Centering of interactions in lower-level mediation models

Haeike Josephy

Tom Loeys

2016 Joint Statistical Meetings

Introduction - Lower level Mediation

1. Mediation

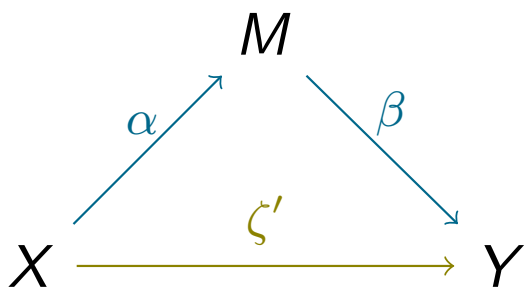
   Unravel causal pathways between exposure X and outcome Y:

   - What is the effect of X on Y?
     = Total Effect

Introduction - Lower level Mediation

1. Mediation

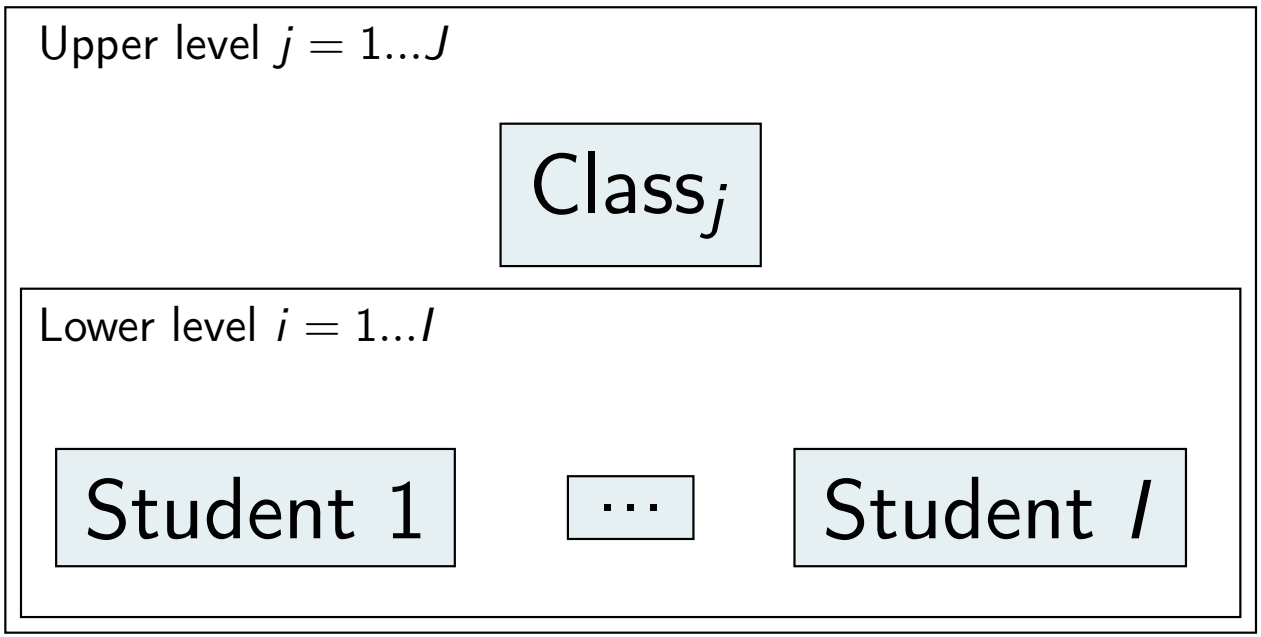
   Unravel causal pathways between exposure X and outcome Y:



   - What part of the effect is mediated by M?
     = Indirect Effect
   - What is the remaining causal effect of X on Y?
     = Direct Effect

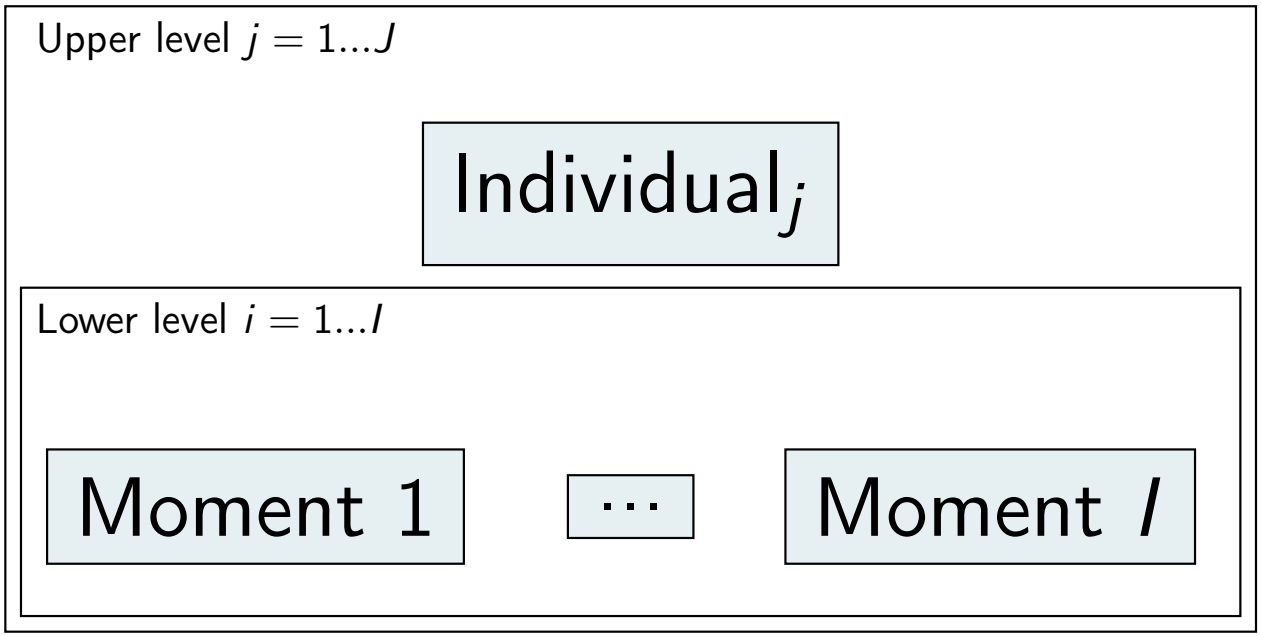
Introduction - Lower level Mediation

2. Multilevel data



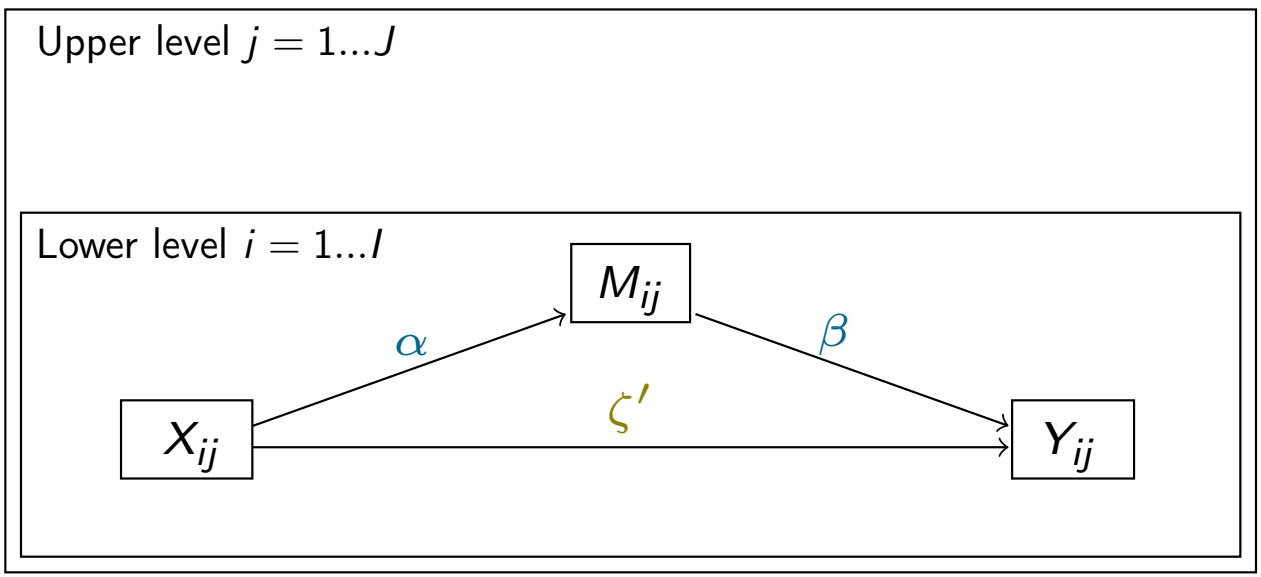
Introduction - Lower level Mediation

2. Multilevel data



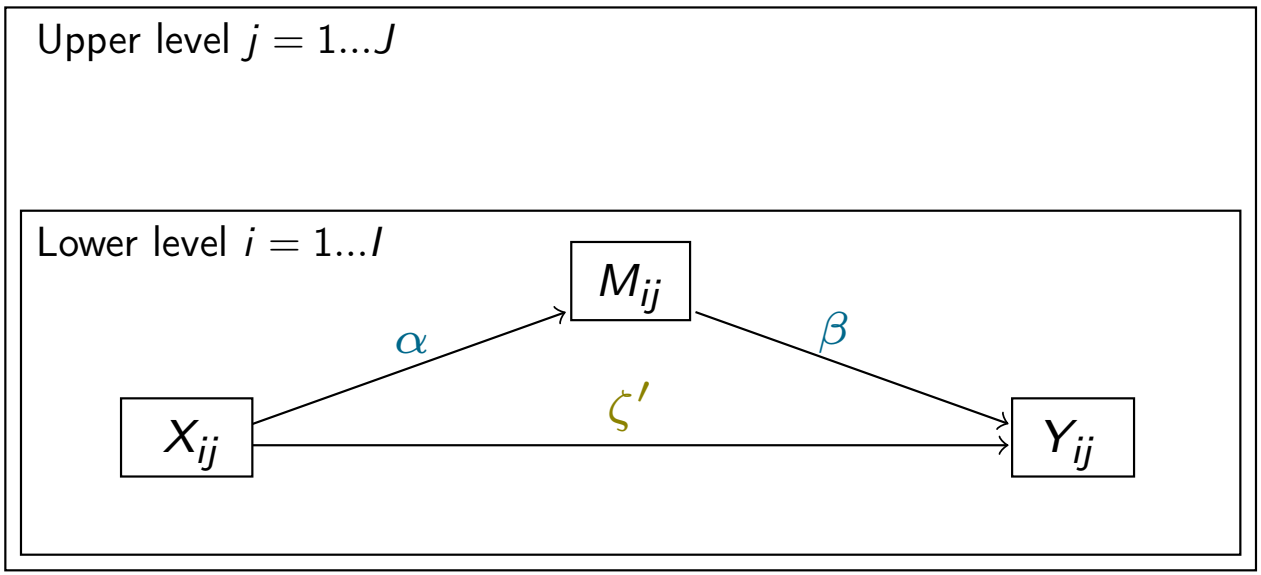
Introduction - Lower level Mediation

3. Lower level Mediation



Introduction - Lower level Mediation

3. Lower level Mediation

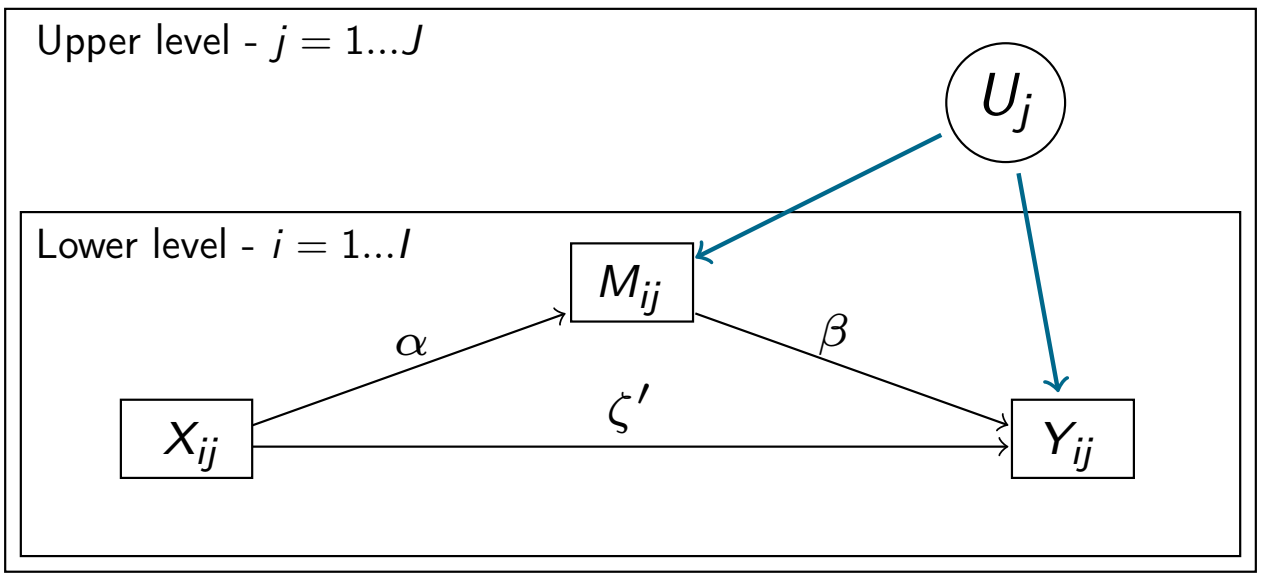


$$
\begin{aligned}
X_{ij} &= \delta_j^x + \epsilon_{ij}^x \\
M_{ij} &= \delta_j^m + \alpha X_{ij} + \epsilon_{ij}^m \\
Y_{ij} &= \delta_j^y + \zeta' X_{ij} + \beta M_{ij} + \epsilon_{ij}^y
\end{aligned} \tag{1}
$$

with all ϵ independently and normally distributed

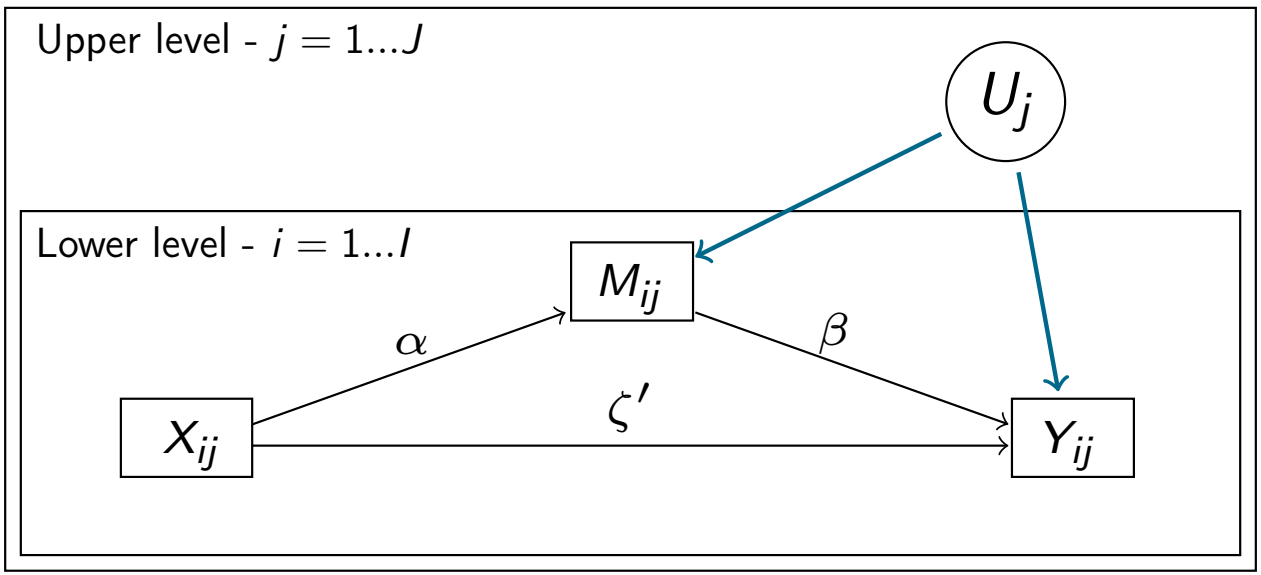
Problem in lower level mediation

Unmeasured upper level confounding of M-Y relation may induce bias in the regression coefficients:



Problem in lower level mediation

Unmeasured upper level confounding of M-Y relation may induce bias in the regression coefficients:



$$
\begin{aligned}
X_{ij} &= \delta_j^x + \epsilon_{ij}^x \\
M_{ij} &= \delta_j^m + \alpha X_{ij} + U_j + \epsilon_{ij}^m \\
Y_{ij} &= \delta_j^y + \zeta' X_{ij} + \beta M_{ij} + U_j + \epsilon_{ij}^y
\end{aligned}
\tag{2}
$$

Dealing with unmeasured upper level M-Y confounding

Possible estimation models:

- Within-cluster centering of X and M in the outcome equation:

$$Y_{ij} = i_{Yj} + c'^{w}(X_{ij} - \bar{X}_j) + c'^{b}\bar{X}_j + b^{w}(M_{ij} - \bar{M}_j) + b^{b}\bar{M}_j + e_{ij} \quad (3)$$

with $\bar{X}_j$ the mean of cluster j

$X_{ij} - \bar{X}_j$ the cluster-mean deviations.

- Joint modeling of M and Y through dummy coding:

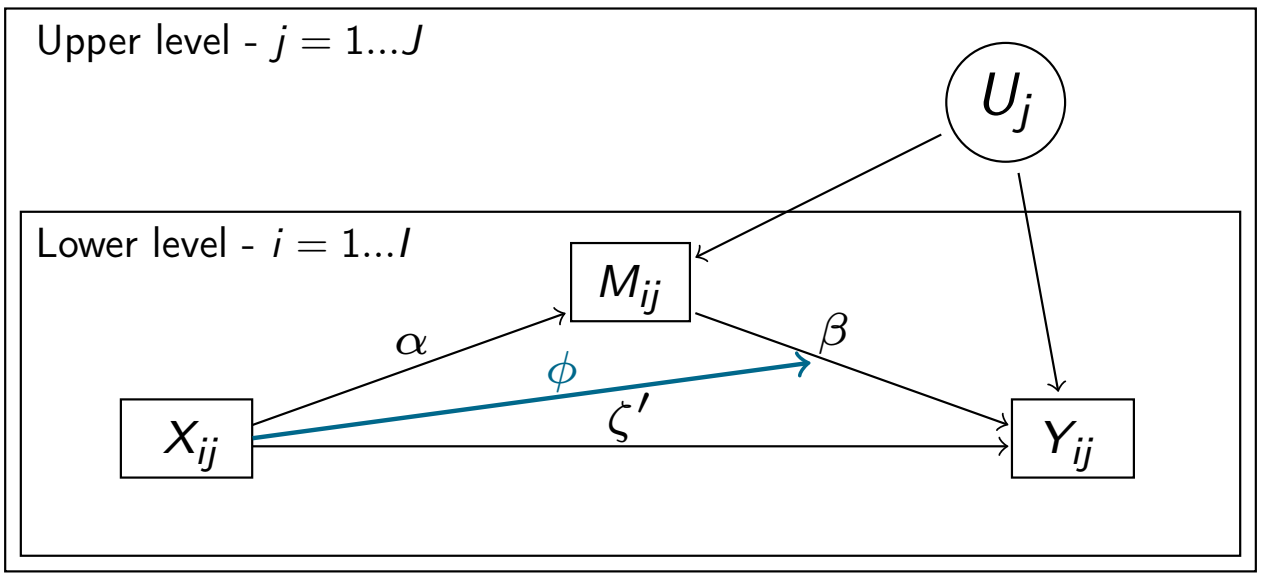
$$Z_{ij} = S_M(i_{Mj} + aX_{ij} + e_{ij}^{M}) + S_Y(i_{Yj} + c'X_{ij} + bM_{ij} + e_{ij}^{Y}) \quad (4)$$

where i_{Mj} and i_{Yj} are allowed to correlate

Lower level Mediation + Moderation

Add one extra complication to out lower level mediation setting:

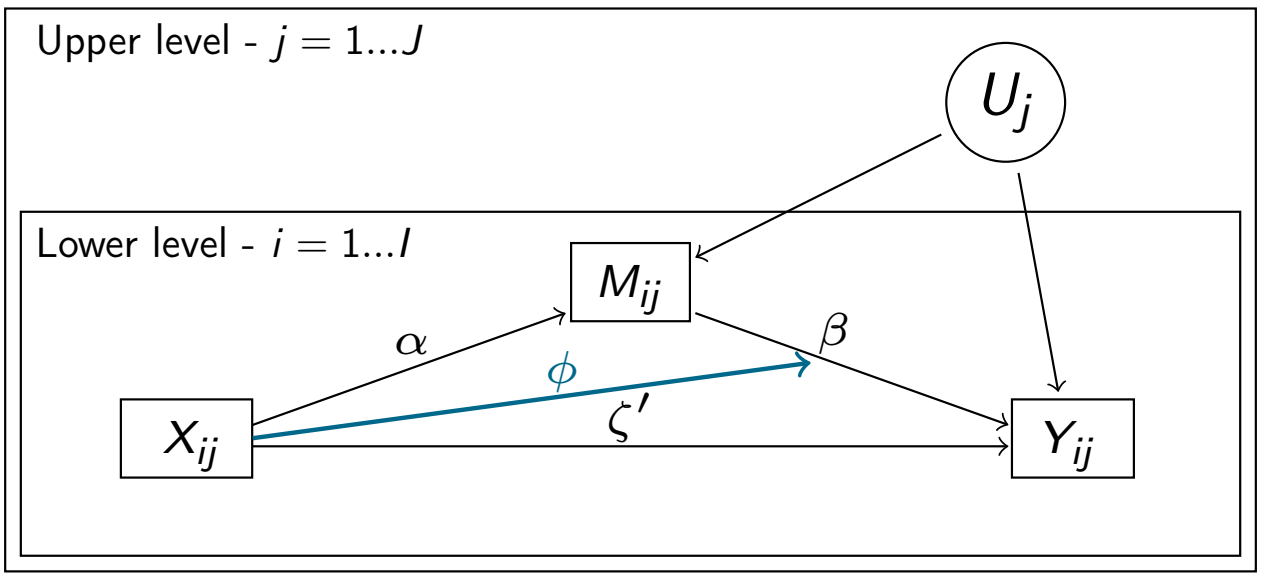
$\Rightarrow$ Lower level Moderated Mediation!



Lower level Mediation + Moderation

Add one extra complication to out lower level mediation setting:

$\Rightarrow$ Lower level Moderated Mediation!



$$
\begin{aligned}
X_{ij} &= \delta_j^x + \epsilon_{ij}^x \\
M_{ij} &= \delta_j^m + \alpha X_{ij} + U_j + \epsilon_{ij}^m \\
Y_{ij} &= \delta_j^y + \zeta' X_{ij} + \beta M_{ij} + \phi X_{ij} M_{ij} + U_j + \epsilon_{ij}^y
\end{aligned}
\tag{5}
$$

Dealing with unmeasured upper level M-Y confounding

THREE ways of centering the interaction within-clusters:

1. Multiply first and center later:

$$\begin{aligned} Y_{ij} =& i_{Yj} + c'^{w}(X_{ij} - \bar{X}_j) + c'^{b}\bar{X}_j + b^{w}(M_{ij} - \bar{M}_j) + b^{b}\bar{M}_j \\ &+ f^{w}(X_{ij}M_{ij} - \overline{XM}_j) + f^{b}\overline{XM}_j + e_{ij} \end{aligned} \tag{6}$$

2. Center first and multiply later:

$$\begin{aligned} Y_{ij} =& i_{Yj} + c'^{w}(X_{ij} - \bar{X}_j) + c'^{b}\bar{X}_j + b^{w}(M_{ij} - \bar{M}_j) + b^{b}\bar{M}_j \\ &+ f^{w}(X_{ij} - \bar{X}_j)(M_{ij} - \bar{M}_j) + f^{b}\bar{X}_j\bar{M}_j + e_{ij} \end{aligned} \tag{7}$$

3. Center first and multiply later, THEN add crosslevel interactions:

$$\begin{aligned} Y_{ij} =& i_{Yj} + c'^{w}(X_{ij} - \bar{X}_j) + c'^{b}\bar{X}_j + b^{w}(M_{ij} - \bar{M}_j) + b^{b}\bar{M}_j \\ &+ f^{w}(X_{ij} - \bar{X}_j)(M_{ij} - \bar{M}_j) + f^{b}\bar{X}_j\bar{M}_j \\ &+ f^{c1}(X_{ij} - \bar{X}_j)\bar{M}_j + f^{c2}\bar{X}_j(M_{ij} - \bar{M}_j) + e_{ij} \end{aligned} \tag{8}$$

Dealing with unmeasured upper level M-Y confounding

ONE joint modeling approach:

4. Joint modeling of M and Y:

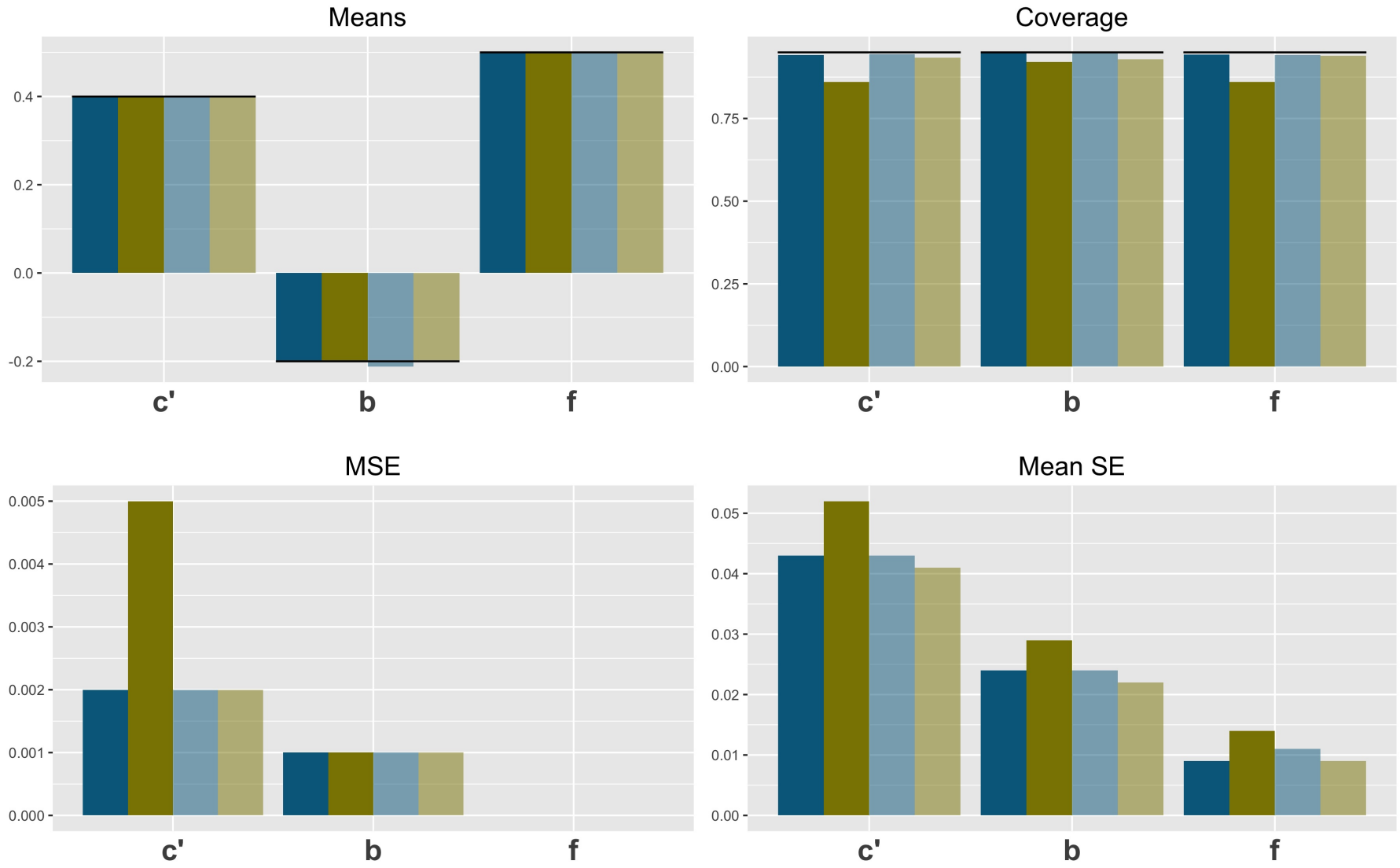
$$Z_{ij} = S_M(i_{Mj} + aX_{ij} + e_{ij}^M) + S_Y(i_{Yj} + c'X_{ij} + bM_{ij} + fX_{ij}M_{ij} + e_{ij}^Y) \quad (6)$$

GOAL: Which of these four methods performs best?

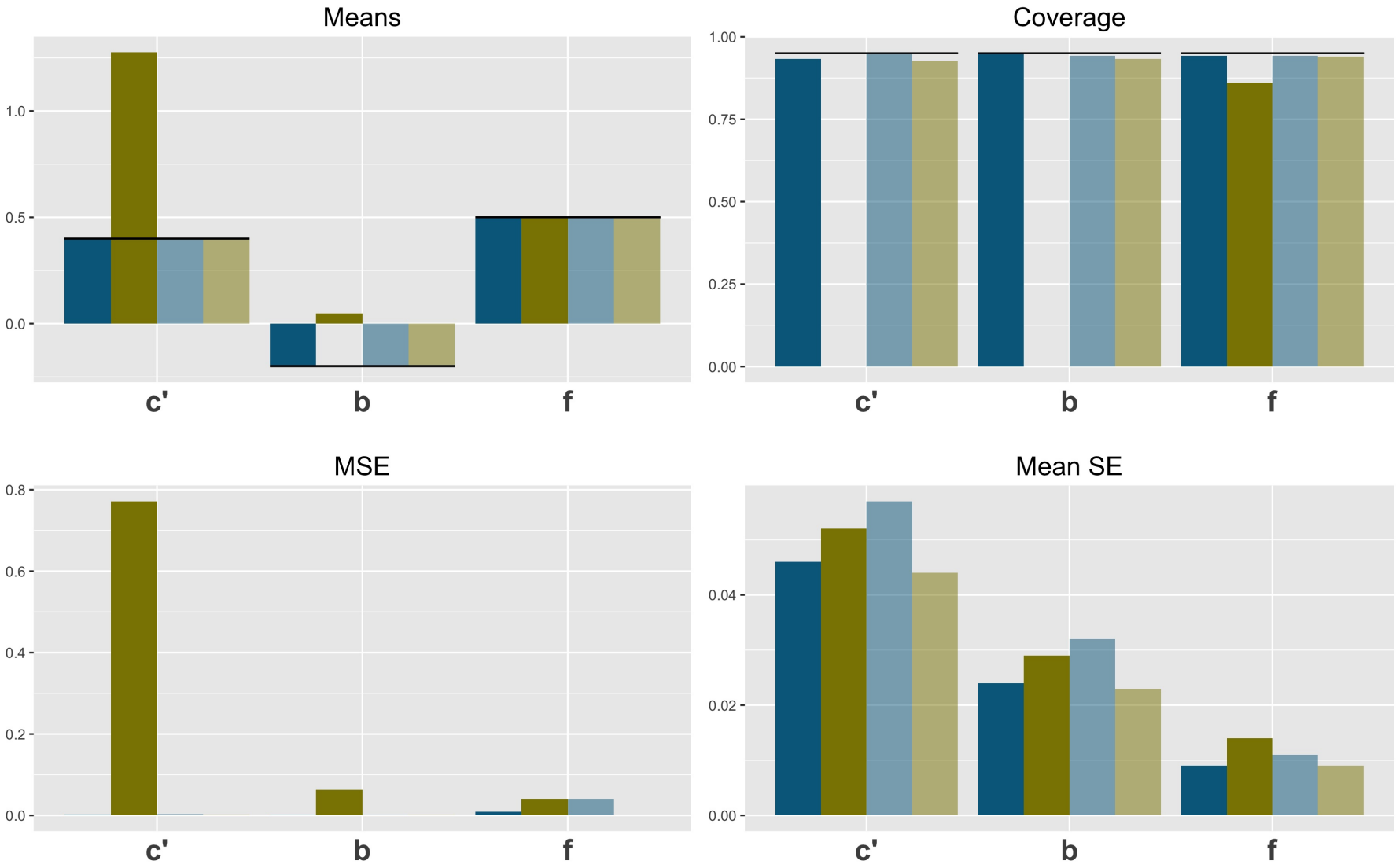
Simulation Studies

- Start with NO unmeasured upper level M-Y confounding $\Rightarrow$ Do all four estimation models estimate the target parameters unbiasedly?
- In a lower level mediation study, focus usually lies in within-effects, so we focus on c'^{w}, b^{w} and f^{w}
- Three different simulation settings of 1000 runs:
  1. No intercepts for normally distributed X and M
  2. Intercepts for normally distributed X and M
  3. Binary distributed X (no intercepts)

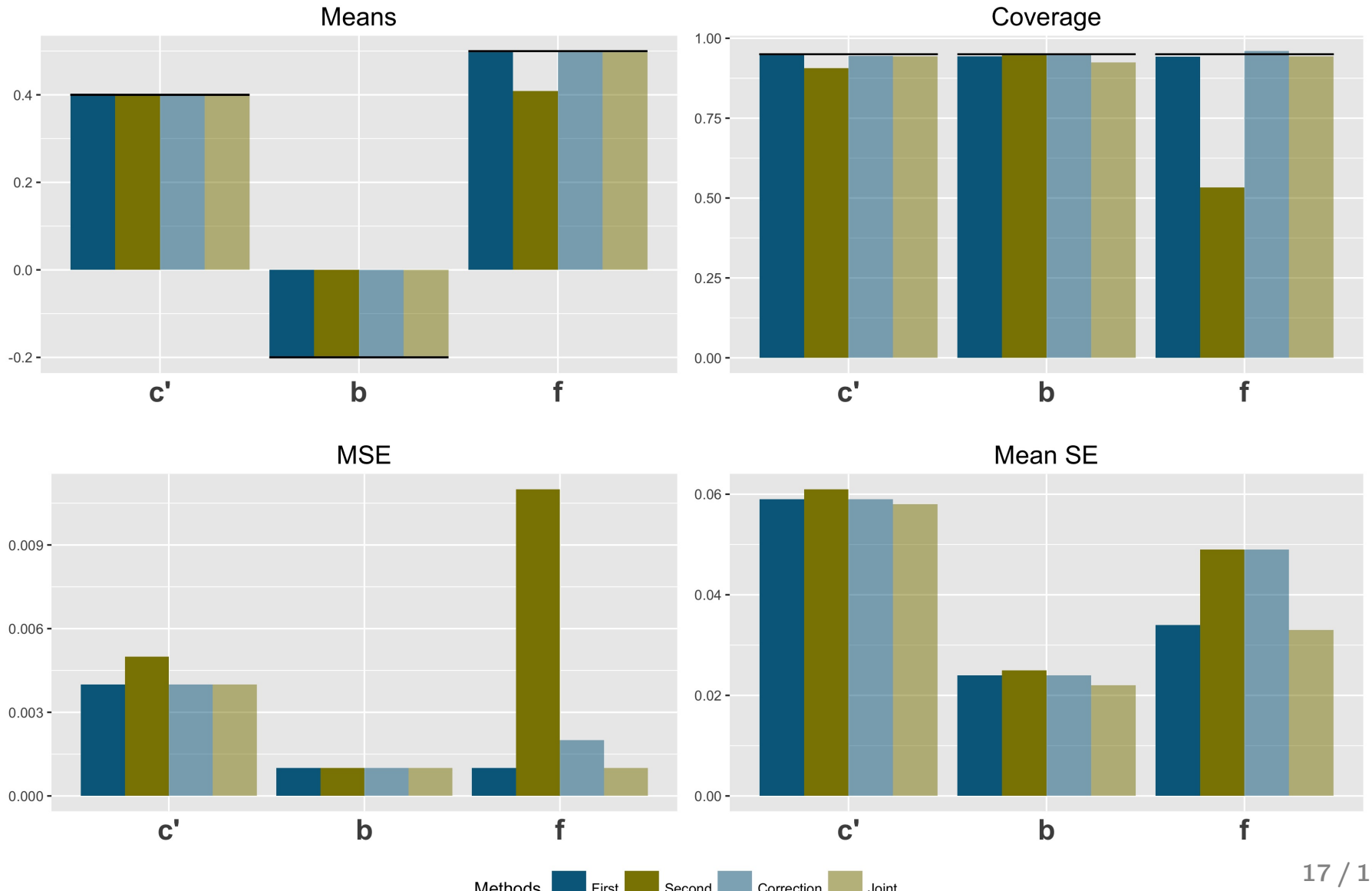
Results: first setting



Results: second setting



Results: third setting



Summary of results

When there is NO unmeasured upper level confounding, we see that:

Sim 1 When X and M are centred:

- No problems

Sim 2 When either X of M show cluster means $\neq 0$:

- The second centering approach does not estimate ζ' and β.
- It estimates $\zeta' + \phi E(\bar{M}_j)$ and $\beta + \phi E(\bar{X}_j)$ instead.
- Inclusion of crosslevel interactions in the estimation model resolves this.

Sim 3 When X is binary:

- The second centering approach biasedly estimates ϕ.
- Inclusion of crosslevel interactions resolves this.

- Overall, the first centering and joint modeling approach provide the smallest SE's (mean and empirical).

Conclusion

- When there is NO unmeasured upper level confounding, we recommend to either center interactions with the first centering approach, of use joint modeling of M and Y:
  - The second centering approach does sometimes estimate the target parameters indirectly, or even with bias.
  - Adding the crosslevel interactions (i.e. the third centering approach) removes these problems, but still yields higher SE's compared to the first centering or the joint modeling approach
- We are currently working on settings WITH unmeasured upper level M-Y confounding, to extend these conclusions.