

## Supplementary Material

### Sample code for implementing linear spline multilevel models using “runmlwin” command

This syntax utilises the user-written command ‘runmlwin’ which must be installed prior to use. The most recent version of MLwiN must be installed to be able to use this command and this package is available for use within Stata and R. Below we demonstrate the basic steps involved in implementing linear spline multilevel modelling using “runmlwin” in Stata. Code below assumes data are in long format and that a variable called “occasion” exists identifying the ordering of observations within individuals. Sample code below applies to length/height from birth to five years.

#### Generate the spline variable

First, three new variables are created: s1 (spline 1 from birth to 6 months), s2 (6 months to 2 year), s3 (2 years to 5 years).

```
mkspline s1_birth_6m 27 s2_6m_2 107 s3_2_max = age_lw
```

#### Generate a constant term

MLwiN does not automatically include a constant term, so this must be generated and included in models.

```
gen cons=1
```

#### Identify the location of MLwiN

```
global MLwiN_path "C:\Program Files\MLwiN v3.05\mlwin.exe"
```

#### Run the multilevel model, sorting the data by person and occasion/age first.

```
sort study_id age
runmlwin length cons s1_birth_6m s2_6m_2 s3_2_max ///
level2 (study_id: cons s1_birth_6m s2_6m_2 s3_2_max , reset(var) residuals (res, var)) ///
level1 (occ: age_lw, reset(var) diag) nopause maxiterations(150)
```

#### Adding covariates

The following assumes covariates are binary and coded 0 and 1 or for covariates with multiple categories, dummy variables have been created. The addition of continuous covariates should be undertaken in the same manner as for categorical covariates but continuous covariates should be centred on the mean so that the baseline trajectory in the model is for the individuals with the mean level of the continuous covariate. Here we demonstrate the steps required for addition of sex as a covariate.

### **Multiply covariate by splines**

Once the covariate is coded in the format of 0/1 representing 0 for the baseline category, we multiply the covariate by the splines, creating interaction terms for inclusion in our model.

```
gen s1_birth_6m_fem = s1_birth_6m*female  
gen s2_6m_2_fem = s2_6m_2*female  
gen s3_2_max_fem = s3_2_max*female
```

### **Run model now including covariate terms**

The model is then ran as before but this time including a term for the covariate in question, here “female” and each of the above female\*spline interaction terms generated. This allows the mean trajectory to differ for females and males. Because in this example the variable female is coded 0 for male and 1 for female the baseline trajectory is now for males with coefficients for “female”, s1\_birth\_6m\_fem, s2\_6m\_2\_fem, s3\_2\_max\_fem representing the difference in the intercept, spline 1 and spline 2 and spline 3 in females compared with males.

```
sort study_id age  
runmlwin length cons s1_birth_6m s2_6m_2 s3_2_max female2*, ///  
level2 (study_id: cons s1_birth_6m s2_6m_2 s3_2_max , reset(var) residuals (res, var)) ///  
level1 (occ: age_lw, reset(var) diag) nopause maxiterations(150)
```