The EffectLiteR Approach for Analyzing Average and Conditional Effects – Supplementary Materials

Axel Mayer\textsuperscript{a}, Lisa Dietzfelbinger\textsuperscript{b}, Yves Rosseel\textsuperscript{a}, Rolf Steyer\textsuperscript{b}

\textsuperscript{a} Department of Data Analysis, Faculty of Psychology and Educational Sciences, Ghent University
\textsuperscript{b} Department of Methodology and Evaluation Research, Institute of Psychology, University of Jena

Supplement A: The EffectLiteR Software Package

A.1 Basic Commands

EffectLiteR is an R package (R Core Team, 2013) that can be used to estimate average and conditional effects of a treatment variable on an outcome variable, taking into account multiple continuous and categorical covariates. EffectLiteR also includes a graphical user interface based on the shiny package (RStudio and Inc., 2014). The R package is available from CRAN and can be installed as usual provided that a recent version of R is installed:

\begin{verbatim}
## Basic commands to install and run EffectLiteR ##
install.packages("EffectLiteR") ## install package
library(EffectLiteR) ## load package
?effectLite ## get help for main function
effectLiteGUI() ## opens the graphical user interface
\end{verbatim}

Users who wish to try out EffectLiteR without installing any software can use the online version of the graphical user interface at https://amayer.shinyapps.io/elrshiny/. The development version of the package along with descriptions of how to install this version is available from https://github.com/amayer2010.
A.2 Screenshot of the Graphical User Interface

A.3 EffectLiteR Input Syntax

```r
## Using EffectLiteR via the command line ##
library(EffectLiteR)

## measurement model
mm <- 'eta =~ 1*y12 + 1*y22 + 1*y32
xi =~ 1*y11 + 1*y21 + 1*y31
y11 + y21 + y31 + y12 + y22 + y32 ~ 0*1

## call effectLite()
## data generation for this example is described in Supplement C.1
m1 <- effectLite(y="eta", x="x", k="k", z="xi", control="0",
                measurement=mm, data=d, fixed.cell=FALSE,
                missing="fiml", syntax.only=FALSE)
print(m1)
```
### Regression Model

\[
E(Y|X,K,Z) = g0(K,Z) + g1(K,Z)I_{X=1} + g2(K,Z)I_{X=2}
\]
\[
g0(K,Z) = g000 + g001Z1 + g010I_{K=1} + g011I_{K=1}Z1
\]
\[
g1(K,Z) = g100 + g101Z1 + g110I_{K=1} + g111I_{K=1}Z1
\]
\[
g2(K,Z) = g200 + g201Z1 + g210I_{K=1} + g211I_{K=1}Z1
\]

### Main Hypotheses

| Wald Chi-Square | df | p-value |
|-----------------|----|---------|
| No average effects | 266 | 2 | 0 |
| No covariate effects in control group | 835 | 3 | 0 |
| No treatment*covariate interaction | 214 | 6 | 0 |
| No treatment effects | 557 | 8 | 0 |

### Average Effects

| Estimate | SE | Est./SE | p-value | Effect Size |
|----------|----|---------|---------|-------------|
| E[g1(K,Z)] | 0.729 | 0.0925 | 7.88 | 3.11e-15 | 0.410 |
| E[g2(K,Z)] | 1.327 | 0.0814 | 16.30 | 0.00e+00 | 0.746 |

### Effects given K=k

| Estimate | SE | Est./SE | p-value | Effect Size |
|----------|----|---------|---------|-------------|
| E[g1(K,Z)|K=0] | 0.476 | 0.144 | 3.30 | 9.57e-04 | 0.268 |
| E[g2(K,Z)|K=0] | 0.694 | 0.116 | 6.01 | 1.91e-09 | 0.390 |
| E[g1(K,Z)|K=1] | 0.953 | 0.118 | 8.07 | 6.66e-16 | 0.536 |
| E[g2(K,Z)|K=1] | 1.885 | 0.109 | 17.35 | 0.00e+00 | 1.061 |
Supplement B: Complete Computations for the Motivating Example

B.1 Multigroup SEM with Stochastic Group Sizes for the Simulated Data Example

Group-invariant $\tau$-equivalent measurement model

$$
\begin{bmatrix}
Y_{12} \\
Y_{22} \\
Y_{32} \\
Y_{11} \\
Y_{21} \\
Y_{31}
\end{bmatrix} =
\begin{bmatrix}
0 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\eta \\
\xi
\end{bmatrix} +
\begin{bmatrix}
\epsilon_{12} \\
\epsilon_{22} \\
\epsilon_{32} \\
\epsilon_{11} \\
\epsilon_{21} \\
\epsilon_{31}
\end{bmatrix}
$$

Group-specific structural model in each of the six groups (3 x 2)

$$
\begin{bmatrix}
\eta \\
\xi
\end{bmatrix} =
\begin{bmatrix}
\alpha_{xk0} \\
\mu_{xk1}
\end{bmatrix} +
\begin{bmatrix}
0 & \alpha_{xk1} \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\eta \\
\xi
\end{bmatrix} +
\begin{bmatrix}
\eta_0 \\
\xi_1
\end{bmatrix}
$$

Group sizes for each of the six groups (3 x 2)

$$
\log(n_{sk}) = \kappa_{sk}
$$

B.2 Effect Functions

Based on the parameters $\alpha_{skz}$, $\mu_{sk1}$, and $\kappa_{sk}$ of the multigroup structural equation model, we can identify average and conditional effects. The parameterization of the regression $E(\eta | X, K, \xi)$ is given by:

$$
E(\eta | X, K, \xi) = g_0(K, \xi) + g_1(K, \xi) \cdot I_{X=1} + g_2(K, \xi) \cdot I_{X=2}
$$

$$
g_0(K, \xi) = \gamma_{000} + \gamma_{010} \cdot K + \gamma_{001} \cdot \xi + \gamma_{011} \cdot K \cdot \xi
$$

$$
g_1(K, \xi) = \gamma_{100} + \gamma_{110} \cdot K + \gamma_{101} \cdot \xi + \gamma_{111} \cdot K \cdot \xi
$$

$$
g_2(K, \xi) = \gamma_{200} + \gamma_{210} \cdot K + \gamma_{201} \cdot \xi + \gamma_{211} \cdot K \cdot \xi
$$

First, we compute the regression coefficients of the effect functions $g_1(K, \xi)$ and $g_2(K, \xi)$ following Equations 17 and 18:

$$
\gamma_{100} = \alpha_{100} - \alpha_{000} \\
\gamma_{101} = \alpha_{101} - \alpha_{001} \\
\gamma_{110} = (\alpha_{110} - \alpha_{100}) - (\alpha_{010} - \alpha_{000}) \\
\gamma_{111} = (\alpha_{111} - \alpha_{101}) - (\alpha_{011} - \alpha_{001}) \\
\gamma_{200} = \alpha_{200} - \alpha_{000} \\
\gamma_{201} = \alpha_{201} - \alpha_{001} \\
\gamma_{210} = (\alpha_{210} - \alpha_{200}) - (\alpha_{010} - \alpha_{000}) \\
\gamma_{211} = (\alpha_{211} - \alpha_{201}) - (\alpha_{011} - \alpha_{001})
$$
B.3 Average and Conditional Effects

**Average Effects** \(AE_{E0}\): In our example, there are two average effects, \(AE_{10}\), the average effect of treatment \(X = 1\) vs. \(X = 0\), and \(AE_{20}\), the average effect of treatment \(X = 2\) vs. \(X = 0\). Following Table 3, these are identified as:

\[
E[g_1(K, \xi)] = \gamma_{100} + \gamma_{110} \cdot E(K) + \gamma_{101} \cdot E(\xi) + \gamma_{111} \cdot E(K \cdot \xi)
\]

\[
E[g_2(K, \xi)] = \gamma_{200} + \gamma_{210} \cdot E(K) + \gamma_{201} \cdot E(\xi) + \gamma_{211} \cdot E(K \cdot \xi)
\]

In order to compute the average effects based on model parameters, we need to compute the unconditional expectations of covariates and product terms (cf. Table 2):

\[
P(X = x, K = k) = \exp(\kappa_{xk}) / \left[ \exp(\kappa_{00}) + \exp(\kappa_{01}) + \exp(\kappa_{10}) + \exp(\kappa_{11}) + \exp(\kappa_{20}) + \exp(\kappa_{21}) \right]
\]

\[
E(\xi) = \mu_{001} \cdot P(X = 0, K = 0) + \mu_{011} \cdot P(X = 0, K = 1) + \mu_{101} \cdot P(X = 1, K = 0) + \\
\mu_{111} \cdot P(X = 1, K = 1) + \mu_{201} \cdot P(X = 2, K = 0) + \mu_{211} \cdot P(X = 2, K = 1)
\]

\[
P(K = 0) = P(X = 0, K = 0) + P(X = 1, K = 0) + P(X = 2, K = 0)
\]

\[
P(K = 1) = P(X = 0, K = 1) + P(X = 1, K = 1) + P(X = 2, K = 1)
\]

\[
E(\xi | K = 1) = \mu_{011} \cdot P(X = 0, K = 1) + \mu_{111} \cdot P(X = 1, K = 1) + \mu_{211} \cdot P(X = 2, K = 1)
\]

Then, the average effects are identified as functions of the regression coefficients of the effect functions and the unconditional expectations of covariates and product terms.

**Conditional Effects** \(CE_{E0,K=k}\): In our example, there are four \((K = k)\)-conditional effects, the two conditional effects of treatment \(X = 1\) vs. \(X = 0\) for males \(CE_{10,K=0}\) and females \(CE_{10,K=1}\), and the two conditional effects of treatment \(X = 2\) vs. \(X = 0\) for males \(CE_{20,K=0}\) and females \(CE_{20,K=1}\). Following Table 3, these are identified as:

\[
E[g_1(K, \xi) | K = 0] = \gamma_{100} + \gamma_{110} \cdot E(\xi | K = 0)
\]

\[
E[g_1(K, \xi) | K = 1] = \gamma_{100} + \gamma_{110} + \gamma_{101} \cdot E(\xi | K = 1) + \gamma_{111} \cdot E(\xi | K = 1)
\]

\[
E[g_2(K, \xi) | K = 0] = \gamma_{200} + \gamma_{201} \cdot E(\xi | K = 0)
\]

\[
E[g_2(K, \xi) | K = 1] = \gamma_{200} + \gamma_{210} + \gamma_{201} \cdot E(\xi | K = 1) + \gamma_{211} \cdot E(\xi | K = 1)
\]

For the identification of the four \((K = k)\)-conditional effects, we need the regression coefficients of the effect functions (see above) and the \((K = k)\)-conditional expectations of \(\xi\):

\[
P(X = x \mid K = k) = P(X = x, K = k) / P(K = k)
\]

\[
E(\xi \mid K = 0) = \mu_{001} \cdot P(X = 0 \mid K = 0) + \mu_{011} \cdot P(X = 1 \mid K = 0) + \mu_{201} \cdot P(X = 2 \mid K = 0)
\]

\[
E(\xi \mid K = 1) = \mu_{011} \cdot P(X = 0 \mid K = 1) + \mu_{111} \cdot P(X = 1 \mid K = 1) + \mu_{211} \cdot P(X = 2 \mid K = 1)
\]

where \(P(K = k)\) and \(P(X = x, K = k)\) have already been computed in the section about the identification of average effects.
Conditional Effects $CE_{t0;X=x}$: There are six different kinds of $(X = x)$-conditional effects in our example, i.e., the three conditional effects of treatment $X = 1$ vs. $X = 0$ given values $x$ of $X$, and the three conditional effects of treatment $X = 2$ vs. $X = 0$ given values $x$ of $X$:

$$E[g_1(K, \xi) | X = 0] = \gamma_{100} + \gamma_{110} \cdot E(I_{K=1} | X = 0) + \gamma_{101} \cdot E(\xi | X = 0) + \gamma_{111} \cdot E(\xi I_{K=1} | X = 0)$$
$$E[g_1(K, \xi) | X = 1] = \gamma_{100} + \gamma_{110} \cdot E(I_{K=1} | X = 1) + \gamma_{101} \cdot E(\xi | X = 1) + \gamma_{111} \cdot E(\xi I_{K=1} | X = 1)$$
$$E[g_1(K, \xi) | X = 2] = \gamma_{100} + \gamma_{110} \cdot E(I_{K=1} | X = 2) + \gamma_{101} \cdot E(\xi | X = 2) + \gamma_{111} \cdot E(\xi I_{K=1} | X = 2)$$
$$E[g_2(K, \xi) | X = 0] = \gamma_{200} + \gamma_{210} \cdot E(I_{K=1} | X = 0) + \gamma_{201} \cdot E(\xi | X = 0) + \gamma_{211} \cdot E(\xi I_{K=1} | X = 0)$$
$$E[g_2(K, \xi) | X = 1] = \gamma_{200} + \gamma_{210} \cdot E(I_{K=1} | X = 1) + \gamma_{201} \cdot E(\xi | X = 1) + \gamma_{211} \cdot E(\xi I_{K=1} | X = 1)$$
$$E[g_2(K, \xi) | X = 2] = \gamma_{200} + \gamma_{210} \cdot E(I_{K=1} | X = 2) + \gamma_{201} \cdot E(\xi | X = 2) + \gamma_{211} \cdot E(\xi I_{K=1} | X = 2)$$

In order to compute the $(X = x)$-conditional effects, we also need the $(X = x)$-conditional expectations of covariates and product terms:

$$E(I_{K=k} | X = x) = P(K = k | X = x) = P(X = x, K = k) / P(X = x)$$
$$E(\xi | X = 0) = \mu_{001} \cdot P(K = 0 | X = 0) + \mu_{011} \cdot P(K = 1 | X = 0)$$
$$E(\xi | X = 1) = \mu_{101} \cdot P(K = 0 | X = 1) + \mu_{111} \cdot P(K = 1 | X = 1)$$
$$E(\xi | X = 2) = \mu_{201} \cdot P(K = 0 | X = 2) + \mu_{211} \cdot P(K = 1 | X = 2)$$
$$E(\xi I_{K=1} | X = 0) = \mu_{011} \cdot P(K = 1 | X = 0)$$
$$E(\xi I_{K=1} | X = 1) = \mu_{111} \cdot P(K = 1 | X = 1)$$
$$E(\xi I_{K=1} | X = 2) = \mu_{211} \cdot P(K = 1 | X = 2)$$

Conditional Effects $CE_{t0;X=x,K=k}$: Finally, we consider $(X = x, K = k)$-conditional effects. In our simulated example, there are twelve different such effects, i.e.,

$$E[g_1(K, \xi) | X = 0, K = 0], E[g_1(K, \xi) | X = 0, K = 1], E[g_1(K, \xi) | X = 1, K = 0], E[g_1(K, \xi) | X = 1, K = 1], E[g_1(K, \xi) | X = 2, K = 0], E[g_1(K, \xi) | X = 2, K = 1], E[g_2(K, \xi) | X = 0, K = 0], E[g_2(K, \xi) | X = 0, K = 1], E[g_2(K, \xi) | X = 1, K = 0], E[g_2(K, \xi) | X = 1, K = 1], E[g_2(K, \xi) | X = 2, K = 0], E[g_2(K, \xi) | X = 2, K = 1].$$

These effects can directly be computed based on regression coefficients of the effect functions and the model parameters for $(X = x, K = k)$-conditional expectations of the continuous latent covariate:

$$E[g_1(K, \xi) | X = x, K = 0] = \gamma_{100} + \gamma_{101} \cdot \mu_{x01}$$
$$E[g_1(K, \xi) | X = x, K = 1] = \gamma_{100} + \gamma_{110} + \gamma_{101} \cdot \mu_{x11} + \gamma_{111} \cdot \mu_{x11}$$
$$E[g_2(K, \xi) | X = x, K = 0] = \gamma_{200} + \gamma_{201} \cdot \mu_{x01}$$
$$E[g_2(K, \xi) | X = x, K = 1] = \gamma_{200} + \gamma_{210} + \gamma_{201} \cdot \mu_{x11} + \gamma_{211} \cdot \mu_{x11}$$
Supplement C: Data Generation

C.1 R Syntax to Generate Data

```r
set.seed(33882)
N <- 1000

design <- expand.grid(k=0:1, x=0:2)
prob <- c(0.1, 0.2, 0.1, 0.1, 0.3, 0.2)
ind <- sample(1:6, size=N, replace=TRUE, prob=prob)

d <- design[ind,
    Ix1 <- as.numeric(x==1)
    Ix2 <- as.numeric(x==2)
    xi <- 1 + 0.5*k + 0.3*Ix1 + 0.4*Ix2 + rnorm(N,0,2)
    y11 <- xi + rnorm(N,0,sqrt(0.7))
    y21 <- xi + rnorm(N,0,sqrt(0.7))
    y31 <- xi + rnorm(N,0,sqrt(0.7))
    eta <- (0.1+0.8*xi) + (0.3+0.5*k+0.2*xi)*Ix1 +
               (0.3+0.5*k+0.2*xi+0.4*k*xi)*Ix2 + rnorm(N,0,0.7)
    y12 <- eta + rnorm(N,0,sqrt(0.8))
    y22 <- eta + rnorm(N,0,sqrt(0.8))
    y32 <- eta + rnorm(N,0,sqrt(0.8))
    x <- as.factor(x)
}

d <- subset(d, select = -c(eta, xi))
```
C.2 True Population Parameters for Data Generation

Cell and marginal probabilities:

|       | $K = 0$ | $K = 1$ | All  |
|-------|---------|---------|------|
| $X = 0$ | 0.1     | 0.2     | 0.3  |
| $X = 1$ | 0.1     | 0.1     | 0.2  |
| $X = 2$ | 0.3     | 0.2     | 0.5  |
| All    | 0.5     | 0.5     | 1.0  |

Manifest variables:

\[
\begin{align*}
Y_{11} &= \xi + \varepsilon_{11} \\
Y_{12} &= \eta + \varepsilon_{12} \\
Y_{21} &= \xi + \varepsilon_{21} \\
Y_{22} &= \eta + \varepsilon_{22} \\
Y_{31} &= \xi + \varepsilon_{31} \\
Y_{32} &= \eta + \varepsilon_{32}
\end{align*}
\]

where $\text{Var}(\varepsilon_{11}) = \text{Var}(\varepsilon_{21}) = \text{Var}(\varepsilon_{31}) = 0.7$ and $\text{Var}(\varepsilon_{12}) = \text{Var}(\varepsilon_{22}) = \text{Var}(\varepsilon_{32}) = 0.8$.

Latent variables:

\[
\begin{align*}
\xi &= 1 + 0.5 \cdot K + 0.3 \cdot I_{X=1} + 0.4 \cdot I_{X=2} + \zeta_1 \\
\eta &= (0.1 + 0.8 \cdot \xi) + (0.3 + 0.5 \cdot K + 0.2 \cdot \xi) \cdot I_{X=1} \\
&\quad + (0.3 + 0.5 \cdot K + 0.2 \cdot \xi + 0.4 \cdot K \cdot \xi) \cdot I_{X=2} + \zeta_0
\end{align*}
\]

where $\text{Var}(\zeta_1) = 4$ and $\text{Var}(\zeta_0) = 0.49$. 
C.3 True Average and Conditional Effects

True average and conditional effects of conventional therapy ($X = 1$ vs. $X = 0$) and of innovative therapy ($X = 2$ vs. $X = 0$) compared to the control group in the simulated data example.

| Effect                           | True value | Effect                           | True value |
|----------------------------------|------------|----------------------------------|------------|
| **True Average and Conditional Effects of Conventional Therapy** |            | **Conditional Effects** |            |
| **Average Effect** $AE_{10}$    |            | $CE_{10,X=x}$                    |            |
| $AE_{10}$                        | 0.852      | $CE_{10,X=0}$                   | 0.9        |
| **Effect Function $g_1(K,\xi)$**|            | $CE_{10,X=1}$                   | 0.86       |
| $\gamma_{100}$                  | 0.3        | $CE_{10,X=2}$                   | 0.82       |
| $\gamma_{101}$                  | 0.2        |                                |            |
| $\gamma_{110}$                  | 0.5        |                                |            |
| $\gamma_{111}$                  | 0.0        |                                |            |
| **Conditional Effects** $CE_{10;K=k}$ |           | $CE_{10,X=x;K=k}$              |            |
| $CE_{10;K=0}$                    | 0.56       | $CE_{10;X=0;K=0}$               | 0.5        |
| $CE_{10;K=1}$                    | 1.144      | $CE_{10;X=1;K=0}$               | 0.56       |
| **True Average and Conditional Effects of Innovative Therapy** |            |                                |            |
| **Average Effect** $AE_{20}$    |            | **Conditional Effects** |            |
| $AE_{20}$                        | 1.196      | $CE_{20,X=x}$                  | 1.3        |
| **Effect Function $g_2(K,\xi)$**|            | $CE_{20,X=1}$                  | 1.22       |
| $\gamma_{200}$                  | 0.3        | $CE_{20,X=2}$                  | 1.24       |
| $\gamma_{201}$                  | 0.2        |                                |            |
| $\gamma_{210}$                  | 0.5        |                                |            |
| $\gamma_{211}$                  | 0.4        |                                |            |
| **Conditional Effects** $CE_{20;K=k}$ |           | $CE_{20,X=x;K=k}$              |            |
| $CE_{20;K=0}$                    | 0.56       | $CE_{20;X=0;K=0}$               | 0.5        |
| $CE_{20;K=1}$                    | 1.832      | $CE_{20;X=1;K=0}$               | 0.56       |

The true effect functions are given by $g_t(K,\xi) = \gamma_{t00} + \gamma_{t10} \cdot K + \gamma_{t01} \cdot \xi + \gamma_{t11} \cdot K \cdot \xi$ for $t = 1, 2$. The reported effects are true average effects $AE_{t0} = E[g_t(K,\xi)]$, conditional effects given gender $CE_{t0;K=k} = E[g_t(K,\xi) | K = k]$, conditional effects given a treatment condition $CE_{t0;X=x} = E[g_t(K,\xi) | X = x]$ and conditional effects given gender and a treatment condition $CE_{t0;X=x;K=k} = E[g_t(K,\xi) | X = x, K = k]$. 
Supplement D: Monte Carlo Study for Sample Size Considerations

```r
library(EffectLiteR)
library(simsem)

set.seed(623523)
nrep <- 5000
dataList <- list(length=nrep)

for(i in 1:nrep){

N <- 150

design <- expand.grid(k=0:1, x=0:2)
prob <- c(0.1, 0.2, 0.1, 0.1, 0.3, 0.2)
ind <- sample(1:6, size=N, replace=TRUE, prob=prob)
dpop <- design[ind ,]

dpop <- within(dpop ,{
    Ix1 <- as.numeric(x==1)
    Ix2 <- as.numeric(x==2)
    eta1 <- 1 + 0.5*k + 0.3*Ix1 + 0.4*Ix2 + rnorm(N,0,2)
    y11 <- eta1 + rnorm(N,0,sqrt(0.7))
    y21 <- eta1 + rnorm(N,0,sqrt(0.7))
    y31 <- eta1 + rnorm(N,0,sqrt(0.7))
    eta2 <- (0.1+0.8*eta1) + (0.3+0.5*k+0.2*eta1)*Ix1 +
    (0.3+0.5*k+0.2*eta1+0.4*k*eta1)*Ix2 + rnorm(N,0,0.7)
    y12 <- eta2 + rnorm(N,0,sqrt(0.8))
    y22 <- eta2 + rnorm(N,0,sqrt(0.8))
    y32 <- eta2 + rnorm(N,0,sqrt(0.8))
    x <- as.factor(x)
    cell <- as.factor(paste0(x,k))
})

dpop <- subset(dpop, select=-c(eta2,eta1,Ix1,Ix2))
dataList[[i]] <- dpop
}

## m1 is the fitted EffectLiteR model from Supplement A.3
analysisModel <- m1@lavaansyntax@model

Output <- sim(model=analysisModel, rawData=dataList, lavaanfun="sem", group="cell",
group.label=c("00","01","10","11","20","21"), fixed.x=FALSE,
group.w.free=TRUE)

summary(Output)
```
References

R Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from http://www.R-project.org/

RStudio and Inc. (2014). shiny: Web application framework for R [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=shiny (R package version 0.10.1)