Online Supporting Materials

Estimation of EEM parameters

In the EEM, the estimator of the slope $\beta_i$ depends on the values of $(\sigma_{uw}, \sigma_{ww}, \sigma_{uu})$. The 2 within-person variances and the covariance for each sample individual can be estimated empirically so that

$$\hat{\sigma}_{wwi} = \frac{1}{n_i-1} \Sigma_{i=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2,$$

$$\hat{\sigma}_{uui} = \frac{1}{n_i-1} \Sigma_{i=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

and

$$\hat{\sigma}_{ui} = \frac{1}{n_i-1} \Sigma_{i=1}^{n_i} (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_i),$$

where $\bar{Y}_i = \frac{1}{n_i} \Sigma_{j=1}^{n_i} Y_{ij}$ and $\bar{X}_i = \frac{1}{n_i} \Sigma_{j=1}^{n_i} X_{ij}$ are the person-level means of the observed daily response and observed daily FGI scores, respectively, for person $i$ and $n_i$ is the number of daily observations available for the $i^{th}$ person. Individuals who are surveyed on 1 day only ($n_i = 1$) do not contribute to the estimation of the within-person variances and covariance, but do contribute to estimation of the between-person variances and to estimation of the slope of the regression of the response on FGI.

A single population-level estimate of the within-person variances and covariance is given by averaging the within-person variances over the $N_1$ individuals with more than 1 observation day:

$$\hat{\sigma}_{ww} = \frac{1}{N_1} \Sigma_{i=1}^{N_1} \hat{\sigma}_{wwi},$$

$$\hat{\sigma}_{uui} = \frac{1}{N_1} \Sigma_{i=1}^{N_1} \hat{\sigma}_{uui},$$

and

$$\hat{\sigma}_{ui} = \frac{1}{N_1} \Sigma_{i=1}^{N_1} \hat{\sigma}_{ui},$$

It is necessary to obtain estimates of the measurement error variances or, equivalently, of the variances of $(\bar{Y}_i - y_i)$ and $(\bar{X}_i - x_i)$. For individuals with $n_i = 2$ days of observation, the 2 estimated within-person variances are:

$$\text{Var}(\bar{Y}_i - y_i) = \hat{\sigma}_{yy} + \frac{\hat{\sigma}_{ww}}{2} - \hat{\sigma}_{yy} = \frac{\hat{\sigma}_{ww}}{2} \quad \text{and} \quad \text{Var}(\bar{X}_i - x_i) = \hat{\sigma}_{xx} + \frac{\hat{\sigma}_{uu}}{2} - \hat{\sigma}_{xx} = \frac{\hat{\sigma}_{uu}}{2}.$$

For individuals with only 1 day of observation, the 2 estimated within-person variances are:

$$\text{Var}(\bar{Y}_i - y_i) = \hat{\sigma}_{ww} \quad \text{and} \quad \text{Var}(\bar{X}_i - x_i) = \hat{\sigma}_{uui}.$$
where $\hat{\sigma}_{yy}$ and $\hat{\sigma}_{xx}$ are estimates of $\sigma_{yy}$ and $\sigma_{xx}$, respectively. The estimates for $\hat{\sigma}_{ww}, \hat{\sigma}_{uu}$ and $\hat{\sigma}_{uw}$, respectively, are the averages of the variances accounting for the number of observation days $n_i$ for each individual:

\[
\hat{\sigma}_{ww} = \frac{1}{2N} \sum_{i=1}^{N} \frac{\hat{\sigma}_{ww}}{n_i} = \hat{\sigma}_{ww} \left( \frac{N_1 + 2N_2}{N} \right),
\]

\[
\hat{\sigma}_{uu} = \frac{1}{2N} \sum_{i=1}^{N} \frac{\hat{\sigma}_{uu}}{n_i} = \hat{\sigma}_{uu} \left( \frac{N_1 + 2N_2}{N} \right),
\]

and

\[
\hat{\sigma}_{uw} = \frac{1}{2N} \sum_{i=1}^{N} \frac{\hat{\sigma}_{uw}}{n_i} = \hat{\sigma}_{uw} \left( \frac{N_1 + 2N_2}{N} \right),
\]

where $N_1$ is the number of individuals with $n_i = 2$ days of observation and $N_2 = N - N_1$ is the number of individuals with only 1 day of intake data.

The estimator for the regression coefficient in the EEM is similar to the OLS estimate in a SLR model with an adjustment that depends on the parameters of the joint distribution in expression 4 and 5 of the manuscript. By letting $\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i$ denote the overall mean (across days and individuals) of observed daily response and $\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$ denote the grand mean of observed FGI scores, the estimate for the slope $\hat{\beta}_1$ of the EEM is given by

\[
\hat{\beta}_1 = \frac{M_{XY} - \hat{\sigma}_{uw}}{M_{XX} - \hat{\sigma}_{uu}},
\]

where

\[
M_{XY} = \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y})
\]

and

\[
M_{XX} = \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2.
\]

When $\hat{\sigma}_{uw} = 0$, the estimate of the slope $\hat{\beta}_1$ is nearly the same as that given in the MEM (3) to be defined later. If both $\hat{\sigma}_{uw}, \hat{\sigma}_{uu} = 0$, then $\hat{\beta}_1$ reduces to the OLS estimate of $\beta_1$. 
Anticipating whether $\tilde{\beta}_1$ will be larger or smaller than the OLS estimate $\hat{\beta}_1$ is difficult because the sign and magnitude of $\beta_1 - \hat{\beta}_1$ depend on the sign of $\bar{u}$ and on the relative sizes of $\bar{u}$ and $\hat{u}$. This is in contrast to the usual measurement error-corrected estimate of $\beta_1$, which can typically be expected to be larger (in absolute value) than the attenuated OLS estimate $\hat{\beta}_1$ (4, 14).

The variance of $\beta_1$ in the EEM is estimated as

$$
\hat{V}ar(\tilde{\beta}_1) = \frac{1}{N} \left[ \frac{s_{yx}}{M_{xx}} + \frac{1}{M_{xx}} \left( \bar{\epsilon}_{uy} + (\bar{\epsilon}_{uu} - \tilde{\beta}_1 \bar{\epsilon}_{uu})^2 \right) \right]
\frac{1}{(N-1)M_{xx}} \left( \bar{\epsilon}_{uu} s_{rr} + (\bar{\epsilon}_{uw} - \tilde{\beta}_1 \bar{\epsilon}_{uu})^2 \right),
$$

where

$$
M_{xx} = M_{XX} - \bar{u}, \quad s_{yy} = \frac{1}{N-2} \sum_{i=1}^{N} \left( \bar{y}_i - \tilde{\beta}_0 - \tilde{\beta}_1 \bar{x}_i \right)^2, \quad \tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1 \bar{x},
$$

and

$$
s_{rr} = \bar{u} - 2 \tilde{\beta}_1 \bar{u} + \tilde{\beta}_1^2 \bar{u}.
$$

Finally, the variance of the error term in the EEM $\bar{\epsilon}_{qq}$ is estimated as

$$
\bar{\epsilon}_{qq} = s_{uv} - s_{rr}. \quad (5)
$$

Note if all measurement error variances and covariances equal 0, expression 5 reduces to the usual mean squared error (MSE) for estimating the error variance in a SLR model.

**Estimation of standard MEM parameters**

In the standard MEM, the unknown $\sigma_{uu}$ can be estimated using expression 2. The estimates for the regression coefficient are similar to the EEM estimate, but now the adjustment depends only on the variance of the measurement error in the predictor

$$
\tilde{\beta}_1 = \frac{m_{xy}}{m_{xx} - \bar{u}},
$$

where

$$
m_{xy} = \frac{NM_{xy}}{N-1},
$$
Online Supporting Materials

\[ m_{XX} = \frac{N m_{XX}}{N-1}, \]

and the error-corrected y-intercept is

\[ \tilde{\beta}_0 = \bar{Y} - \tilde{\beta}_1 \bar{X}, \]

with \( \bar{Y} \) and \( \bar{X} \) defined as before.

The estimate of the variance of the error-corrected slope is

\[ \text{Var}(\tilde{\beta}_1) = \frac{1}{N-1} \left[ \frac{m_{XX} \sigma_{YY} + \tilde{\beta}_1^2 \sigma_{uu}}{\sigma_{XX}^2} \right], \]

where the between-person variance of the usual predictor \( \sigma_{xx} \) is estimated as

\[ \hat{\sigma}_{xx} = m_{XX} - \hat{\sigma}_{uu}. \] \hspace{1cm} (6)

Lastly, model error \( \sigma_{qq} \) is estimated by

\[ \hat{\sigma}_{qq} = m_{yy} - m_{XY} \tilde{\beta}_1 \] \hspace{1cm} (7),

where

\[ m_{yy} = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \bar{Y})^2. \]

The variance estimates that result from equation 6 and 7 are not restricted to be positive and can thus take on negative values. This is often an indication that the true variances are close to 0. In samples that yield either,

\[ \hat{\sigma}_{xx} = \frac{m_{XX}^2}{m_{yy}}, \] \hspace{1cm} (8),

\[ \beta_1 = \frac{m_{yy}}{m_{XY}} \] \hspace{1cm} (9)

or \( \hat{\sigma}_{qq} = 0 \), the estimate for the error-corrected y-intercept is still as given above.
Supplemental Figure 1  (A) Original distributions for BLUP of MPA (B) and the normal scaled BLUP of MPA
Supplemental Figure 2  Studentized residual plots for the EEM regressing usual calcium intake on usual (A) FGI-13R, (B) FGI-21 and (C) FGI-21R
Supplemental Figure 3  Studentized residual plots for the SLR model of BLUP of MPA on (A) BLUP of FGI-13, (B) BLUP of FGI-21 and (C) BLUP of FGI-21R
Supplemental Figure 4  Studentized residual plots for the SLR model of BLUP of MPA on observed (A) FGI-13R, (B) FGI-21 and (C) FGI-21R