STA2101 Assignment 6*

Ahmed Nawaz Amanullah †

March 2021

1

1.1 Problem

In the following regression model, the explanatory variables X_1 and X_2 are random variables. The true model is

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i, \tag{1}$$

independently for $i = 1, ..., n$, where $\epsilon_i \sim N(0, \sigma^2)$.
The mean and covariance matrix of the explanatory variables are given by

$$E\begin{pmatrix} X_{i,1} \\ X_{i,2} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ and } Var\begin{pmatrix} X_{i,1} \\ X_{i,2} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{pmatrix} \tag{2}$$

The explanatory variables $X_{i,1}$ and $X_{i,2}$ are independent of ϵ_i.
Unfortunately $X_{i,2}$ which has an impact on Y_i and is correlated with $X_{i,1}$, is not part of the data set. Since $X_{i,2}$ is not observed, it is absorbed by the intercept and error term, as follows.

†with help from the Overleaf team

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i && (3)\\
&= (\beta_0 + \beta_2\mu_2) + \beta_1 X_{i,1} + (\beta_2 X_{i,2} - \beta_2\mu_2 + \epsilon_i) && (4)\\
&= \beta_0' + \beta_1 X_{i,1} + \epsilon_i' && (5)
\end{aligned}
$$

The primes just denote a new β_0 and a new ϵ_i. It was necessary to add and subtract $\beta_2\mu_2$ in order to obtain $E(\epsilon_i') = 0$. And of course there could be more than one omitted variable. They would all get swallowed by the intercept and error term, the garbage bins of regression analysis.

(a) What is $Cov(X_{i,1}, \epsilon_i')$?

(b) Calculate the variance-covariance matrix of $(X_{i,1}, Y_i)$ under the true model. Is it possible to have non-zero covariance between $X_{i,1}$ and Y_i when $\beta_1 = 0$?

(c) Suppose we want to estimate β_1. The least squares estimator is

$$
\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_1)(Y_i - \bar{Y})}{\sum_{i=1}^n (X_{i,1} - \bar{X}_1)^2}. \qquad (6)
$$

You may just use this formula; you don't have to derive it. Is $\hat{\beta}_1$ a consistent estimator of β_1 if the true model holds? Answer Yes or no and show your work. You may use the consistency of the sample variance and covariance without proof.

(d) Are there *any* points in the parameter space for which $\hat{\beta}_1 \stackrel{p}{\rightarrow} \beta_1$ when the true model holds.

1.2 Solution

(a)

$$
\begin{aligned}
&Cov(X_{i,1}, \epsilon_i') && (7)\\
&= E[(X_{i,1} - E(X_{i,1}))(\epsilon_i' - E(\epsilon_i'))] && (8)\\
&= E[(X_{i,1} - E(X_{i,1}))\epsilon_i'] && (9)\\
&= E[(X_{i,1} - \mu_1)(\beta_2 X_{i,2} - \beta_2\mu_2 + \epsilon_i)] && (10)\\
&= \beta_2 E[(X_{i,1} - \mu_1)(X_{i,2} - \mu_2)] + E[(X_{i,1} - \mu_1)\epsilon_i] && (11)
\end{aligned}
$$

$$= \beta_2 Cov(X_{i,1}, X_{i,2}) + Cov(X_{i,1}, \epsilon_i) \tag{12}$$
$$= \beta_2 \phi_{12} + 0 \tag{13}$$
$$= \beta_2 \phi_{12} \tag{14}$$

(b) We want

$$Cov \begin{pmatrix} X_{i,1} \\ Y_i \end{pmatrix} = \begin{pmatrix} Var(X_{i,1}) & Cov(X_{i,1}, Y_i) \\ Cov(X_{i,1}, Y_i) & Var(Y_i) \end{pmatrix} \tag{15}$$

We are already given $Var(X_{i,1}) = \phi_{11}$. Let's then try to derive $Cov(X_{i,1}, Y_i)$,

$$Cov(X_{i,1}, Y_i) \tag{16}$$
$$= Cov(X_{i,1}, \beta_0' + \beta_1 X_{i,1} + \epsilon_i') \tag{17}$$
$$= E[(X_{i,1} - E(X_{i,1}))(\beta_0' + \beta_1 X_{i,1} + \epsilon_i' - E(\beta_0' + \beta_1 X_{i,1} + \epsilon_i'))] \tag{18}$$
$$= \beta_1 E[(X_{i,1} - \mu_1)^2] + E[(X_{i,1} - \mu_1)\epsilon_i'] \tag{19}$$
$$= \beta_1 Var(X_{i,1}) + Cov(X_{i,1}, \epsilon_i') \tag{20}$$
$$= \beta_1 \phi_{11} + \beta_2 \phi_{12} \tag{21}$$

We still have $Var(Y_i)$ left to derive,

$$Var(Y_i) \tag{22}$$
$$= E[(\beta_0' + \beta_1 X_{i,1} + \epsilon_i' - E(\beta_0' + \beta_1 X_{i,1} + \epsilon_i'))^2] \tag{23}$$
$$= E[(\beta_1 (X_{i,1} - \mu_1) + \epsilon_i')^2] \tag{24}$$
$$= E[\beta_1^2 (X_{i,1} - \mu_1)^2 + 2\beta_1 (X_{i,1} - \mu_1)\epsilon_i' + (\epsilon_i')^2] \tag{25}$$
$$= \beta_1^2 E[(X_{i,1} - \mu_1)^2] + 2\beta_1 E[(X_{i,1} - \mu_1)\epsilon_i'] + E[(\epsilon_i')^2] \tag{26}$$
$$= \beta_1^2 Var(X_{i,1}) + 2\beta_1 Cov(X_{i,1}, \epsilon_i') + E[(\beta_2 X_{i,2} - \beta_2 \mu_2 + \epsilon_i)^2] \tag{27}$$
$$= \beta_1^2 \phi_{11} + 2\beta_1 \beta_2 \phi_{12} + E[\beta_2^2 (X_{i,2} - \mu_2)^2 + 2\beta_2 (X_{i,2} - \mu_2)\epsilon_i + \epsilon_i^2] \tag{28}$$
$$= \beta_1^2 \phi_{11} + 2\beta_1 \beta_2 \phi_{12} + \beta_2^2 E[(X_{i,2} - \mu_2)^2] + 2\beta_2 E[(X_{i,2} - \mu_2)\epsilon_i] + E[\epsilon_i^2] \tag{29}$$
$$= \beta_1^2 \phi_{11} + 2\beta_1 \beta_2 \phi_{12} + \beta_2^2 Var(X_{i,2}) + 2\beta_2 Cov(X_{i,2}, \epsilon_i) + Var(\epsilon_i) \tag{30}$$
$$= \beta_1^2 \phi_{11} + 2\beta_1 \beta_2 \phi_{12} + \beta_2^2 \phi_{22} + 0 + \sigma^2 \tag{31}$$
$$= \beta_1^2 \phi_{11} + 2\beta_1 \beta_2 \phi_{12} + \beta_2^2 \phi_{22} + \sigma^2 \tag{32}$$

Thus, the variance-covariance matrix of $(X_{i,1}, Y_i)$ under the true model is

$$Cov \begin{pmatrix} X_{i,1} \\ Y_i \end{pmatrix} = \begin{pmatrix} \phi_{11} & \beta_1 \phi_{11} + \beta_2 \phi_{12} \\ \beta_1 \phi_{11} + \beta_2 \phi_{12} & \beta_1^2 \phi_{11} + 2\beta_1 \beta_2 \phi_{12} + \beta_2^2 \phi_{22} + \sigma^2 \end{pmatrix} \tag{33}$$

Since $Cov(X_{i,1}, Y_i) = \beta_1\phi_{11} + \beta_2\phi_{12}$, when $\beta_1 = 0$, $Cov(X_{i,1}, Y_i) = \beta_2\phi_{12}$. Thus, under the true model, it is possible to have non-zero covariance between $X_{i,1}$ and Y_i when $\beta_1 = 0$ (specifically this is the case when both $X_{i,1}$ and $X_{i,2}$ have non-zero covariance **and** allowing for $X_{i,1}$, $X_{i,2}$ and Y_i are related, i.e. $\beta_2 \neq 0$).

(c)

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_1)(Y_i - \bar{Y})}{\sum_{i=1}^n (X_{i,1} - \bar{X}_1)^2} && (34)\\
&= \frac{\frac{1}{n}\sum_{i=1}^n (X_{i,1} - \bar{X}_1)(Y_i - \bar{Y})}{\frac{1}{n}\sum_{i=1}^n (X_{i,1} - \bar{X}_1)^2} && (35)\\
&\xrightarrow{p} \frac{Cov(X_{i,1}, Y_i)}{Var(X_{i,1})} && (36)\\
&= \frac{\beta_1\phi_{11} + \beta_2\phi_{12}}{\phi_{11}} && (37)\\
&= \beta_1 + \beta_2\frac{\phi_{12}}{\phi_{11}} \neq \beta_1 && (38)
\end{aligned}$$

So, **no**, $\hat{\beta}_1$ is not a consistent estimator of β_1 if the true model holds.

N.B. In the above rough work we used the consistency of sample variance and covariance to get $\frac{\frac{1}{n}\sum_{i=1}^n (X_{i,1}-\bar{X}_1)(Y_i-\bar{Y})}{\frac{1}{n}\sum_{i=1}^n (X_{i,1}-\bar{X}_1)^2} \xrightarrow{p} \frac{Cov(X_{i,1},Y_i)}{Var(X_{i,1})}$.

(d) We observe that $\beta_1 + \beta_2\frac{\phi_{12}}{\phi_{11}} = \beta_1$ iff $\beta_2\frac{\phi_{12}}{\phi_{11}} = 0$. So, if either $X_{i,1}$ and $X_{i,2}$ have zero covariance, i.e. $\phi_{12} = 0$ **or** allowing for $X_{i,1}$, $X_{i,2}$ and Y_i are not related, i.e. $\beta_2 = 0$, then $\hat{\beta}_1 \xrightarrow{p} \beta_1 + \beta_2\frac{\phi_{12}}{\phi_{11}} = \beta_1$ when the true model holds. Thus, there are points in the parameter space for which $\hat{\beta}_1 \xrightarrow{p} \beta_1$ when the true model holds (specifically, as mentioned earlier, points with either $\beta_2 = 0$ or $\phi_{12} = 0$).

2

2.1 Problem

Independently for $i = 1, ..., n$, let $Y_i = \beta X_i + \epsilon_i$, where $X_i \sim N(\mu, \sigma_x^2)$ and $\epsilon_i \sim N(0, \sigma_\epsilon^2)$. Because of omitted variables that influence both X_i and Y_i, we have $Cov(X_i, \epsilon_i) = c \neq 0$.

(a) The least squares estimator of β is $\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}$. Is this estimator consistent? Answer Yes or No and prove your answer.

(b) Give the parameter space for this model. There are some constraints on c.

(c) First consider points in the parameter space where $\mu \neq 0$. Give an estimator of β that converges almost surely to the right answer for that part of the parameter space. If you are not sure how to proceed, try calculating the expected value and covariance matrix of (X_i, Y_i).

(d) What happens in the rest of the parameter space, that is where $\mu = 0$? Is a consistent estimator possible there? So we see that parameters may be identifiable in some parts of the parameter space but not all.

2.2 Solution

(a) To begin, we want the population moments $E(X_i Y_i)$ and $E(X_i^2)$. Starting with $E(X_i Y_i)$,

$$\begin{aligned}
E(X_i Y_i) &= E(X_i(\beta X_i + \epsilon_i)) && (39)\\
&= \beta E(X_i^2) + E(X_i \epsilon_i) && (40)\\
&= \beta(Var(X_i) + E(X_i)^2) + Cov(X_i, \epsilon_i) + E(X_i)E(\epsilon_i) && (41)\\
&= \beta(\sigma_x^2 + \mu^2) + c + 0 && (42)\\
&= \beta\sigma_x^2 + \beta\mu^2 + c && (43)
\end{aligned}$$

Next $E(X_i^2)$,

$$\begin{aligned}
E(X_i^2) &= Var(X_i) + E(X_i)^2 && (44)\\
&= \sigma_x^2 + \mu^2 && (45)
\end{aligned}$$

So, by law of large numbers and function continuity,

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2} && (46)\\
&= \frac{\frac{1}{n}\sum_{i=1}^n X_i Y_i}{\frac{1}{n}\sum_{i=1}^n X_i^2} && (47)\\
&\xrightarrow{a.s.} \frac{E(X_i Y_i)}{E(X_i^2)} && (48)
\end{aligned}$$

$$= \frac{\beta(\sigma_x^2 + \mu^2) + c}{\sigma_x^2 + \mu^2} \tag{49}$$

$$= \beta + \frac{c}{\sigma_x^2 + \mu^2} \neq \beta \tag{50}$$

So, **no**, the least squares estimator of β, $\hat{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}$, is not a consistent estimator of β.

(b) $\Theta = \{(\beta, \mu, \sigma_x^2, \sigma_\epsilon^2, c) : -\infty < \beta < \infty, -\infty < \mu < \infty, \sigma_x^2 > 0, \sigma_\epsilon^2 > 0, -\infty < c < \infty, c \neq 0, -\sigma_x \sigma_\epsilon \leq c \leq \sigma_x \sigma_\epsilon\}$

N.B. Unsure if $c \neq 0$ ought to be part of the parameter space, included it because question mentions it. Also, we got $-\sigma_x \sigma_\epsilon \leq c \leq \sigma_x \sigma_\epsilon$ from the fact that correlation between X_i and ϵ_i is bound between -1 and 1 inclusive.

(c) Let's begin with the hint, starting with computing the expected value matrix of (X_i, Y_i),

$$E \begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} E(X_i) \\ E(Y_i) \end{pmatrix} = \begin{pmatrix} \mu \\ \beta\mu \end{pmatrix} \tag{51}$$

Next, the covariance matrix of (X_i, Y_i),

$$Cov \begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} Var(X_i) & Cov(X_i, Y_i) \\ Cov(X_i, Y_i) & Var(Y_i) \end{pmatrix} \tag{52}$$

We are already given $Var(X_i) = \sigma_x^2$. Moving on to $Cov(X_i, Y_i)$,

$$Cov(X_i, Y_i) = E[(X_i - E(X_i))(Y_i - E(Y_i))] \tag{53}$$

$$= E(X_i Y_i) - E(X_i)E(Y_i) \tag{54}$$

$$= \beta\sigma_x^2 + \beta\mu^2 + c - \mu(\beta\mu) \tag{55}$$

$$= \beta\sigma_x^2 + c \tag{56}$$

Finally, $Var(Y_i)$,

$$Var(Y_i) = E[(Y_i - E(Y_i))^2] \tag{57}$$

$$= E[(\beta X_i + \epsilon_i - \beta\mu)^2] \tag{58}$$

$$= \beta^2 E[(X_i - \mu)^2] + 2\beta E[(X_i - \mu)\epsilon_i] + E[\epsilon_i^2] \tag{59}$$

$$
\begin{aligned}
&= \beta^2 Var(X_i) + 2\beta Cov(X_i, \epsilon_i) + Var(\epsilon_i) && (60)\\
&= \beta^2 \sigma_x^2 + 2\beta c + \sigma_\epsilon^2 && (61)
\end{aligned}
$$

Thus the the covariance matrix of (X_i, Y_i),

$$
\begin{aligned}
Cov\begin{pmatrix} X_i \\ Y_i \end{pmatrix} &= \begin{pmatrix} Var(X_i) & Cov(X_i, Y_i) \\ Cov(X_i, Y_i) & Var(Y_i) \end{pmatrix} && (62)\\
&= \begin{pmatrix} \sigma_x^2 & \beta\sigma_x^2 + c \\ \beta\sigma_x^2 + c & \beta^2\sigma_x^2 + 2\beta c + \sigma_\epsilon^2 \end{pmatrix} && (63)
\end{aligned}
$$

By inspecting our expected value and covariance matrices, we note that $\hat{\beta} = \frac{\bar{Y}}{\bar{X}} = \frac{\frac{1}{n}\sum_{i=1}^n Y_i}{\frac{1}{n}\sum_{i=1}^n X_i} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i}$ suits our purposes. We can show this as follows, using law of large numbers and function continuity,

$$
\begin{aligned}
\hat{\beta} &= \frac{\bar{Y}}{\bar{X}} && (64)\\
&\xrightarrow{a.s.} \frac{E(Y)}{E(X)} && (65)\\
&= \frac{\beta\mu}{\mu} && (66)\\
&= \beta && (67)
\end{aligned}
$$

Thus for parameter space where $\mu \neq 0$, the estimator $\hat{\beta} = \frac{\bar{Y}}{\bar{X}} = \frac{\frac{1}{n}\sum_{i=1}^n Y_i}{\frac{1}{n}\sum_{i=1}^n X_i} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i}$ converges almost surely to the right answer (i.e. β).

(d) Provided $\mu = 0$, the moment structure equations for the means evaluate to zero and they are of no further use. So we are left with the three covariance structure equations and four unknowns. We note that while we can solve for σ_x^2, we are still left with two equations and three unknowns. So, we fail the parameter count rule and the set of points in the parameter space where there is a unique solution occupies a set of volume zero and it is impossible to identify the whole parameter vector. More specifically, since we have two equations and three unknowns (and one of them is β), there are infinitely many of the three remaining parameters that give rise to the same distribution of sample data, making β not identifiable and so a consistent estimator for β is not possible where $\mu = 0$.

3

3.1 Problem

We know that omitted explanatory variables are a big problem, because they induce non-zero covariance between the explanatory variables and the error terms ϵ_i. The residuals have a lot in common with the ϵ_i terms in a regression model, though they are not the same thing. A reasonable idea is to check for correlation between explanatory variables and the ϵ_i values by looking at the correlation between the residuals and explanatory variables.

Accordingly, for a multiple regression model with an intercept so that $\sum_{i=1}^{n} e_i = 0$, calculate the sample correlation r between explanatory variable j and the residuals $e_1, ..., e_n$. Use this formula for the correlation: $r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$. Simplify. What can the sample correlations between residuals and x variables tell you about the correlation between ϵ and the x variables?

3.2 Solution

$$r_{x_j,e} = \frac{\sum_{i=1}^{n}(x_{i,j}-\bar{x}_j)(e_i-\bar{e})}{\sqrt{\sum_{i=1}^{n}(x_{i,j}-\bar{x}_j)^2}\sqrt{\sum_{i=1}^{n}(e_i-\bar{e})^2}} \tag{68}$$

$$= \frac{\sum_{i=1}^{n} e_i(x_{i,j}-\bar{x}_j)}{\sqrt{\sum_{i=1}^{n}(x_{i,j}-\bar{x}_j)^2}\sqrt{\sum_{i=1}^{n} e_i^2}} \tag{69}$$

$$= \frac{(\sum_{i=1}^{n} e_i x_{i,j}) - \bar{x}_j \sum_{i=1}^{n} e_i}{\sqrt{\sum_{i=1}^{n}(x_{i,j}-\bar{x}_j)^2}\sqrt{\sum_{i=1}^{n} e_i^2}} \tag{70}$$

$$= \frac{\sum_{i=1}^{n} e_i x_{i,j}}{\sqrt{\sum_{i=1}^{n}(x_{i,j}-\bar{x}_j)^2}\sqrt{\sum_{i=1}^{n} e_i^2}} \tag{71}$$

$$= \frac{\mathbf{X}_j^T\mathbf{e}}{\sqrt{\sum_{i=1}^{n}(x_{i,j}-\bar{x}_j)^2}\sqrt{\sum_{i=1}^{n} e_i^2}} \tag{72}$$

where $\mathbf{X}_j$ is the $n \times 1$ column vector related to explanatory variable j of the $n \times p$ design matrix $\mathbf{X}$ and $\mathbf{e}$ is a $n \times 1$ column vector consisting of the residuals. Thus, the quantity $\mathbf{X}_j^T\mathbf{e}$ is the jth element of the matrix product $\mathbf{X}^T\mathbf{e}$ which evaluates to

$$\mathbf{X}^T\mathbf{e} = \mathbf{X}^T(\mathbf{y} - \hat{\mathbf{y}}) \tag{73}$$

$$
\begin{aligned}
&= \mathbf{X}^T(\mathbf{y} - \hat{\mathbf{y}}) && (74)\\
&= \mathbf{X}^T\mathbf{y} - \mathbf{X}^T\hat{\mathbf{y}} && (75)\\
&= \mathbf{X}^T\mathbf{y} - \mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} && (76)\\
&= \mathbf{X}^T\mathbf{y} - \mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} && (77)\\
&= \mathbf{X}^T\mathbf{y} - \mathbf{X}^T\mathbf{y} = \mathbf{0} && (78)
\end{aligned}
$$

Thus,

$$
r_{x_j,e} = \frac{\mathbf{X}_j^T\mathbf{e}}{\sqrt{\sum_{i=1}^{n}(x_{i,j} - \bar{x}_j)^2}\sqrt{\sum_{i=1}^{n} e_i^2}} = 0 \qquad (79)
$$

and the sample correlations between residuals and x variables tells us nothing about the correlation between ϵ and the x variables.

4

4.1 Problem

This question explores the consequences of ignoring measurement error in the response variable. Independently for $i = 1, ..., n$, let

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_i + \epsilon_i && (80)\\
V_i &= Y_i + \epsilon_i, && (81)
\end{aligned}
$$

where $Var(X_i) = \phi$, $E(X_i) = \mu_x$, $Var(e_i) = \omega$, $Var(\epsilon_i) = \psi$ and X_i, e_i, ϵ_i are all independent. The explanatory variable X_i is observable, but the response variable Y_i is latent. Instead of Y_i, we can see V_i, which is Y_i plus a piece of random noise. Call this the *true model.*

(a) Make a path diagram of the true model.

(b) Strictly speaking, the distributions of X_i, e_i and ϵ_i are unknown parameters because they are unspecified. But suppose we are interested in identifying just the Greek-letter parameters. Does the true model pass the test of the Parameter Count Rule? Answer Yes or No and give the numbers.

(c) Calculate the variance-covariance matrix of the observable variables as a function of the model parameters. Show your work.

(d) Suppose that the analyst assumes that V_i is the same thing as Y_i, and fits the naive model $V_i = \beta_0 + \beta_1 X_i + \epsilon_i$, in which

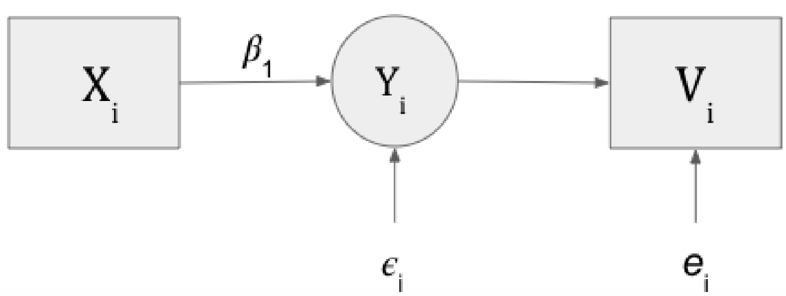
$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(V_i - \bar{V})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}. \tag{82}$$

Assuming the *true* model (not the naive model), is $\hat{\beta}_1$ a consistent estimator of β_1? Answer Yes or No and show your work.

(e) Why does this prove that β_1 is identifiable?

4.2 Solution

(a) The path diagram for the true model is provided below



(b) We have two observable variables, so we have **five moment structure equations** (two from expected values and three from variance-covariances). Provided we are interested in identifying just the Greek-letter parameters, we have β_0, β_1, ϕ, μ_x, ω, ψ, that is **six parameters**. So, **No**, the true model does not pass the Parameter Count Rule.

(c) We want

$$Cov\begin{pmatrix} X_i \\ V_i \end{pmatrix} = \begin{pmatrix} Var(X_i) & Cov(X_i, V_i) \\ Cov(X_i, V_i) & Var(V_i) \end{pmatrix} \tag{83}$$

We are already provided $Var(X_i) = \phi$. So, let's move on to $Cov(X_i, V_i)$,

$$
\begin{aligned}
Cov(X_i, V_i) &= Cov(X_i, Y_i + e_i) && (84)\\
&= Cov(X_i, Y_i) + Cov(X_i, e_i) && (85)\\
&= Cov(X_i, Y_i) + 0 && (86)\\
&= Cov(X_i, \beta_0 + \beta_1 X_i + \epsilon_i) && (87)\\
&= Cov(X_i, \beta_1 X_i + \epsilon_i) && (88)\\
&= Cov(X_i, \beta_1 X_i) + Cov(X_i, \epsilon_i) && (89)\\
&= \beta_1 Cov(X_i, X_i) && (90)\\
&= \beta_1 \phi && (91)
\end{aligned}
$$

Finally, let's compute $Var(V_i)$,

$$
\begin{aligned}
Var(V_i) &= Var(Y_i + e_i) && (92)\\
&= E[(Y_i + e_i - E(Y_i + e_i))^2] && (93)\\
&= E[(\beta_0 + \beta_1 X_i + \epsilon_i + e_i - E(\beta_0 + \beta_1 X_i + \epsilon_i))^2] && (94)\\
&= E[(\beta_1 X_i - \beta_1 \mu_x + \epsilon_i + e_i)^2] && (95)\\
&= E[\beta_1^2 (X_i - \mu_x)^2 + 2\beta_1 (X_i - \mu_x)(\epsilon_i + e_i) + (\epsilon_i + e_i)^2] && (96)\\
&= \beta_1^2 Var(X_i) + 2\beta_1 (Cov(X_i, \epsilon_i) + Cov(X_i, e_i)) + E[(\epsilon_i + e_i)^2] && (97)\\
&= \beta_1^2 \phi + 0 + E[\epsilon_i^2 + 2\epsilon_i e_i + e_i^2] && (98)\\
&= \beta_1^2 \phi + Var(\epsilon_i) + 2Cov(\epsilon_i, e_i) + Var(e_i) && (99)\\
&= \beta_1^2 \phi + \psi + 0 + \omega && (100)\\
&= \beta_1^2 \phi + \psi + \omega && (101)
\end{aligned}
$$

With this we have the variance-covariance matrix as,

$$
\begin{aligned}
Cov \begin{pmatrix} X_i \\ V_i \end{pmatrix} &= \begin{pmatrix} Var(X_i) & Cov(X_i, V_i) \\ Cov(X_i, V_i) & Var(V_i) \end{pmatrix} && (102)\\
&= \begin{pmatrix} \phi & \beta_1 \phi \\ \beta_1 \phi & \beta_1^2 \phi + \psi + \omega \end{pmatrix} && (103)
\end{aligned}
$$

(d) Assuming the consistency of the sample covariance and variance moments (as we did in question 1, or apply SLLN and function continuity (do we need to assume anything?)),

$$
\hat{\beta_1} = \frac{\sum_{i=1}^n (X_i - \bar{X})(V_i - \bar{V})}{\sum_{i=1}^n (X_i - \bar{X})^2} \tag{104}
$$

$$
\begin{aligned}
&= \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(V_i - \bar{V})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} && (105)\\
&\xrightarrow{a.s.} \frac{E[(X_i - E(X_i))(V_i - E(V_i))]}{E[(X_i - E(X_i))^2]} && (106)\\
&= \frac{Cov(X_i, V_i)}{Var(X_i)} && (107)\\
&= \frac{\beta_1 \phi}{\phi} && (108)\\
&= \beta_1 && (109)
\end{aligned}
$$

And thus we say that **Yes**, $\hat{\beta}_1$ is a consistent estimator of β_1.

(e) This is because consistency is impossible without identifiability (i.e. according to theorem discussed in class, if parameter vector is unidentifiable, consistent estimation is impossible), and since we have shown that $\hat{\beta}_1$ is a consistent estimator of β_1, β_1 is identifiable.

5

5.1 Problem

This question explores the consequences of ignoring measurement error in the explanatory variable when there is only one explanatory variable. Independently for $i = 1, ..., n$, let

$$
\begin{aligned}
Y_i &= \beta X_i + \epsilon_i && (110)\\
W_i &= X_i + e_i && (111)
\end{aligned}
$$

where all random variables are normal with expected value zero, $Var(X_i) = \phi > 0$, $Var(\epsilon_i) = \psi > 0$, $Var(e_i) = \omega > 0$ and ϵ_i, e_i and X_i are all independent. The variables W_i and Y_i are observable, while X_i is latent. Error terms are never observable.

(a) What is the parameter vector $\boldsymbol{\theta}$ for this model?

(b) Denote the covariance matrix of the observable variable by $\boldsymbol{\Sigma} = [\sigma_{ij}]$. The unique σ_{ij} values are the moments, and there is a covariance structure equation for each one. Calculate the variance-covariance matrix

$\boldsymbol{\Sigma}$ of the observable variables, expressed as a function of the model parameters. You now have the covariance structure equations.

(c) Does this model pass the test of the parameter count rule? Answer Yes or No and give the numbers.

(d) Are there any points in the parameter space where the parameter β is identifiable? Are there infinitely many, or just one point?

(e) The naive estimator of β is

$$\hat{\beta}_n = \frac{\sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i^2}. \tag{112}$$

Is $\hat{\beta}_n$ a consistent estimator of β? Answer Yes or No. To what does $\hat{\beta}_n$ converge?

(f) Are there any points in the parameter space for which $\hat{\beta}_n$ converges to the right answer? Compare your answer to the set of points where β is identifiable.

(g) Suppose the reliability of W_i were known[1], or to be more realisitic, suppose that a good estimate of the reliability were available; call it r^2_{wx}. How could you use r^2_{wx} to improve $\hat{\beta}_n$? Give the formula for an improved estimator of β.

5.2 Solution

(a) $\boldsymbol{\theta} = (\beta, \phi, \psi, \omega)$

(b) We want

$$Cov \begin{pmatrix} Y_i \\ W_i \end{pmatrix} = \begin{pmatrix} Var(Y_i) & Cov(Y_i, W_i) \\ Cov(Y_i, W_i) & Var(W_i) \end{pmatrix} \tag{113}$$

Starting with $Var(Y_i)$,

$$Var(Y_i) = Var(\beta X_i + \epsilon_i) \tag{114}$$

[1]As a reminder, the reliability of an observed measurement is the proportion of its variace (variance?) that comes from the "true" latent variable it is measuring. Here, the reliability of W_i is $\frac{\phi}{\phi+\omega}$.

$$
\begin{aligned}
&= E[(\beta X_i + \epsilon_i - E(\beta X_i + \epsilon_i))^2] && (115)\\
&= E[(\beta X_i + \epsilon_i)^2] && (116)\\
&= E[\beta^2 X_i^2 + 2\beta X_i \epsilon_i + \epsilon_i^2] && (117)\\
&= \beta^2 E(X_i^2) + 2\beta E(X_i \epsilon_i) + E(\epsilon_i^2) && (118)\\
&= \beta^2 Var(X_i) + 2\beta Cov(X_i, \epsilon_i) + Var(\epsilon_i) && (119)\\
&= \beta^2 \phi + 0 + \psi && (120)\\
&= \beta^2 \phi + \psi && (121)
\end{aligned}
$$

Moving onto $Cov(Y_i, W_i)$,

$$
\begin{aligned}
Cov(Y_i, W_i) &= Cov(\beta X_i + \epsilon_i, X_i + e_i) && (122)\\
&= Cov(\beta X_i, X_i) + Cov(\beta X_i, e_i) + Cov(\epsilon_i, X_i) + Cov(\epsilon_i, e_i) && (123)\\
&= \beta Cov(X_i, X_i) + 0 + 0 + 0 && (124)\\
&= \beta \phi && (125)
\end{aligned}
$$

Finally, $Var(W_i)$,

$$
\begin{aligned}
Var(W_i) &= Var(X_i + e_i) && (126)\\
&= E[(X_i + e_i - E(X_i + e_i))^2] && (127)\\
&= E[(X_i + e_i)^2] && (128)\\
&= E[X_i^2 + 2X_i e_i + e_i^2] && (129)\\
&= E(X_i^2) + 2E(X_i e_i) + E(e_i^2) && (130)\\
&= Var(X_i) + 2Cov(X_i, e_i) + Var(e_i) && (131)\\
&= \phi + 0 + \omega && (132)\\
&= \phi + \omega && (133)
\end{aligned}
$$

With this we have the variance-covariance matrix as,

$$
\begin{aligned}
Cov \begin{pmatrix} Y_i \\ W_i \end{pmatrix} &= \begin{pmatrix} Var(Y_i) & Cov(Y_i, W_i) \\ Cov(Y_i, W_i) & Var(W_i) \end{pmatrix} && (134)\\
&= \begin{pmatrix} \beta^2\phi + \psi & \beta\phi \\ \beta\phi & \phi + \omega \end{pmatrix} && (135)
\end{aligned}
$$

(c) We have **three moment structure equations** (all of them covariance structure equations since the means are useless) and **four parameters**. So, **No**, model does not pass parameter count rule.

(d) Well, we know from (c) that the parameter vector as a whole is unidentifiable. However, there may be sets where β is identifiable that have volume zero.

What if $\beta = 0$? Since $\phi > 0$ according to the parameter space, $Cov(Y_i, W_i) = 0$ iff $\beta = 0$. So β is identifiable if $\beta = 0$. This is a 4D hyperplane (?) in the parameter space.

Additionally, what if we make two of the parameters related? Doing so, we will be able to use substitution to create a system of 3 equations with 3 unknowns and we may be able to solve for a unique β. So, let's say $\omega = \alpha\phi$ for a $\alpha > 0$. Then,

$$\begin{aligned}
Cov\begin{pmatrix} Y_i \\ W_i \end{pmatrix} &= \begin{pmatrix} Var(Y_i) & Cov(Y_i, W_i) \\ Cov(Y_i, W_i) & Var(W_i) \end{pmatrix} && (136) \\
&= \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} && (137) \\
&= \begin{pmatrix} \beta^2\phi + \psi & \beta\phi \\ \beta\phi & \phi + \omega \end{pmatrix} && (138) \\
&= \begin{pmatrix} \beta^2\phi + \psi & \beta\phi \\ \beta\phi & \phi + \alpha\phi \end{pmatrix} && (139) \\
&= \begin{pmatrix} \beta^2\phi + \psi & \beta\phi \\ \beta\phi & (1+\alpha)\phi \end{pmatrix} && (140)
\end{aligned}$$

And we note that β is identifiable in this case (specifically we can solve for β, getting $\beta = \frac{(1+\alpha)v_{12}}{v_{22}}$).

Thus, there are points in the parameter space where β is identifiable. Since $\beta = 0$ and $\omega = \alpha\phi$ are hyperplanes (?) in the 4D parameter space, they consist of infinitely many points. So there are infinitely many points where β is identifiable.

(e) Using SLLN and function continuity (or continuous mapping?),

$$\begin{aligned}
\hat{\beta}_n &= \frac{\sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i^2} && (141) \\
&= \frac{\frac{1}{n}\sum_{i=1}^n W_i Y_i}{\frac{1}{n}\sum_{i=1}^n W_i^2} && (142) \\
&\xrightarrow{a.s.} \frac{E(W_i Y_i)}{E(W_i^2)} && (143)
\end{aligned}$$

$$= \frac{Cov(W_i, Y_i)}{Var(W_i^2)} \tag{144}$$

$$= \frac{\beta\phi}{\phi + \omega} \tag{145}$$

$$\neq \beta \tag{146}$$

So $\hat{\beta_n}$ converges to $\frac{\beta\phi}{\phi+\omega}$. Consequently, **No**, $\hat{\beta_n}$ is not a consistent estimator of β.

(f) From (e), we have that

$$\hat{\beta_n} \xrightarrow{a.s.} \frac{\beta\phi}{\phi + \omega} \tag{147}$$

To have $\frac{\beta\phi}{\phi+\omega} = \beta$ or $\beta(\frac{\phi}{\phi+\omega} - 1) = 0$ we must have $\frac{\phi}{\phi+\omega} = 1$ or $\beta = 0$. But if $\frac{\phi}{\phi+\omega} = 1$, then $\omega = 0$, which is not in the parameter space (we are given $\omega > 0$ for the parameter space). The other possibility is $\beta = 0$. Then $\frac{\beta\phi}{\phi+\omega} = \beta = 0$ and $\hat{\beta_n}$ converges to the right answer. So there are points in the parameter space for which $\hat{\beta_n}$ converges to the right answer, with them matching one of the set of points found in (d) (specifically the set of points with $\beta = 0$) where β is identifiable. This makes sense because, according to theorem discussed in class, if parameter vector is unidentifiable, consistent estimation is impossible.

(g) We are provided a good estimate of reliability, r_{wx}^2. Let's assume being a good estimate means it is consistent, i.e. $r_{wx}^2 \xrightarrow{a.s.} \frac{\phi}{\phi+\omega}$. Then let the improved estimator be

$$\hat{\beta_n} = \frac{\sum_{i=1}^n W_i Y_i}{r_{wx}^2 \sum_{i=1}^n W_i^2} \tag{148}$$

Then by SLLN and function continuity (or continuous mapping?),

$$\hat{\beta_n} = \frac{\sum_{i=1}^n W_i Y_i}{r_{wx}^2 \sum_{i=1}^n W_i^2} \tag{149}$$

$$= \frac{\frac{1}{n}\sum_{i=1}^n W_i Y_i}{r_{wx}^2 \frac{1}{n}\sum_{i=1}^n W_i^2} \tag{150}$$

$$\xrightarrow{a.s.} \frac{E(W_i Y_i)}{\frac{\phi}{\phi+\omega} E(W_i^2)} \tag{151}$$

$$= \frac{Cov(W_i, Y_i)}{\frac{\phi}{\phi+\omega} Var(W_i^2)} \tag{152}$$

$$= \frac{\beta\phi}{\frac{\phi}{\phi+\omega}(\phi+\omega)} \tag{153}$$

$$= \beta \tag{154}$$

and we note that our improved estimator is consistent.

6

6.1 Problem

The improved version of $\hat{\beta}_n$ in the last question is an example of *correction for attenuation* (weakening) caused by measurement error. Here is the version that applies to correlation. Independently for $i = 1, ..., n$, let

$$\begin{array}{l} D_{i,1} = F_{i,1} + e_{i,1} \\ D_{i,2} = F_{i,2} + e_{i,2} \end{array} \quad cov\begin{pmatrix} F_{i,1} \\ F_{i,2} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{pmatrix} \quad cov\begin{pmatrix} e_{i,1} \\ e_{i,2} \end{pmatrix} = \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix}$$

To make this concrete, it would be natural for psychologists to be interested in the correlation between intelligence and self-esteem, but what they want to know is the correlation between *true* intelligence and *true* self-esteem, not just the (correlation?) between score on an IQ test and score on a self-esteem questionnaire. So for subject i, let $F_{i,1}$ represent true intelligence and $F_{i,2}$ represent true self-esteem, while $D_{i,1}$ is the subject's score on an intelligence test and $D_{i,2}$ (?) is score on a self-esteem questionnaire.

(a) Make a path diagram of this model.

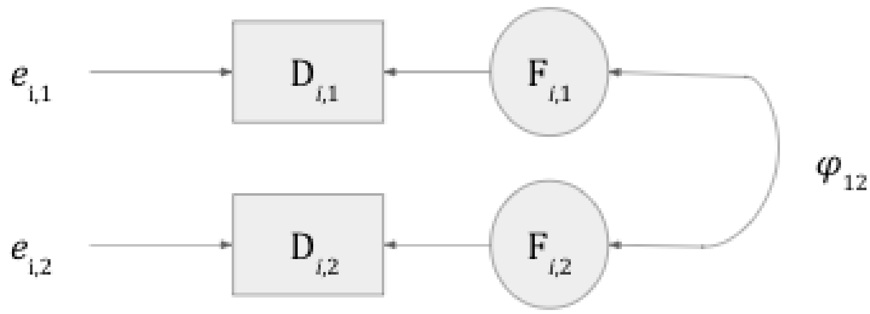
(b) Show that $|Corr(D_{i,1}, D_{i,2})| \leq |Corr(F_{i,1}, F_{i,2})|$. That is, measurement error weakens (attenuates) the correlation.

(c) Suppose the reliability of $D_{i,1}$ is ρ_1^2 and the reliability of $D_{i,2}$ is ρ_2^2. How could you apply ρ_1^2 and ρ_2^2 to $Corr(D_{i,1}, D_{i,2})$, to obtain $Corr(F_{i,1}, F_{i,2})$?

(d) You obtain a sample correlation between IQ score and self-esteem score of $r = 0.25$, which is disappointingly low. From other data, the estimated reliability of the IQ test is $r_1^2 = 0.90$, and the estimated reliability of the self-esteem scale is $r_2^2 = 0.75$. Give an estimate of the correlation between true intelligence and true self-esteem. The answer is a number.

6.2 Solution

(a) The path diagram for the model is provided below



(b) Firstly, we have

$$Corr(F_{i,1}, F_{i,2}) = \frac{Cov(F_{i,1}, F_{i,2})}{\sqrt{Var(F_{i,1})Var(F_{i,2})}} \tag{155}$$

$$= \frac{\phi_{12}}{\sqrt{\phi_{11}\phi_{22}}} \tag{156}$$

Now, we need to compute the variance-covariance matrix of $(D_{i,1}, D_{i,2})$,

$$Cov\begin{pmatrix} D_{i,1} \\ D_{i,2} \end{pmatrix} = \begin{pmatrix} Var(D_{i,1}) & Cov(D_{i,1}, D_{i,2}) \\ Cov(D_{i,1}, D_{i,2}) & Var(D_{i,2}) \end{pmatrix} \tag{157}$$

Starting with $Var(D_{i,1})$,

$$Var(D_{i,1}) = Var(F_{i,1} + e_{i,1}) \tag{158}$$

$$= Cov(F_{i,1} + e_{i,1}, F_{i,1} + e_{i,1}) \tag{159}$$

$$= Cov(F_{i,1}, F_{i,1}) + Cov(F_{i,1}, e_{i,1}) + Cov(e_{i,1}, F_{i,1}) + Cov(e_{i,1}, e_{i,1}) \tag{160}$$

$$= \phi_{11} + \omega_1 + 2Cov(F_{i,1}, e_{i,1}) \tag{161}$$

Assuming the error terms and the *true* variables are independent (i.e. $Cov(F_{i,j}, e_{i,k}) = 0$ for $j = 1, 2$ and $k = 1, 2$) (we are not provided this), then $Var(D_{i,1}) = \phi_{11} + \omega_1$. Moving onto $Cov(D_{i,1}, D_{i,2})$,

$$Cov(D_{i,1}, D_{i,2}) = Cov(F_{i,1} + e_{i,1}, F_{i,2} + e_{i,2}) \tag{162}$$

$$= Cov(F_{i,1}, F_{i,2}) + Cov(F_{i,1}, e_{i,2}) + Cov(e_{i,1}, F_{i,2}) + Cov(e_{i,1}, e_{i,2}) \tag{163}$$

$$= \phi_{12} + 0 + Cov(F_{i,1}, e_{i,2}) + Cov(e_{i,1}, F_{i,2}) \tag{164}$$

Once again, assuming the error terms and the *true* variables are independent (i.e. $Cov(F_{i,j}, e_{i,k}) = 0$ for $j = 1, 2$ and $k = 1, 2$) (we are not provided this), then $Cov(D_{i,1}, D_{i,2}) = \phi_{12}$. Finally, $Var(D_{i,2})$,

$$Var(D_{i,2}) = Var(F_{i,2} + e_{i,2}) \tag{165}$$

$$= Cov(F_{i,2} + e_{i,2}, F_{i,2} + e_{i,2}) \tag{166}$$

$$= Cov(F_{i,2}, F_{i,2}) + Cov(F_{i,2}, e_{i,2}) + Cov(e_{i,2}, F_{i,2}) + Cov(e_{i,2}, e_{i,2}) \tag{167}$$

$$= \phi_{22} + \omega_2 + 2Cov(F_{i,2}, e_{i,2}) \tag{168}$$

Once again, assuming the error terms and the *true* variables are independent (i.e. $Cov(F_{i,j}, e_{i,k}) = 0$ for $j = 1, 2$ and $k = 1, 2$) (we are not provided this), then $Var(D_{i,2}) = \phi_{22} + \omega_2$. Now, we can finally derive the correlation of the observed variables,

$$Corr(D_{i,1}, D_{i,2}) = \frac{Cov(D_{i,1}, D_{i,2})}{\sqrt{Var(D_{i,1})Var(D_{i,2})}} \tag{169}$$

$$= \frac{\phi_{12}}{\sqrt{(\phi_{11} + \omega_1)(\phi_{22} + \omega_2)}} \tag{170}$$

Since $\omega_j \geq 0$ for $j = 1, 2$, $\phi_{jj} + \omega_j \geq \phi_{jj}$ and so $\frac{1}{\sqrt{(\phi_{11}+\omega_1)(\phi_{22}+\omega_2)}} \leq \frac{1}{\sqrt{\phi_{11}\phi_{22}}}$. Thus, we have $|Corr(D_{i,1}, D_{i,2})| \leq |Corr(F_{i,1}, F_{i,2})|$, that is, measurement error weakens (attenuates) the correlation, as required.

(c) We know $\rho_j^2 = \frac{\phi_{jj}}{\phi_{jj}+\omega_j}$ for $j = 1, 2$. So, we divide $Corr(D_{i,1}, D_{i,2})$ by the square root of both reliabilities to get $Corr(F_{i,1}, F_{i,2})$,

$$\frac{Corr(D_{i,1}, D_{i,2})}{\sqrt{\rho_1^2\rho_2^2}} = \frac{\phi_{12}}{\sqrt{(\phi_{11} + \omega_1)(\phi_{22} + \omega_2)\frac{\phi_{11}}{\phi_{11}+\omega_1}\frac{\phi_{22}}{\phi_{22}+\omega_2}}} \tag{171}$$

$$= \frac{\phi_{12}}{\sqrt{\phi_{11}\phi_{22}}} \tag{172}$$

$$= Corr(F_{i,1}, F_{i,2}) \tag{173}$$

So, in conclusion, $\frac{Corr(D_{i,1},D_{i,2})}{\sqrt{\rho_1^2\rho_2^2}} = Corr(F_{i,1}, F_{i,2})$.

(d) We compute $\frac{0.25}{\sqrt{0.90*0.75}}$. Plugging this into R console (or a calculator), we have an estimate of the correlation between true intelligence and true self-esteem of approximately **0.30**.

7

7.1 Problem

This is a simplified version of the situation where one is attempting to "control" for explanatory variables that are measured with error. People do this all the time, and it doesn't work. Independently for $i = 1, ..., n$, let

$$Y_i = \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i \tag{174}$$
$$W_i = X_{i,1} + e_i, \tag{175}$$

where $V\begin{pmatrix} X_{i,1} \\ X_{i,2} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{pmatrix}$, $V(\epsilon_i) = \psi$, $V(e_1) = \omega$, all expected values are zero, and the error terms ϵ_i and e_i are independent of one another, and also independent of $X_{i,1}$ and $X_{i,2}$. The variable $X_{i,1}$ is latent, while the variables W_i, Y_i and $X_{i,2}$ are observable. What people usually do in situations like this is fit a model like $Y_i = \beta_1 W_i + \beta_2 X_{i,2} + \epsilon_i$, and test $H_0 : \beta_2 = 0$. That is, they ignore the measurement error in variables for which they are "controlling." ("controlling". ?)

(a) Suppose $H_0 : \beta_2 = 0$ is true. Does the ordinary least squares estimator

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n W_i^2 \sum_{i=1}^n X_{i,2} Y_i - \sum_{i=1}^n W_i X_{i,2} \sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i^2 \sum_{i=1}^n X_{i,2}^2 - (\sum_{i=1}^n W_i X_{i,2})^2} \tag{176}$$

converge to the true value of $\beta_2 = 0$ as $n \to \infty$ everywhere in the parameter space? Answer Yes or No and show your work.

(b) Under what conditions (that is, for what values of other parameters) does $\hat{\beta}_2 \xrightarrow{p} 0$ when $\beta_2 = 0$?

7.2 Solution

(a) To solve this question, we need to first derive the relevant population moments. Beginning with $E(W_i^2)$,

$$\begin{aligned}
E(W_i^2) &= E[(X_{i,1} + e_i)^2] && (177)\\
&= E(X_{i,1}^2 + 2X_{i,1}e_i + e_i^2) && (178)\\
&= E(X_{i,1}^2) + 2E(X_{i,1}e_i) + E(e_i^2) && (179)\\
&= Var(X_{i,1}) + E(X_{i,1})^2 + 2E(X_{i,1}e_i) + E(e_i^2) && (180)\\
&= \phi_{11} + 0 + 2Cov(X_{i,1}, e_i) + Var(e_i) + E(e_i)^2 && (181)\\
&= \phi_{11} + 0 + \omega + 0 && (182)\\
&= \phi_{11} + \omega && (183)
\end{aligned}$$

Next, $E(X_{i,2}Y_i)$,

$$\begin{aligned}
E(X_{i,2}Y_i) &= E(X_{i,2}(\beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i)) && (184)\\
&= \beta_1 E(X_{i,2}X_{i,1}) + \beta_2 E(X_{i,2}^2) + E(X_{i,2}\epsilon_i) && (185)\\
&= \beta_1 Cov(X_{i,2}, X_{i,1}) + \beta_2 Var(X_{i,2}^2) + Cov(X_{i,2}, \epsilon_i) && (186)\\
&= \beta_1\phi_{12} + \beta_2\phi_{22} + 0 && (187)\\
&= \beta_1\phi_{12} + \beta_2\phi_{22} && (188)
\end{aligned}$$

Then, $E(W_iX_{i,2})$,

$$\begin{aligned}
E(W_iX_{i,2}) &= E((X_{i,1} + e_i)X_{i,2}) && (189)\\
&= E(X_{i,1}X_{i,2}) + E(e_iX_{i,2}) && (190)\\
&= Cov(X_{i,1}, X_{i,2}) + Cov(e_i, X_{i,2}) && (191)\\
&= \phi_{12} + 0 && (192)\\
&= \phi_{12} && (193)
\end{aligned}$$

Now, $E(W_iY_i)$,

$$\begin{aligned}
E(W_iY_i) &= E((X_{i,1} + e_i)(\beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i)) && (194)\\
&= \beta_1 E(X_{i,1}^2) + \beta_2 E(X_{i,1}X_{i,2}) + E(X_{i,1}\epsilon_i) + E(e_i(\beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i)) && (195)\\
&= \beta_1 Var(X_{i,1}) + \beta_2 Cov(X_{i,1}, X_{i,2}) + Cov(X_{i,1}, \epsilon_i) + E(e_i(\beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i)) && (196)
\end{aligned}$$

$$= \beta_1\phi_{11} + \beta_2\phi_{12} + 0 + \beta_1 E(e_i X_{i,1}) + \beta_2 E(e_i X_{i,2}) + E(e_i\epsilon_i) \tag{197}$$

$$= \beta_1\phi_{11} + \beta_2\phi_{12} + \beta_1 Cov(e_i, X_{i,1}) + \beta_2 Cov(e_i, X_{i,2}) + Cov(e_i, \epsilon_i) \tag{198}$$

$$= \beta_1\phi_{11} + \beta_2\phi_{12} + 0 + 0 + 0 \tag{199}$$

$$= \beta_1\phi_{11} + \beta_2\phi_{12} \tag{200}$$

With this our preparatory population computations are complete. Now, by SLLN and function continuity (or continuous mapping?),

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n W_i^2 \sum_{i=1}^n X_{i,2}Y_i - \sum_{i=1}^n W_i X_{i,2} \sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i^2 \sum_{i=1}^n X_{i,2}^2 - (\sum_{i=1}^n W_i X_{i,2})^2} \tag{201}$$

$$= \frac{\sum_{i=1}^n W_i^2 \sum_{i=1}^n X_{i,2}Y_i - \sum_{i=1}^n W_i X_{i,2} \sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i^2 \sum_{i=1}^n X_{i,2}^2 - (\sum_{i=1}^n W_i X_{i,2})^2} \frac{\frac{1}{n^2}}{\frac{1}{n^2}} \tag{202}$$

$$= \frac{(\frac{1}{n}\sum_{i=1}^n W_i^2)(\frac{1}{n}\sum_{i=1}^n X_{i,2}Y_i) - (\frac{1}{n}\sum_{i=1}^n W_i X_{i,2})(\frac{1}{n}\sum_{i=1}^n W_i Y_i)}{(\frac{1}{n}\sum_{i=1}^n W_i^2)(\frac{1}{n}\sum_{i=1}^n X_{i,2}^2) - (\frac{1}{n}\sum_{i=1}^n W_i X_{i,2})^2} \tag{203}$$

$$\xrightarrow{a.s.} \frac{(E(W_i^2))(E(X_{i,2}Y_i)) - (E(W_i X_{i,2}))(E(W_i Y_i))}{(E(W_i^2))(E(X_{i,2}^2)) - (E(W_i X_{i,2}))^2} \tag{204}$$

$$= \frac{(\phi_{11} + \omega)(\beta_1\phi_{12} + \beta_2\phi_{22}) - (\phi_{12})(\beta_1\phi_{11} + \beta_2\phi_{12})}{(\phi_{11} + \omega)(\phi_{22}) - (\phi_{12})^2} \tag{205}$$

$$= \frac{\beta_1\phi_{11}\phi_{12} + \beta_2\phi_{11}\phi_{22} + \omega\beta_1\phi_{12} + \omega\beta_2\phi_{22} - \beta_1\phi_{11}\phi_{12} - \beta_2\phi_{12}^2}{(\phi_{11} + \omega)(\phi_{22}) - (\phi_{12})^2} \tag{206}$$

$$= \frac{\beta_2\phi_{11}\phi_{22} + \omega\beta_1\phi_{12} + \omega\beta_2\phi_{22} - \beta_2\phi_{12}^2}{\phi_{11}\phi_{22} + \omega\phi_{22} - \phi_{12}^2} \tag{207}$$

$$= \frac{\beta_2(\phi_{11}\phi_{22} + \omega\phi_{22} - \phi_{12}^2) + \omega\beta_1\phi_{12}}{\phi_{11}\phi_{22} + \omega\phi_{22} - \phi_{12}^2} \tag{208}$$

$$= \beta_2 + \frac{\omega\beta_1\phi_{12}}{\phi_{11}\phi_{22} + \omega\phi_{22} - \phi_{12}^2} \tag{209}$$

$$\neq \beta_2 \tag{210}$$

We note that $\hat{\beta}_2$ is not a consistent estimator of β_2. Thus, supposing $H_0 : \beta_2 = 0$ is true, **No**, the ordinary least squares estimator, $\hat{\beta}_2$, does not converge to the true value of $\beta_2 = 0$ as $n \to \infty$ everywhere in the parameter space (it converges to $\frac{\omega\beta_1\phi_{12}}{\phi_{11}\phi_{22}+\omega\phi_{22}-\phi_{12}^2}$ supposing $H_0 : \beta_2 = 0$ is true).

(b) Since $\hat{\beta}_2$ converges to $\frac{\omega\beta_1\phi_{12}}{\phi_{11}\phi_{22}+\omega\phi_{22}-\phi_{12}^2}$ supposing $H_0 : \beta_2 = 0$ is true, $\hat{\beta}_2 \xrightarrow{p} 0$ when $\beta_2 = 0$ provided that we have any of the following conditions

i. $\omega = 0$ or the measurement error in the variable being "controlled" has no variance (i.e. there is no measurement error in W_i).

ii. $\beta_1 = 0$ or there is no relation between variable being "controlled" and the response variable.

iii. $\phi_{12} = 0$ or there is no correlation (or covariance) between the explanatory variables $X_{i,1}$ and $X_{i,2}$.

8

8.1 Problem

Finally we have a solution, though as usual there is a little twist. Independently for $i = 1, ..., n$, let

$$
\begin{aligned}
Y_i &= \beta X_i + \epsilon_i && (211)\\
V_i &= Y_i + e_i && (212)\\
W_{i,1} &= X_i + e_{i,1} && (213)\\
W_{i,2} &= X_i + e_{i,2} && (214)\\
& && (215)
\end{aligned}
$$

where

- Y_i is a latent variable.
- V_i, $W_{i,1}$ and $W_{i,2}$ are all observable variables.
- X_i is a normally distributed *latent* variable with mean zero and variance $\phi > 0$.
- ϵ_i is normally distributed with mean zero and variance $\psi > 0$.

- e_i is normally distribute with mean zero and variance $\omega > 0$
- $e_{i,1}$ is normally distributed with mean zero and variance $\omega_1 > 0$.
- $e_{i,2}$ is normally distributed with mean zero and variance $\omega_2 > 0$.
- $X_i, \epsilon_i, e_i, e_{i,1}$ and $e_{i,2}$ are all independent of one another.

(a) Make a path diagram of this model.

(b) What is the parameter vector $\boldsymbol{\theta}$ for this model?

(c) Does the model pass the test of the Parameter Count Rule? Answer Yes or No and give the numbers.

(d) Calculate the variance-covariance matrix of the observable variables as a function of the model parameters. Show your work.

(e) Is the parameter vector identifiable at every point in the parameter space? Answer Yes or No and prove your answer.

(f) Some parameters are identifiable, while others are not. Which ones are identifiable?

(g) If β (the paramter (parameter?) of main interest) is identifiable, propose a Method of Moments estimator for it and prove that your proposed estimator is consistent.

(h) Suppose the sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ is

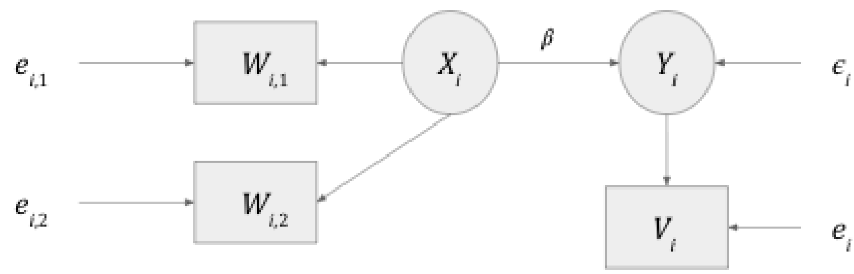
$$\begin{array}{cccc} & W1 & W2 & V \\ W1 & 38.53 & 21.39 & 19.85 \\ W2 & 21.39 & 35.50 & 19.00 \\ V & 19.85 & 19.00 & 28.81 \end{array} \tag{216}$$

Give a reasonable estimate of β. There is more than one right answer. The answer is a number. (Is this the Method of Moments estimate you proposed? It does not have to be.) **Circle you answer.**

(i) Describe how you could re-parameterize this model to make the parameters all identifiable, allowing you (to?) do maximum likelihood.

8.2 Solution

(a) The path diagram for the model is provided below



(b) $\boldsymbol{\theta} = (\beta, \phi, \psi, \omega, \omega_1, \omega_2)$

(c) Since we have mean zero for everything the mean equations are useless. We have 3 observable variables, so we have 6 covariance structure equations. From the previous sub-question, we have 6 parameters. So, **Yes**, model passes the test of the Parameter Count Rule.

(d) We want

$$Cov\begin{pmatrix} V_i \\ W_{i,1} \\ W_{i,2} \end{pmatrix} = \begin{pmatrix} Var(V_i) & Cov(V_i, W_{i,1}) & Cov(V_i, W_{i,2}) \\ Cov(V_i, W_{i,1}) & Var(W_{i,1}) & Cov(W_{i,1}, W_{i,2}) \\ Cov(V_i, W_{i,2}) & Cov(W_{i,1}, W_{i,2}) & Var(W_{i,2}) \end{pmatrix} \tag{217}$$

Starting with $Var(V_i)$,

$$\begin{aligned}
Var(V_i) &= Var(Y_i + e_i) && (218)\\
&= Var(\beta X_i + \epsilon_i + e_i) && (219)\\
&= E[(\beta X_i + \epsilon_i + e_i)^2] && (220)\\
&= E[(\beta X_i + \epsilon_i)^2 + 2(\beta X_i + \epsilon_i)e_i + e_i^2] && (221)\\
&= E[(\beta X_i + \epsilon_i)^2] + 2(\beta E(X_i e_i) + E(\epsilon_i e_i)) + E(e_i^2) && (222)\\
&= E[(\beta X_i + \epsilon_i)^2] + 2(0 + 0) + \omega && (223)\\
&= E(\beta^2 X_i^2 + 2\beta X_i \epsilon_i + \epsilon_i^2) + \omega && (224)\\
&= \beta^2 E(X_i^2) + 2\beta E(X_i \epsilon_i) + E(\epsilon_i^2) + \omega && (225)\\
&= \beta^2 \phi + 0 + \psi + \omega && (226)\\
&= \beta^2 \phi + \psi + \omega && (227)
\end{aligned}$$

Next, $Cov(V_i, W_{i,1})$,

$$
\begin{aligned}
Cov(V_i, W_{i,1}) &= Cov(\beta X_i + \epsilon_i + e_i, X_i + e_{i,1}) && (228)\\
&= Cov(\beta X_i + \epsilon_i + e_i, X_i) + Cov(\beta X_i + \epsilon_i + e_i, e_{i,1}) && (229)\\
&= \beta Cov(X_i, X_i) + Cov(\epsilon_i, X_i) + Cov(e_i, X_i) + Cov(\beta X_i + \epsilon_i + e_i, e_{i,1}) && (230)\\
&= \beta\phi + 0 + 0 + Cov(\beta X_i + \epsilon_i + e_i, e_{i,1}) && (231)\\
&= \beta\phi + \beta Cov(X_i, e_{i,1}) + Cov(\epsilon_i, e_{i,1}) + Cov(e_i, e_{i,1}) && (232)\\
&= \beta\phi + 0 + 0 + 0 && (233)\\
&= \beta\phi && (234)
\end{aligned}
$$

Then, $Cov(V_i, W_{i,2})$,

$$
\begin{aligned}
Cov(V_i, W_{i,2}) &= Cov(\beta X_i + \epsilon_i + e_i, X_i + e_{i,2}) && (235)\\
&= Cov(\beta X_i + \epsilon_i + e_i, X_i) + Cov(\beta X_i + \epsilon_i + e_i, e_{i,2}) && (236)\\
&= \beta Cov(X_i, X_i) + Cov(\epsilon_i, X_i) + Cov(e_i, X_i) + Cov(\beta X_i + \epsilon_i + e_i, e_{i,2}) && (237)\\
&= \beta\phi + 0 + 0 + Cov(\beta X_i + \epsilon_i + e_i, e_{i,2}) && (238)\\
&= \beta\phi + \beta Cov(X_i, e_{i,2}) + Cov(\epsilon_i, e_{i,2}) + Cov(e_i, e_{i,2}) && (239)\\
&= \beta\phi + 0 + 0 + 0 && (240)\\
&= \beta\phi && (241)
\end{aligned}
$$

Now, $Var(W_{i,1})$,

$$
\begin{aligned}
Var(W_{i,1}) &= Var(X_i + e_{i,1}) && (242)\\
&= E[(X_i + e_{i,1})^2] && (243)\\
&= E[X_i^2 + 2X_i e_{i,1} + e_{i,1}^2] && (244)\\
&= E(X_i^2) + 2E(X_i e_{i,1}) + E(e_{i,1}^2) && (245)\\
&= \phi + 0 + \omega_1 && (246)\\
&= \phi + \omega_1 && (247)
\end{aligned}
$$

Moving on, $Cov(W_{i,1}, W_{i,2})$,

$$\begin{aligned}
Cov(W_{i,1}, W_{i,2}) &= Cov(X_i + e_{i,1}, X_i + e_{i,2}) && (248)\\
&= Cov(X_i, X_i) + Cov(X_i, e_{i,2}) + Cov(e_{i,1}, X_i) + Cov(e_{i,1}, e_{i,2}) && (249)\\
&= \phi + 0 + 0 + 0 && (250)\\
&= \phi && (251)
\end{aligned}$$

Finally, $Var(W_{i,2})$,

$$\begin{aligned}
Var(W_{i,2}) &= Var(X_i + e_{i,2}) && (252)\\
&= E[(X_i + e_{i,2})^2] && (253)\\
&= E[X_i^2 + 2X_ie_{i,2} + e_{i,2}^2] && (254)\\
&= E(X_i^2) + 2E(X_ie_{i,2}) + E(e_{i,2}^2) && (255)\\
&= \phi + 0 + \omega_2 && (256)\\
&= \phi + \omega_2 && (257)
\end{aligned}$$

With this, our variance-covariance computations are complete. The matrix obtained is

$$\begin{aligned}
Cov\begin{pmatrix} V_i \\ W_{i,1} \\ W_{i,2} \end{pmatrix} &= \begin{pmatrix} Var(V_i) & Cov(V_i, W_{i,1}) & Cov(V_i, W_{i,2}) \\ Cov(V_i, W_{i,1}) & Var(W_{i,1}) & Cov(W_{i,1}, W_{i,2}) \\ Cov(V_i, W_{i,2}) & Cov(W_{i,1}, W_{i,2}) & Var(W_{i,2}) \end{pmatrix} && (258)\\
&= \begin{pmatrix} \beta^2\phi + \psi + \omega & \beta\phi & \beta\phi \\ \beta\phi & \phi + \omega_1 & \phi \\ \beta\phi & \phi & \phi + \omega_2 \end{pmatrix} && (259)
\end{aligned}$$

(e) Let

$$\begin{aligned}
Cov\begin{pmatrix} V_i \\ W_{i,1} \\ W_{i,2} \end{pmatrix} &= \begin{pmatrix} Var(V_i) & Cov(V_i, W_{i,1}) & Cov(V_i, W_{i,2}) \\ Cov(V_i, W_{i,1}) & Var(W_{i,1}) & Cov(W_{i,1}, W_{i,2}) \\ Cov(V_i, W_{i,2}) & Cov(W_{i,1}, W_{i,2}) & Var(W_{i,2}) \end{pmatrix} && (260)\\
&= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix} && (261)
\end{aligned}$$

Then from our solution to the previous sub-question we have a system of equations,

$$\sigma_{11} = \beta^2\phi + \psi + \omega \tag{262}$$
$$\sigma_{12} = \beta\phi \tag{263}$$
$$\sigma_{13} = \beta\phi \tag{264}$$
$$\sigma_{22} = \phi + \omega_1 \tag{265}$$
$$\sigma_{23} = \phi \tag{266}$$
$$\sigma_{33} = \phi + \omega_2 \tag{267}$$

We can solve for ϕ, ω_1, ω_2 and β as $\phi = \sigma_{23}$, $\omega_1 = \sigma_{22} - \phi = \sigma_{22} - \sigma_{23}$, $\omega_2 = \sigma_{33} - \phi = \sigma_{33} - \sigma_{23}$ and $\beta = \frac{\sigma_{12}}{\phi} = \frac{\sigma_{12}}{\sigma_{23}}$ (although with β, we can also use $\beta = \frac{\sigma_{13}}{\phi} = \frac{\sigma_{13}}{\sigma_{23}}$, and the system of equations enforces the constraint $\sigma_{12} = \sigma_{13}$). Then we are left with one equation ($\sigma_{11} = \beta^2\phi + \psi + \omega$) and two unknowns ($\psi$ and ω). We can't recover unique values for either ψ or ω from this single remaining equation, and in fact any pair of (ψ, ω) values within the line $\sigma_{11} = \beta^2\phi + \psi + \omega$ (note that we substitute in the solutions we have for β and ϕ above to have a line in the (ψ, ω) space) would do (meaning we can't recover either ψ or ω from the data). So, **No**, the parameter vector is not identifiable at every point in the parameter space.

(f) ϕ, ω_1, ω_2 and β are identifiable, as we could solve for them in the previous sub-question. ψ and ω are not identifiable, as we could not solve for them in the previous sub-question.

(g) One Method of Moments estimator (solution is not unique) is $\tilde{\beta} = \frac{\hat{\sigma_{12}}}{\hat{\sigma_{23}}} = \frac{\sum_{i=1}^{n}(V_i - \bar{V})(W_{i,1} - \bar{W}_1)}{\sum_{i=1}^{n}(W_{i,1} - \bar{W}_1)(W_{i,2} - \bar{W}_2)}$. Let's show this is consistent. Assuming the consistency of the sample covariance and variance moments (as we did in question 1, or apply SLLN and function continuity (do we need to assume anything?)),

$$\tilde{\beta} = \frac{\hat{\sigma_{12}}}{\hat{\sigma_{23}}} \tag{268}$$
$$= \frac{\sum_{i=1}^{n}(V_i - \bar{V})(W_{i,1} - \bar{W}_1)}{\sum_{i=1}^{n}(W_{i,1} - \bar{W}_1)(W_{i,2} - \bar{W}_2)} \tag{269}$$
$$= \frac{\frac{1}{n}\sum_{i=1}^{n}(V_i - \bar{V})(W_{i,1} - \bar{W}_1)}{\frac{1}{n}\sum_{i=1}^{n}(W_{i,1} - \bar{W}_1)(W_{i,2} - \bar{W}_2)} \tag{270}$$

$$\xrightarrow{a.s.} \frac{E[(V_i - E[V_i])(W_{i,1} - E[W_{i,1}])]}{E[(W_{i,1} - E[W_{i,1}])(W_{i,2} - E[W_{i,2}])]} \tag{271}$$

$$= \frac{Cov(V_i, W_{i,1})}{Cov(W_{i,1}, W_{i,2})} \tag{272}$$

$$= \frac{\beta\phi}{\phi} \tag{273}$$

$$= \beta \tag{274}$$

(h) From the previous sub-question we have $\tilde{\beta} = \frac{\hat{\sigma_{12}}}{\hat{\sigma_{23}}}$. Plugging in, we have $\tilde{\beta} = \frac{19.85}{21.39} \approx 0.93$.

(i) Idea is to combine the two unidentifiable parameters into a single one, i.e. $\psi + \omega = \gamma$, which is akin to performing reparameterization:

$$V_i = Y_i + e_i \tag{275}$$

$$= \beta X_i + \epsilon_i + e_i \tag{276}$$

$$= \beta X_i + \epsilon_i' \tag{277}$$

Clearly, $\epsilon_i' \sim N(0, Var(\epsilon_i + e_i)) = N(0, \psi + \omega) = N(0, \gamma)$, and $X_i, \epsilon_i', e_{i,1}$ and $e_{i,2}$ are all independent of one another. Now our covariance structure equations look like

$$\sigma_{11} = \beta^2\phi + \gamma \tag{278}$$

$$\sigma_{12} = \beta\phi \tag{279}$$

$$\sigma_{13} = \beta\phi \tag{280}$$

$$\sigma_{22} = \phi + \omega_1 \tag{281}$$

$$\sigma_{23} = \phi \tag{282}$$

$$\sigma_{33} = \phi + \omega_2 \tag{283}$$

which solves to give

$$\phi = \sigma_{23} \tag{284}$$

$$\omega_1 = \sigma_{22} - \sigma_{23} \tag{285}$$

$$\omega_2 = \sigma_{33} - \sigma_{23} \tag{286}$$

$$\beta = \frac{\sigma_{12}}{\sigma_{23}} \tag{287}$$

$$\gamma = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{23}} \tag{288}$$

making all parameters identifiable as required (note that the system of equations enforces the constraint $\sigma_{12} = \sigma_{13}$).