that is, it leaves out the control variable c. When the correct model is that of (3.12) and (3.13), the residual e_i in (3.14) is expressed as

$$e_i = \beta_2 \, c_i + \epsilon_{yi}. \tag{3.15}$$

Because m_i in (3.13) is a function of c_i, the predictor m_i and the residual e_i in (3.14) are correlated because both are a function of c_i. This correlation cannot be identified and estimated. When the correlation is ignored, that is, assumed to be zero, the estimated κ_1 slope is a biased estimate of the slope β_1. Another way of portraying the problem is that the omitted confounder causes the two residuals in the model of Figure 3.5 (b) to be correlated. The residual correlation is not identified because the slope in the regression of y on m already makes the relationship between y and m fit perfectly. When the residual correlation is fixed at zero as shown in the figure, the non-zero true value causes biased estimates. A sensitivity analysis that studies the indirect and direct effects at different values of the residual correlation is described in Section 3.5. When the exposure is randomized, the confounder variable c and the exposure variable are uncorrelated and the regression of m on x correctly estimates γ_1 also when c is left out of the model.

Table 3.19 shows the input corresponding to the model in Figure 3.5 (a). The R^2 for the y regression is 0.28 and the R^2 for the m regression is 0.27. These values are computed as follows. Because x and c are uncorrelated with variances 0.5^2 and one,

$$V(m) = 0.5^2 \times 0.5^2 + 0.4^2 + 0.59 = 0.81, \tag{3.16}$$

$$R^2(m) = (0.5^2 \times 0.5^2 + 0.4^2)/0.81 = 0.27. \tag{3.17}$$

Inserting m in the y equation,

$$\begin{aligned} V(y) &= V(0.5(0.5x + 0.4c + \epsilon_m) + 0.3c + \epsilon_y) \\ &= 0.25^2 \times 0.5^2 + (0.2 + 0.3)^2 + 0.5^2 \times 0.59 + 0.54 = 0.95, \end{aligned} \tag{3.18}$$

$$R^2(y) = (0.25^2 \times 0.5^2 + (0.2 + 0.3)^2)/0.95 = 0.28. \tag{3.19}$$

The parameter values imply that the correlation between the residual e of (3.15) and the mediator m in the model of Figure 3.5 (b) is 0.17. This is obtained as follows using (3.15) and (3.13) and using the fact that e and c have zero means, c has variance one, and m has variance 0.81,

$$Cov(e, m) = E(e \times m) = \gamma_2 \beta_2 E(c^2) = 0.12, \tag{3.20}$$

$$V(e) = 0.3^2 \times 1 + 0.54 = 0.63, \tag{3.21}$$

$$Corr(e, m) = Cov(e, m)/\sqrt{V(e)} \times \sqrt{V(m)} = 0.17, \tag{3.22}$$