Appendix 1

In the following, we consider the model presented in Figure A1 as the data generating mechanism. See main text for details on the model and notation.

![Figure A1](image)

Figure A1: Structural relations between an exposure (X), an outcome (Y), and two confounders (Z and U) of the exposure-outcome relation.

Bias due to two unmeasured confounders

Let D be the design matrix, which in case of two unmeasured confounders is simply the vector matrix X. The estimator of the relation between X and Y can then be expressed as:

$$
\hat{\beta}_{yx} = (D'D)^{-1}D'Y.
$$

Since $E[Y] = X\beta_{yx} + Z\beta_{yz} + U\beta_{yu}$, $E[Z] = X\beta_{xz}\frac{Var(Z)}{Var(X)} + \beta_{zu}\beta_{xu}\frac{Var(U)}{Var(X)}$, and $E[U] = X\beta_{yu}\frac{Var(U)}{Var(X)} + \beta_{zu}\beta_{xz}$, it follows that the expected value of the total effect of X on Y based on a model without U and Z is given by:

$$
E[\hat{\beta}_{yx}] = \beta_{yx} + \beta_{yz}(\beta_{xz}\frac{Var(Z)}{Var(X)} + \beta_{zu}\beta_{xu}\frac{Var(U)}{Var(X)}) + \beta_{yu}\frac{Var(U)}{Var(X)}(\beta_{zu} + \beta_{xz} + \beta_{xz} + \beta_{zu} + \beta_{xz})
$$

Hence, the bias in the OLS estimator of the relation between X and Y, when omitting U and Z is given by:

$$
bias(\hat{\beta}_{yx}) = \beta_{yz}(\beta_{xz}\frac{Var(Z)}{Var(X)} + \beta_{zu}\beta_{xu}\frac{Var(U)}{Var(X)}) + \beta_{yu}\frac{Var(U)}{Var(X)}(\beta_{zu} + \beta_{xz} + \beta_{xz} + \beta_{zu} + \beta_{xz})
$$

(1)
Bias due to one unmeasured confounder

When Z is a measured confounder, while U is unmeasured, the design matrix D is the two column matrix of X and Z.

The expected value of the estimated total effect of X on Y, when adjusting for Z (but without adjustment for U) is then given by the first element of 

\((D' D)^{-1} D' Y\):

\[
\hat{\beta} = (D' D)^{-1} D' Y
= \begin{pmatrix}
Z' Z X' Y - X' Z Z' Y \\
- Z' X X' Y + X' X Z' Y
\end{pmatrix} \frac{1}{X' X Z' Z - X' Z Z' X}.
\]

The first row of \(\hat{\beta}\) corresponds to the estimated effect of X on Y and the second row to the effect of Z on Y. Using the same arguments as before, we can derive an expression for the expected value of the total effect of X on Y, conditional on Z \((\beta_{yx|z})\), but ignoring U:

\[
E[\beta_{yx|z}] = \beta_{yx} + \beta_{xu}(\beta_{yu} V a r(U)(1 - \rho_{zu}^2)) V a r(X)(1 - \rho_{xz}^2)
\]

The expected value of the estimated effect of Z on Y, conditional on X, is given by the second row of \(\hat{\beta}\). The expression for the expected value of the effect of Z on Y, conditional on X \((\beta_{yz|x})\), but ignoring U is given by:

\[
E[\beta_{yz|x}] = \beta_{yu}' \left( \frac{\rho_{zu} - \rho_{xz} \rho_{xu}}{1 - \rho_{xz}^2} \right),
\]

where \(\beta_{yu}'\) represents the conditional (or direct) effect of U on Y, if both are standardized. \(\rho\) again represents (marginal) correlations.