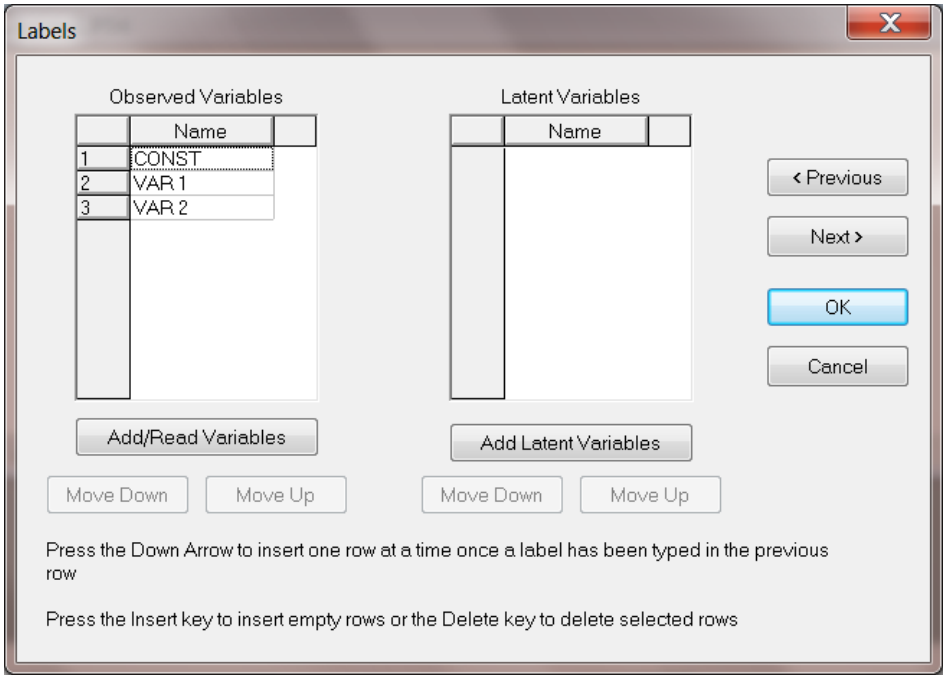
Save the new path diagram in the **Complex survey sample examples** folder as **Survey1.pth**. Click **Save** when done. From the **Setup** menu, select **Title and Comments**.



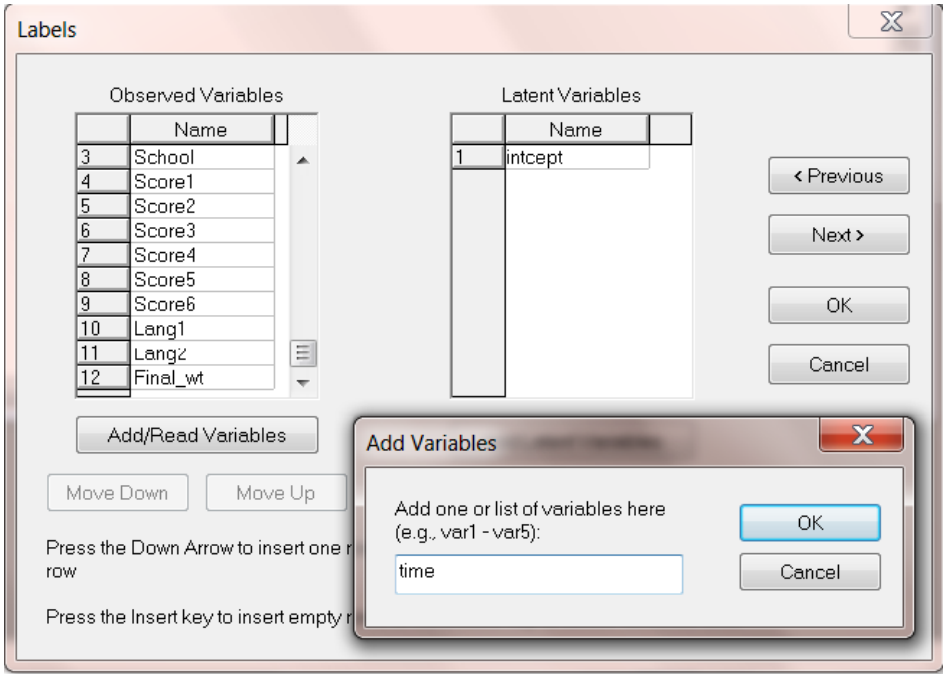
This action loads the **Title and Comments** dialog box shown below. Enter a title and any (optional) comments as shown below.



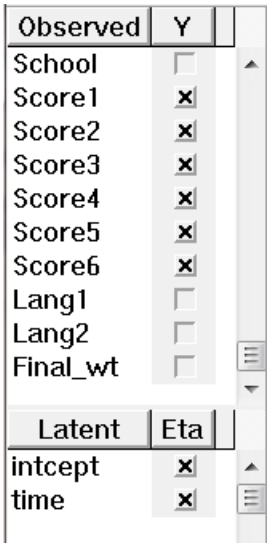
Click **Next** to proceed to the **Groups** dialog and, since this is a single-group example, click **Next** again to activate the **Labels** dialog shown below. To add a list of observed variables, click the **Add/Read Variables** button below the **Observed Variables** list box.



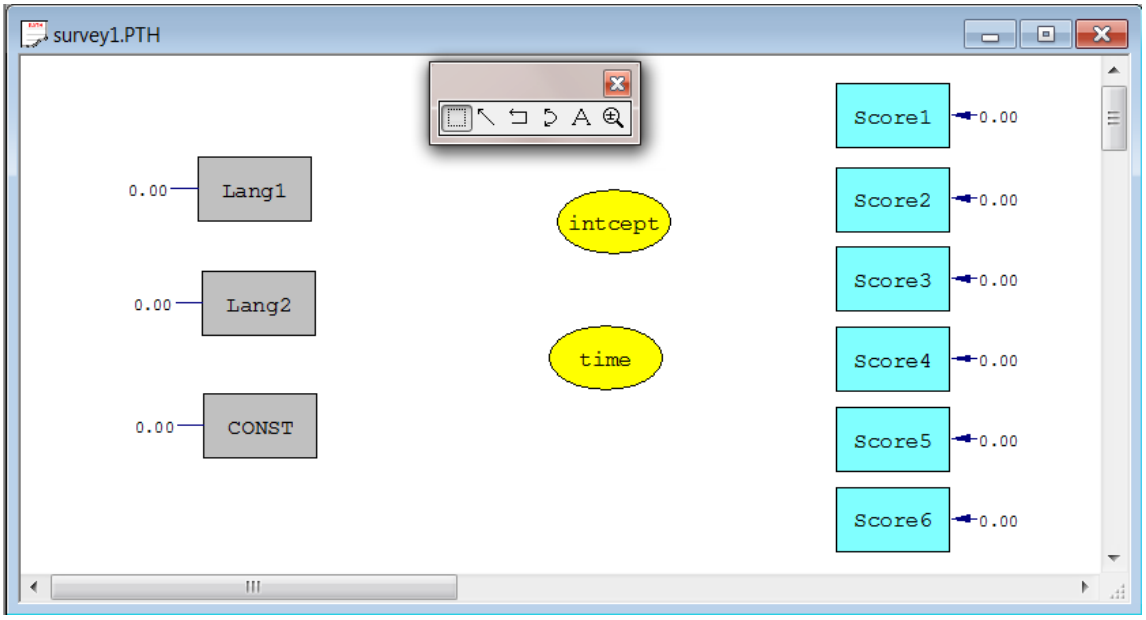
From the **Add/Read Variables** dialog box, click the **Read from file:** radio button and select **Lisrel System File**. Next, use the **Browse** button to locate and select **surveysem.LSF**. Click the **OK** button once the desired LSF is selected. The observed variable names will be displayed in the **Labels** dialog box. To add a list of latent variables, click the **Add Latent Variables** button and type the names of the latent variables one at a time.



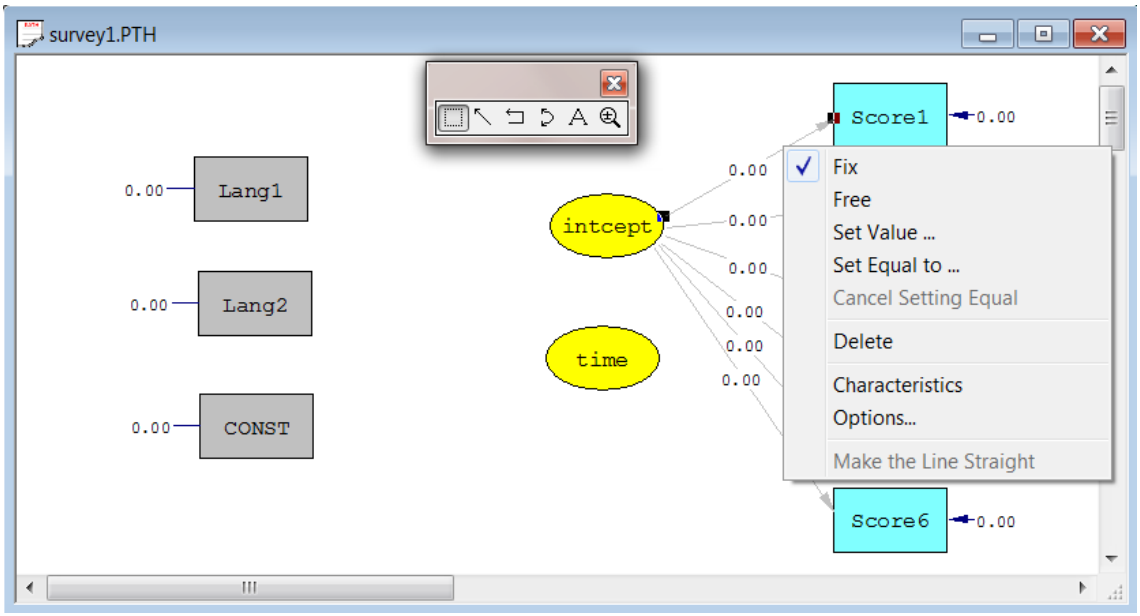
Click **OK** when the latent variables intcept and time are entered. The left hand side of the path diagram window should display the observed and latent variables. If not, select the **View**, **Toolbars** option and from the drop-down list **Select Variables**. Click on the check boxes on the right hand side of the variable names to define Score1 to Score6 as Y (dependent) variables and intcept and time as Eta (endogenous latent) variables.



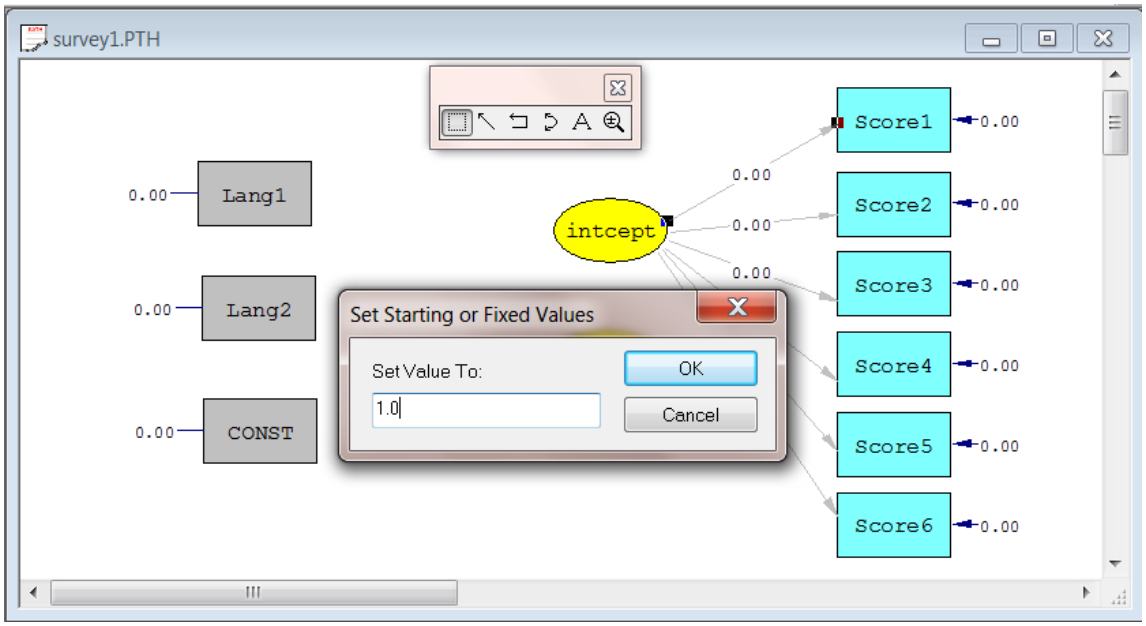
We start drawing the path diagram by dragging the names Score1 to Score6 to the path diagram window. Next, drag intcept and time to the middle of the path diagram window. A variable is dragged to the path diagram window by left-clicking on the variable name and then moving it with the left mouse button held down. Finally, drag Lang1, Lang2 and CONST to the left of intcept and time. Click the one-sided arrow on the drawing bar and connect arrows from intcept to the variables Score1, …, Score6.



With the left mouse button down, start in the ellipse and drag the arrow to within a rectangle representing one of the Score variables before releasing the mouse button. Unselect the arrow by clicking on the square on the drawing bar. Once this is done, move the mouse pointer to each arrow and right-click. From the pop-up menu, select **Fix**. Repeat this for each path from intcept to a Score value.

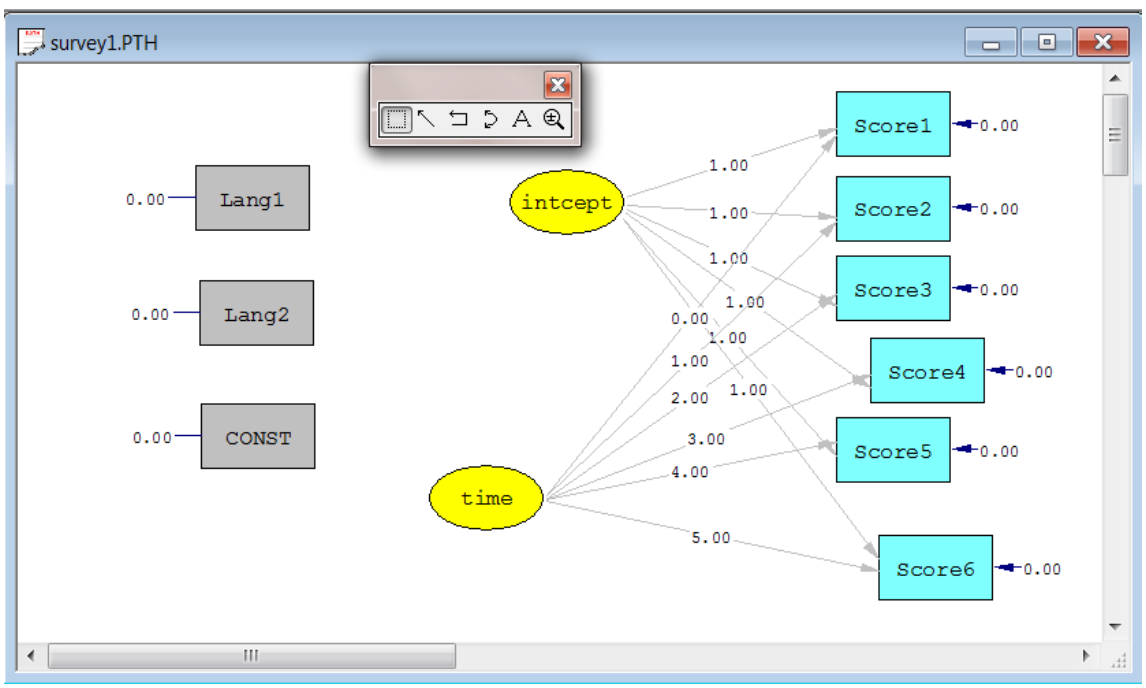


Once all the paths are fixed, start with the path from intcept to Score1. Right-click on the arrow, and select the **Set Value** option from the pop-up menu.

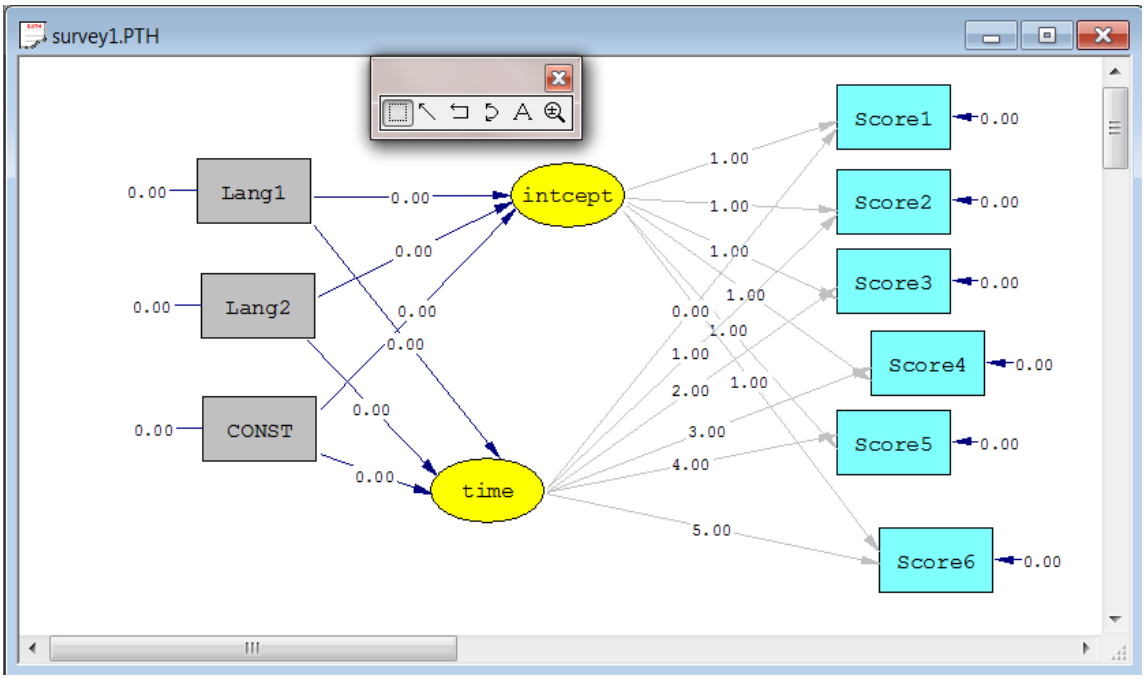


Change the value of each path to 1.0.

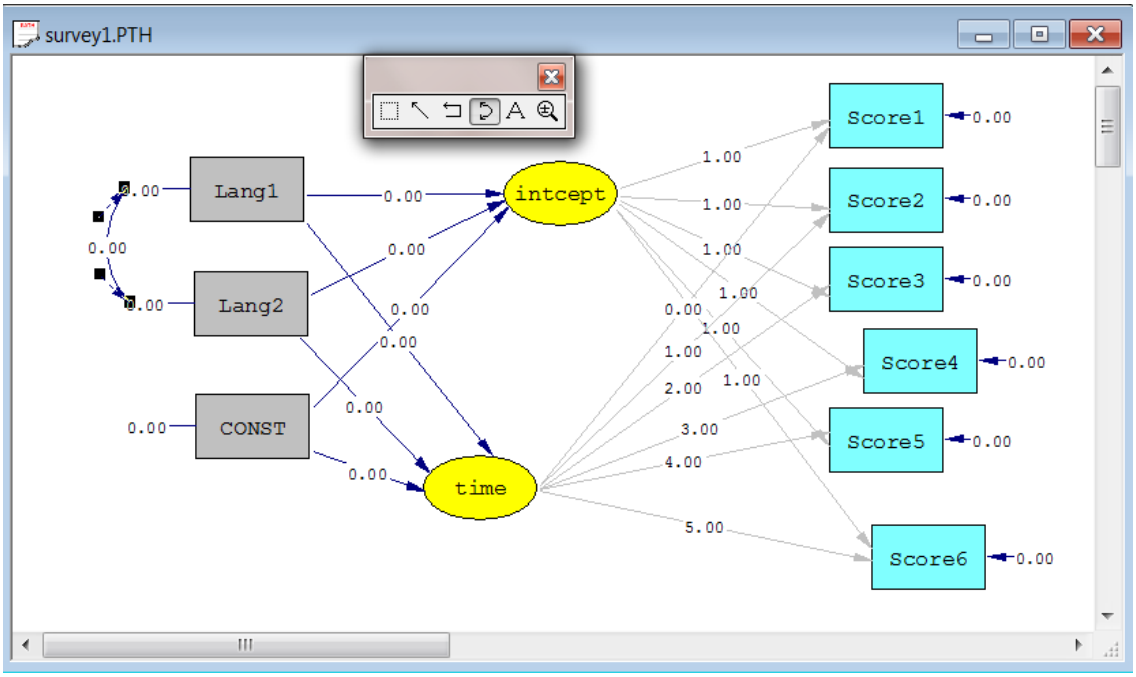
Repeat the above procedure by drawing paths from time to Score2, Score3, …, Score6. Fix these paths and set the path coefficients to 1.0, 2.0, 3.0, 4.0, and 5.0 respectively, as shown below.



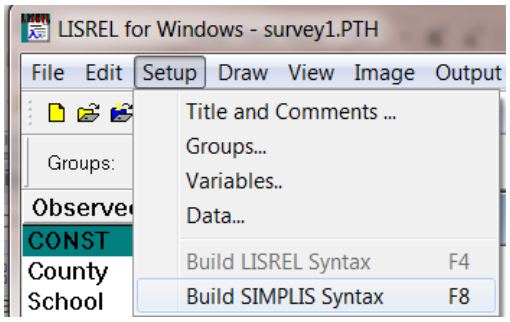
The path diagram is completed by drawing arrows from Lang1, Lang2 and the SIMPLIS variable CONST to intcept and time as shown. Once this is done, select the two-sided arrow (error covariance or factor correlation) to add a covariance path between Lang1 and Lang2.



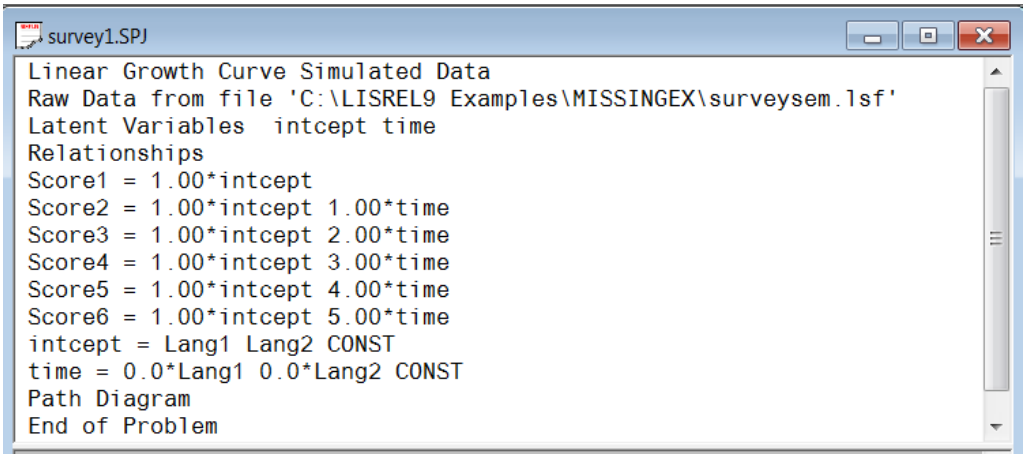
To draw this path, left-click on the horizontal link between 0.00 and the Lang1 rectangle. Drag the path to the line connecting 0.00 and Lang2 before releasing the mouse button.



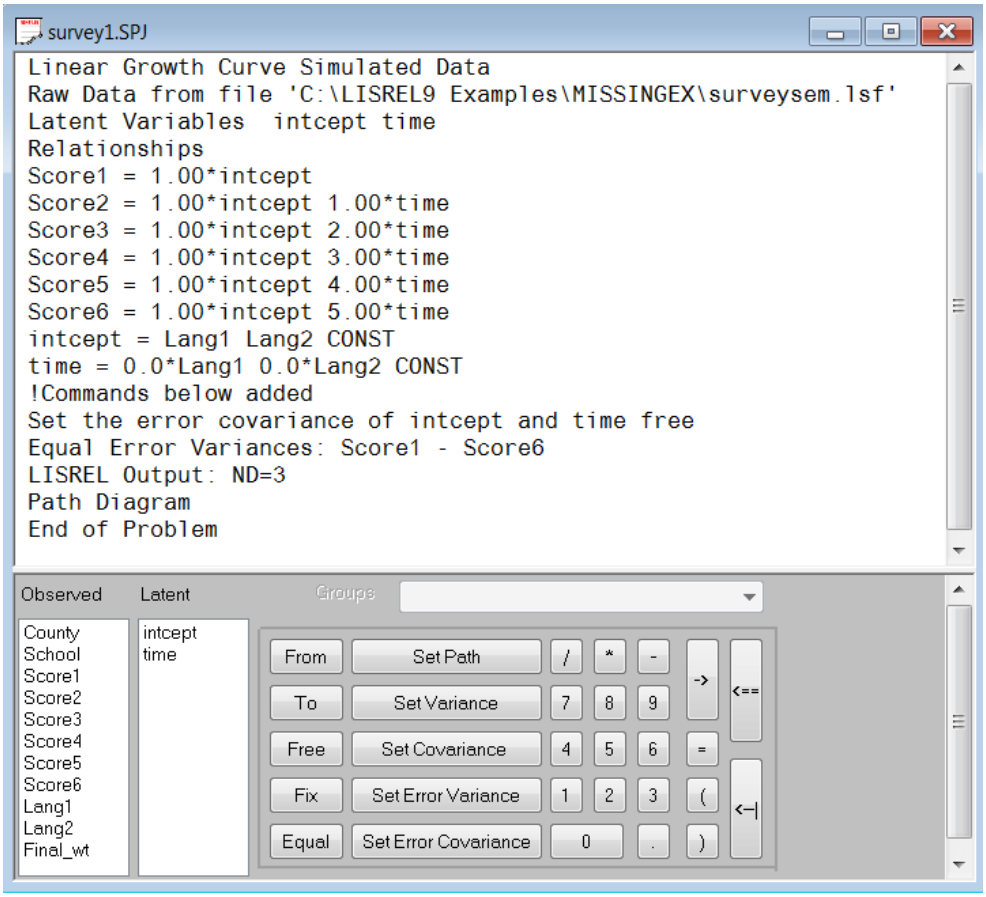
To build the corresponding SIMPLIS syntax, select **Setup**, **Build SIMPLIS Syntax** from the main menu.



The syntax shown below is generated.



In our model, we assume that the coefficient of the latent variable time is not influenced by Lang1 or Lang2 and we change the syntax by adding 0.00* in front of Lang1 and Lang2 as shown below.



Since it is assumed that the error variances associated with the dependent variables are homogeneous, we manually add the command

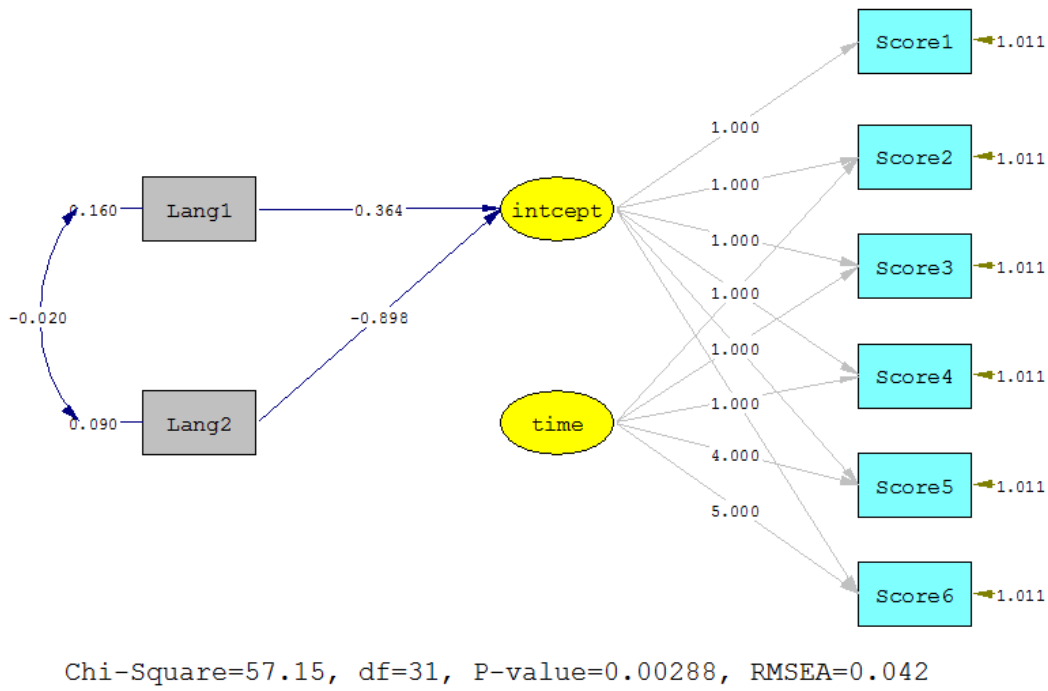
Equal Error Variances: Score1 – Score6

In addition, it is assumed that intcept and time are correlated. The equivalent SIMPLIS command is also typed in, as shown in the syntax file.

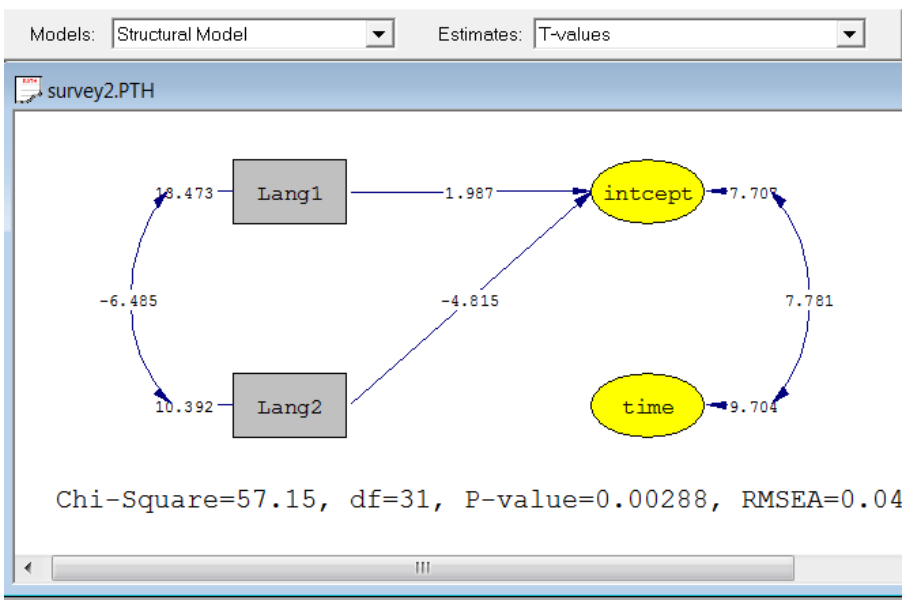
Once these modifications to the command file are completed, click the **Run LISREL** icon to obtain a path diagram.

2.1 Discussion of results

The path diagram shows the parameter estimates and χ^2 goodness of fit statistic (χ^2 = 57.15, df = 31) under the assumption that stratification and selection of clusters do not affect standard errors or model fit.

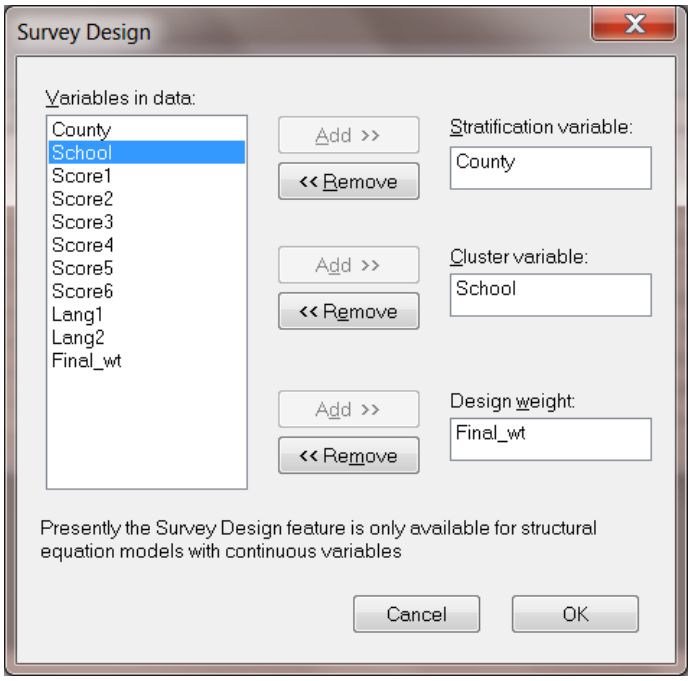


In the display below, the structural part of the model is shown. The numeric values are the *t*-values (*t*: parameter estimate/std. error) for each path. For example; the *t*-value for $\hat{\gamma}_1 = 0.36$ equals 1.99.



3. Incorporating stratum and cluster variables in the model

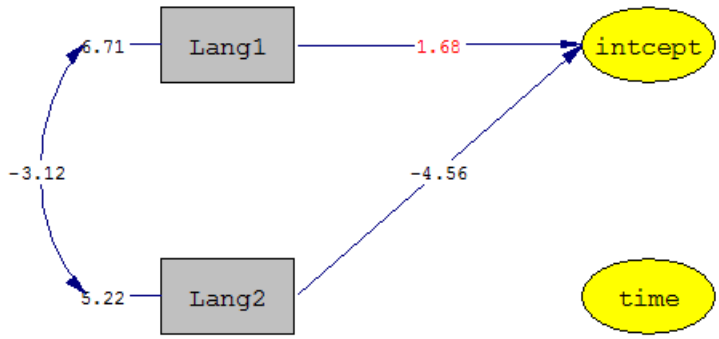
Next, select the **Survey Design** option from the **Data** menu and add County and School respectively as stratification and cluster variables. Click **OK** and then select the **File**, **Save** option.



Once this is done, make **survey1.spj** the active window before clicking the **Run Lisrel** icon.

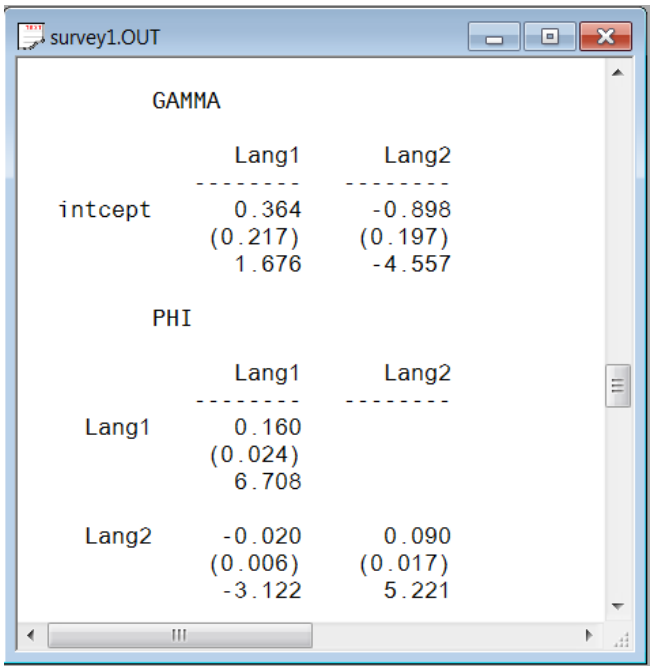
3.1 Discussion of results

From the path diagram below we note that the χ^2 statistic has increased from 57.15 to 62.76 and that the *t*-value corresponding to $\hat{\gamma}_1$ decreased from 1.99 to 1.68. In the latter case, the *t*-value indicates a non-significant coefficient for the Lang1, intcept path.



Chi-Square=62.76, df=31, P-value=0.00063, RMSEA=0.046

Finally, selected parts of the LISREL output are shown below.



survey1.OUT

```
        GAMMA

                Lang1       Lang2
              --------    --------
intcept         0.364      -0.898
               (0.217)     (0.197)
                1.676      -4.557

        PHI

                Lang1       Lang2
              --------    --------
  Lang1         0.160
               (0.024)
                6.708

  Lang2        -0.020       0.090
               (0.006)     (0.017)
               -3.122       5.221
```

The 95% confidence intervals for the estimated parameters show that the population parameter values are all contained in the respective intervals.

survey1.OUT

```
        PSI

            intcept       time
           --------   --------
intcept       2.005
             (0.285)
               7.032

   time       0.661      0.418
             (0.087)    (0.039)
               7.598     10.863


        THETA-EPS

             Score1     Score2     Score3     Score4     Score5     Score6
           --------   --------   --------   --------   --------   --------
               1.011      1.011      1.011      1.011      1.011      1.011
             (0.039)    (0.039)    (0.039)    (0.039)    (0.039)    (0.039)
              25.895     25.895     25.895     25.895     25.895     25.895

        ALPHA

            intcept       time
           --------   --------
               0.958      0.523
             (0.148)    (0.047)
               6.492     11.025
```

The χ^2 goodness-of-fit statistic (see below) equals 62.757. This value is obtained as the difference between the deviance values for the saturated and fitted models, multiplied by the chi-squared scale factor. That is (9484.714 - 9408.501)*0.82344. The RMSEA value of 0.0462 indicates that the model provides a reasonable fit to the data.

survey1.OUT

```
              Global Goodness of Fit Statistics, FIML case

              Chi-Square Scale Factor =      0.82344
              Number of Strata   =               48
              Number of Clusters =              144


                -2ln(L) for the saturated model =        9408.501
                -2ln(L) for the fitted model    =        9484.714

Degrees of Freedom = 31
Full Information ML Chi-Square                        62.757 (P = 0.0006)
Root Mean Square Error of Approximation (RMSEA)       0.0462
90 Percent Confidence Interval for RMSEA              (0.0295 ; 0.0626)
P-Value for Test of Close Fit (RMSEA < 0.05)          0.626
```