# Centering lower-level interactions in multilevel models 

Haeike Josephy<br>Tom Loeys

International Meetings of the Psychometric Society 2017

## Longitudinal diary study on sexual behavior in Flanders

- Info on 66 heterosexual couples
- Here, we only focus on men's data
- Daily measures during 3 weeks:
- Daily morning reports on sexual and intimate behavior: amount of intimate acts (kissing, cuddling and caressing) measured on a 7 -point scale
- Daily evening reports on positive relationship feelings: average score of 9 items (happy, satisfied, understood, ...) measured on a 7 -point scale


## Longitudinal diary study on sexual behavior in Flanders

- Info on 66 heterosexual couples
- Here, we only focus on men's data
- Daily measures during 3 weeks:
- Daily morning reports on sexual and intimate behavior: amount of intimate acts (kissing, cuddling and caressing) measured on a 7-point scale
- Daily evening reports on positive relationship feelings: average score of 9 items (happy, satisfied, understood, ...) measured on a 7-point scale

Question: What is the contribution of intimacy to next-day positive relationship feelings?

## Standard multilevel modeling

In our example:

- $X_{i j}$ : daily measurement of intimacy of individual $j$ at time $i$
- $Y_{i j}$ : next day's positive relational feelings

Standard analysis by a multilevel model with random intercept $b_{j}$ :

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, b_{j}\right)=\gamma_{0}+\gamma X_{i j}+b_{j} \tag{1}
\end{equation*}
$$



## Standard multilevel modeling

In our example:

- $X_{i j}$ : daily measurement of intimacy of individual $j$ at time $i$
- $Y_{i j}$ : next day's positive relational feelings

Standard analysis by a multilevel model with random intercept $b_{j}$ :

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, b_{j}\right)=\gamma_{0}+\gamma X_{i j}+b_{j} \tag{1}
\end{equation*}
$$



Unfortunately, there may be upper level endogeneity!

## Standard multilevel modeling

In our example:

- $X_{i j}$ : daily measurement of intimacy of individual $j$ at time $i$
- $Y_{i j}$ : next day's positive relational feelings

Standard analysis by a multilevel model with random intercept $b_{j}$ :

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, b_{j}\right)=\gamma_{0}+\gamma X_{i j}+b_{j} \tag{1}
\end{equation*}
$$


(1) assumes that $b_{j}$ and $X_{i j}$ are independent $\Rightarrow$ biased estimator for $\gamma$ under upper-level endogeneity!

## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{w} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- $\gamma_{w}$ captures the within-subject effect
- $\gamma_{B}$ captures the between-subject effect


## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{W} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- As $X_{i j}-\bar{X}_{j}$ removes all upper level effects, it no longer depends on $b_{j}$ in case of upper level endogeneity.


## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{w} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- As $X_{i j}-\bar{X}_{j}$ removes all upper level effects, it no longer depends on $b_{j}$ in case of upper level endogeneity.
- The OLS-estimator for $\gamma_{w}$ will converge to (in balanced designs):

$$
\begin{equation*}
\hat{\gamma} w=\frac{\operatorname{cov}\left(Y_{i j}, X_{i j}-\bar{X}_{j}\right)}{\operatorname{var}\left(X_{i j}-\bar{X}_{j}\right)} \rightarrow \gamma \tag{3}
\end{equation*}
$$

## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{w} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- As $X_{i j}-\bar{X}_{j}$ removes all upper level effects, it no longer depends on $b_{j}$ in case of upper level endogeneity.
- The OLS-estimator for $\gamma_{w}$ will converge to (in balanced designs):

$$
\begin{align*}
& \hat{\gamma} w=\frac{\operatorname{cov}\left(Y_{i j}, X_{i j}-\bar{X}_{j}\right)}{\operatorname{var}\left(X_{i j}-\bar{X}_{j}\right)} \rightarrow \gamma  \tag{3}\\
& \Rightarrow \text { No bias! }
\end{align*}
$$

## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{w} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- As $X_{i j}-\bar{X}_{j}$ removes all upper level effects, it no longer depends on $b_{j}$ in case of upper level endogeneity.
- The OLS-estimator for $\gamma_{B}$ will converge to (in balanced designs):

$$
\begin{equation*}
\hat{\gamma}_{B}=\frac{\operatorname{cov}\left(\bar{Y}_{j}, \bar{X}_{j}\right)}{\operatorname{var}\left(\bar{X}_{j}\right)} \rightarrow \gamma+\frac{\operatorname{cov}\left(b_{j}, \bar{X}_{j}\right)}{\operatorname{var}\left(\bar{X}_{j}\right)} \tag{3}
\end{equation*}
$$

## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{w} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- As $X_{i j}-\bar{X}_{j}$ removes all upper level effects, it no longer depends on $b_{j}$ in case of upper level endogeneity.
- The OLS-estimator for $\gamma_{B}$ will converge to (in balanced designs):

$$
\begin{equation*}
\hat{\gamma}_{B}=\frac{\operatorname{cov}\left(\bar{Y}_{j}, \bar{X}_{j}\right)}{\operatorname{var}\left(\bar{X}_{j}\right)} \rightarrow \gamma+\frac{\operatorname{cov}\left(b_{j}, \bar{X}_{j}\right)}{\operatorname{var}\left(\bar{X}_{j}\right)} \tag{3}
\end{equation*}
$$

$\Rightarrow$ Bias under upper level endogeneity!

## Centring of lower level effects

A solution to the upper-level endogeneity problem is to separate withinfrom between-effects:

$$
\begin{align*}
& E\left(Y_{i j} \mid X_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{w} X_{i j}^{c}+\gamma_{B} \bar{X}_{j}+b_{j}  \tag{2}\\
& \quad \text { with } \bar{X}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} X_{i j} \text { and } X_{i j}^{c}=X_{i j}-\bar{X}_{j}
\end{align*}
$$

- As $X_{i j}-\bar{X}_{j}$ removes all upper level effects, it no longer depends on $b_{j}$ in case of upper level endogeneity.
- The OLS-estimator for $\gamma_{B}$ will converge to (in balanced designs):

$$
\begin{equation*}
\hat{\gamma}_{B}=\frac{\operatorname{cov}\left(\bar{Y}_{j}, \bar{X}_{j}\right)}{\operatorname{var}\left(\bar{X}_{j}\right)} \rightarrow \gamma+\frac{\operatorname{cov}\left(b_{j}, \bar{X}_{j}\right)}{\operatorname{var}\left(\bar{X}_{j}\right)} \tag{3}
\end{equation*}
$$

$\Rightarrow$ Bias under upper level endogeneity!
(note that $\bar{X}_{j}$ can also be excluded from (2), as $\bar{X}_{j} \Perp X_{i j}^{c}$ )

## Revisited: Longitudinal study on sexual behavior in Flanders

New question: Does the effect of intimacy on next-day positive relationship feelings differ according to whether or not the participant has masturbated the previous day?

- $X_{i j}$ : daily measurement of intimacy of individual $j$ at time $i$
- $Y_{i j}$ : next day's positive relational feelings
- $Z_{i j}: 1$ when individual $j$ has masturbated on day $i, 0$ if not


## Revisited: Longitudinal study on sexual behavior in Flanders

New question: Does the effect of intimacy on next-day positive relationship feelings differ according to whether or not the participant has masturbated the previous day?

- $X_{i j}$ : daily measurement of intimacy of individual $j$ at time $i$
- $Y_{i j}$ : next day's positive relational feelings
- $Z_{i j}: 1$ when individual $j$ has masturbated on day $i, 0$ if not

Adjusted multilevel model:

$$
\begin{gather*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}+\gamma_{2} Z_{i j}+\gamma_{3} X_{i j} Z_{i j}+b_{j}  \tag{4}\\
\Rightarrow \text { Lower level interaction term! }
\end{gather*}
$$

## Revisited: Longitudinal study on sexual behavior in Flanders

New question: Does the effect of intimacy on next-day positive relationship feelings differ according to whether or not the participant has masturbated the previous day?

- $X_{i j}$ : daily measurement of intimacy of individual $j$ at time $i$
- $Y_{i j}$ : next day's positive relational feelings
- $Z_{i j}: 1$ when individual $j$ has masturbated on day $i, 0$ if not

Adjusted multilevel model:

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}+\gamma_{2} Z_{i j}+\gamma_{3} X_{i j} Z_{i j}+b_{j} \tag{4}
\end{equation*}
$$

- $b_{j}$ may again be correlated with $X, Y$ and/or $Z$ in case of upper level endogeneity
- The 'naive' model may then again yield biased estimators for the $\gamma$ 's


## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

$$
X_{i j} * Z_{i j}
$$

## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

$$
{\overline{x z_{j}}}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}}{\overleftarrow{x_{i j} z_{i j}}}_{x_{i j *} * z_{i j}}^{(X Z)_{i j}^{c}=x_{i j} z_{i j}-\overline{X Z}_{j}}
$$

## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

- Center first, take the product next (C1P2):


## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

$$
{\overline{x z_{j}}}_{j}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}}{\overleftarrow{x_{i j} z_{i j}}}_{x_{i j *} * z_{i j}}^{(X Z)_{i j}^{c}=x_{i j} z_{i j}-\overline{X Z}_{j}}
$$

- Center first, take the product next (C1P2):

$$
x_{i j} \quad z_{i j}
$$

## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

- Center first, take the product next (C1P2):



## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

- Center first, take the product next (C1P2):



## Centring of lower-level interactions

There a two possible centring approaches:

- Product first, center next (P1C2):

- Center first, take the product next (C1P2):


Question: which approach should we take? Do they differ in any way?

## The P1C2-approach

- P1C2 model:

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3}(X Z)_{i j}^{c}+b_{j} \tag{5}
\end{equation*}
$$

with $(X Z)_{i j}^{c}=X_{i j} z_{i j}-\overline{X Z_{j}}$ (and $\overline{X Z_{j}}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} x_{i j} z_{i j}$ )
$\Rightarrow \hat{\gamma}_{1}, \hat{\gamma}_{2}$ and $\hat{\gamma}_{3}$ are unbiased estimators for $\gamma_{1}, \gamma_{2}$ and $\gamma_{3}$

## The P1C2-approach

- P1C2 model:

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3}(X Z)_{i j}^{c}+b_{j} \tag{5}
\end{equation*}
$$

with $(X Z)_{i j}^{c}=X_{i j} z_{i j}-\overline{X Z_{j}}$ (and $\overline{X Z_{j}}=\frac{1}{n_{j}} \sum_{i=1}^{n_{j}} x_{i j} z_{i j}$ )
$\Rightarrow \hat{\gamma}_{1}, \hat{\gamma}_{2}$ and $\hat{\gamma}_{3}$ are unbiased estimators for $\gamma_{1}, \gamma_{2}$ and $\gamma_{3}$

- P1C2+ model:

$$
\begin{aligned}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)= & \gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3}(X Z)_{i j}^{c}+ \\
& \gamma_{4} \bar{X}_{j}+\gamma_{5} \bar{Z}_{j}+\gamma_{6} \overline{X Z_{j}}+b_{j}
\end{aligned}
$$

$\Rightarrow$ in balanced designs, the estimated within-effects $\hat{\gamma}_{1}, \hat{\gamma}_{2}$ and $\hat{\gamma}_{3}$ are identical in both models

## The C1P2-approach

- C1P2 model:

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c}+b_{j} \tag{6}
\end{equation*}
$$

## The C1P2-approach

- C1P2 model:

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c}+b_{j} \tag{6}
\end{equation*}
$$

- C1P2+ model:

$$
\begin{aligned}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)= & \gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c} \\
& +\gamma_{4} \bar{X}_{j}+\gamma_{5} \bar{Z}_{j}+\gamma_{6} \bar{X}_{j} \bar{Z}_{j}+b_{j}
\end{aligned}
$$

## The C1P2-approach

- C1P2 model:

$$
\begin{equation*}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c}+b_{j} \tag{6}
\end{equation*}
$$

- C1P2+ model:

$$
\begin{aligned}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)= & \gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c} \\
& +\gamma_{4} \bar{X}_{j}+\gamma_{5} \bar{Z}_{j}+\gamma_{6} \bar{X}_{j} \bar{Z}_{j}+b_{j}
\end{aligned}
$$

- C1P2++ model:

$$
\begin{aligned}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)=\gamma_{0} & +\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c}+\gamma_{4} \bar{X}_{j}+\gamma_{5} \bar{Z}_{j} \\
& +\gamma_{6} \bar{X}_{j} \bar{Z}_{j}+\gamma_{7} \bar{X}_{j} Z_{i j}^{c}+\gamma_{8} \bar{Z}_{j} X_{i j}^{c}+b_{j}
\end{aligned}
$$

## Example - Results

|  | Intimacy $=X$ |  | Masturbation $=Z$ |  | Interaction $=X Z$ |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Estimate | $p$-value | Estimate | $p$-value | Estimate | $p$-value |
| P1C2 | $0.079(0.015)$ | $<.001$ | $-0.151(0.079)$ | .057 | $.0 .075(0.039)$ | .054 |
| P1C2+ | $0.079(0.015)$ | $<.001$ | $-0.150(0.079)$ | .059 | .$-0.075(0.039)$ | .054 |
| C1P2 | $0.080(0.015)$ | $<.001$ | $-0.163(0.080)$ | .042 | $.0 .102(0.050)$ | .042 |
| C1P2+ | $0.080(0.015)$ | $<.001$ | $-0.160(0.080)$ | .045 | $.0 .098(0.045)$ | .049 |
| C1P2++ | $0.080(0.015)$ | $<.001$ | $-0.167(0.080)$ | .037 | .$-0.096(0.050)$ | .056 |

- Different approaches lead to different estimates
- Different approaches lead to different conclusions (at the $5 \%$ significance level)!


## Simulation Study - Settings



| Simulation | $\alpha_{1}$ | Distribution of $X$ |
| :--- | ---: | ---: |
| Sim 1 | 0.0 | $N(0,1)$ |
| Sim 2 | 0.0 | $B(1,0.5)-0.5$ |
| Sim 3 | -0.2 | $B(1,0.5)-0.5$ |
| Sim 4 | -1.5 | $B(1,0.5)-0.5$ |

## Simulation study - Results

- $X$ and $Z$ are grand-mean centred to facilitate interpretation
- Focus on within-effects only


## Simulation study - Results

- $X$ and $Z$ are grand-mean centred to facilitate interpretation
- Focus on within-effects only
- $\mathrm{P} 1 \mathrm{C} 2=\mathrm{P} 1 \mathrm{C} 2+$ and $\mathrm{C} 1 \mathrm{P} 2=\mathrm{C} 1 \mathrm{P} 2+$


## Simulation study - Results

- $X$ and $Z$ are grand-mean centred to facilitate interpretation
- Focus on within-effects only
- $\mathrm{P} 1 \mathrm{C} 2=\mathrm{P} 1 \mathrm{C} 2+$ and $\mathrm{C} 1 \mathrm{P} 2=\mathrm{C} 1 \mathrm{P} 2+$
- Bias for interaction effect in C1P2=C1P2+ when $Z$ is a mediator:

$$
\begin{equation*}
E\left(\hat{\gamma}_{3}\right)=\gamma_{3} \frac{\operatorname{cov}\left[X_{i j} Z_{i j}, X_{i j}^{c} Z_{i j}^{c}\right]}{\operatorname{var}\left[X_{i j}^{c} Z_{i j}^{c}\right]} \tag{7}
\end{equation*}
$$

## Simulation study - Results

- $X$ and $Z$ are grand-mean centred to facilitate interpretation
- Focus on within-effects only
- $\mathrm{P} 1 \mathrm{C} 2=\mathrm{P} 1 \mathrm{C} 2+$ and $\mathrm{C} 1 \mathrm{P} 2=\mathrm{C} 1 \mathrm{P} 2+$
- Bias for interaction effect in C1P2=C1P2+ when $Z$ is a mediator:

$$
\begin{equation*}
E\left(\hat{\gamma}_{3}\right)=\gamma_{3} \frac{\operatorname{cov}\left[X_{i j} Z_{i j}, X_{i j}^{c} Z_{i j}^{c}\right]}{\operatorname{var}\left[X_{i j}^{c} Z_{i j}^{c}\right]} \tag{7}
\end{equation*}
$$

- No bias for P1C2 = P1C2+ and C1P2++


## Simulation study - Results

- $X$ and $Z$ are grand-mean centred to facilitate interpretation
- Focus on within-effects only
- $\mathrm{P} 1 \mathrm{C} 2=\mathrm{P} 1 \mathrm{C} 2+$ and $\mathrm{C} 1 \mathrm{P} 2=\mathrm{C} 1 \mathrm{P} 2+$
- Bias for interaction effect in C1P2=C1P2+ when $Z$ is a mediator:

$$
\begin{equation*}
E\left(\hat{\gamma}_{3}\right)=\gamma_{3} \frac{\operatorname{cov}\left[X_{i j} Z_{i j}, X_{i j}^{c} Z_{i j}^{c}\right]}{\operatorname{var}\left[X_{i j}^{c} Z_{i j}^{c}\right]} \tag{7}
\end{equation*}
$$

- No bias for P1C2 = P1C2+ and C1P2++
- Precision of interaction effect estimator is about 30\% smaller for P1C2 compared to C1P2++


## Conclusions

1. P1C2 yields more precise estimators of the interaction effect compared to the C1P2-approaches
2. In contrast to C1P2, P1C2 is not affected by misspecification or omission of upper level effects (i.e. upper level endogeneity)

## Conclusions

1. P1C2 yields more precise estimators of the interaction effect compared to the C1P2-approaches
2. In contrast to C1P2, P1C2 is not affected by misspecification or omission of upper level effects (i.e. upper level endogeneity)

Thank you!

## Possible complication when the predictors are NOT centred

- Again consider C1P2 or C1P2+:

$$
\begin{aligned}
& E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)= \gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c}+b_{j} \\
& \qquad \begin{aligned}
E\left(Y_{i j} \mid X_{i j}, Z_{i j}, u_{j}\right)= & \gamma_{0}+\gamma_{1} X_{i j}^{c}+\gamma_{2} Z_{i j}^{c}+\gamma_{3} X_{i j}^{c} Z_{i j}^{c} \\
& +\gamma_{4} \bar{X}_{j}+\gamma_{5} \bar{Z}_{j}+\gamma_{6} \bar{X}_{j} \bar{Z}_{j}+b_{j}
\end{aligned}
\end{aligned}
$$

- Bias in main effects for these approaches:

$$
\begin{aligned}
& E\left[\hat{\gamma}_{1}\right]=\beta_{1}+\beta_{3} E\left(\bar{Z}_{j}\right) \\
& E\left[\hat{\gamma}_{2}\right]=\beta_{2}+\beta_{3} E\left(\bar{X}_{j}\right)
\end{aligned}
$$

## Results - Simulation Study

| Estimator |  | $\gamma_{1}$ |  |  | Power | Estimate ( $s d_{E}$ ) | $s e^{\hat{\gamma}_{2}}$ | Coverage | Power | Estimate ( $s d_{E}$ ) | se $\chi^{13}$ | Coverage | Power |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Estimate ( $s d_{E}$ ) | se | Coverage |  |  |  |  |  |  |  |  |  |
| $\stackrel{-}{\text { E }}$ | P1C2 | 0.101 (0.031) | 0.028 | 0.92 | 0.93 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.101 (0.019) | 0.020 | 0.95 | 1.00 |
|  | C1P2 | 0.101 (0.032) | 0.028 | 0.92 | 0.92 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.101 (0.029) | 0.029 | 0.95 | 0.94 |
|  | C1P2++ | 0.101 (0.031) | 0.028 | 0.93 | 0.93 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.101 (0.029) | 0.029 | 0.95 | 0.94 |
| NE | P1C2 | 0.103 (0.056) | 0.055 | 0.95 | 0.46 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.101 (0.040) | 0.039 | 0.95 | 0.72 |
|  | C1P2 | 0.103 (0.056) | 0.055 | 0.95 | 0.47 | 0.100 (0.027) | 0.028 | 0.96 | 1.00 | -0.099 (0.058) | 0.058 | 0.95 | 0.41 |
|  | C1P2++ | 0.103 (0.056) | 0.056 | 0.95 | 0.46 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | $\underline{-0.099(0.058)}$ | 0.058 | 0.95 | 0.41 |
| ${ }_{i}^{m}$ | P1C2 | 0.103 (0.056) | 0.055 | 0.95 | 0.46 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.101 (0.040) | 0.039 | 0.95 | 0.72 |
|  | C1P2 | 0.103 (0.056) | 0.055 | 0.95 | 0.46 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.099 (0.058) | 0.058 | 0.95 | 0.41 |
|  | C1P2++ | 0.103 (0.056) | 0.055 | 0.95 | 0.46 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | $\underline{-0.099(0.058)}$ | 0.058 | 0.95 | 0.41 |
| $\begin{aligned} & \pm \\ & i= \end{aligned}$ | P1C2 | 0.103 (0.068) | 0.069 | 0.95 | 0.33 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.101 (0.040) | 0.039 | 0.95 | 0.72 |
|  | C1P2 | 0.103 (0.068) | 0.069 | 0.95 | 0.33 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.090 (0.055) | 0.055 | 0.95 | 0.37 |
|  | C1P2++ | 0.103 (0.068) | 0.069 | 0.96 | 0.33 | 0.150 (0.027) | 0.028 | 0.96 | 1.00 | -0.099 (0.055) | 0.056 | 0.96 | 0.41 |

