Ignoring Measurement Error in Regression[1]

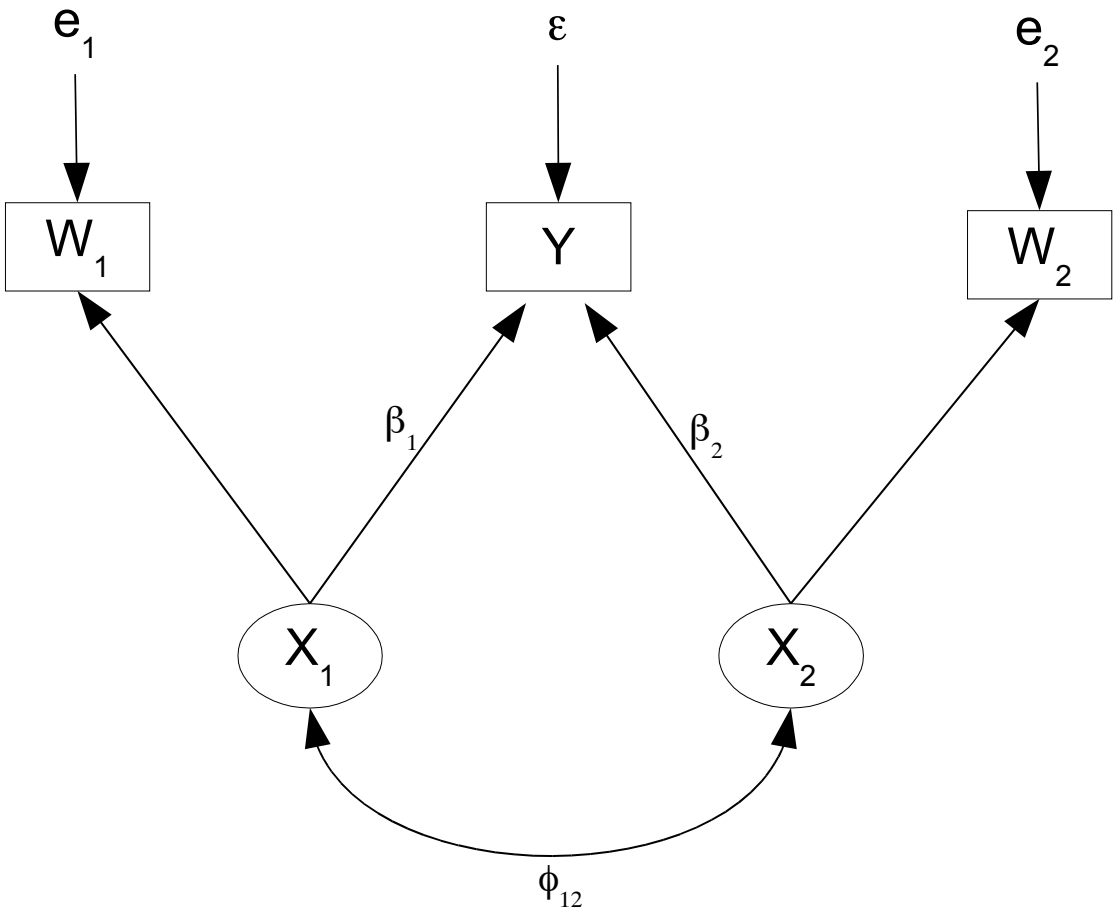
STA2053 Fall 2022

[1] See last slide for copyright information.

Measurement Error

- Snack food consumption
- Exercise
- Income
- Cause of death
- Even amount of drug that reaches animal's blood stream in an experimental study
- Is there anything that is *not* measured with error?

Measurement error in two explanatory variables



Want to assess the relationship of X_2 to Y controlling for X_1 by testing $H_0 : \beta_2 = 0$.

Statement of the model

Independently for $i = 1, \ldots, n$

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i \\
W_{i,1} &= X_{i,1} + e_{i,1} \\
W_{i,2} &= X_{i,2} + e_{i,2},
\end{aligned}
$$

where

$E(X_{i,1}) = \mu_1$, $E(X_{i,2}) = \mu_2$, $E(\epsilon_i) = E(e_{i,1}) = E(e_{i,2}) = 0$,

$Var(\epsilon_i) = \psi$, $Var(e_{i,1}) = \omega_1$, $Var(e_{i,2}) = \omega_2$,

The errors $\epsilon_i, e_{i,1}$ and $e_{i,2}$ are all independent,

$X_{i,1}$ and $X_{i,2}$ are independent of $\epsilon_i, e_{i,1}$ and $e_{i,2}$, and

$$
cov \begin{pmatrix} X_{i,1} \\ X_{i,1} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{pmatrix}.
$$

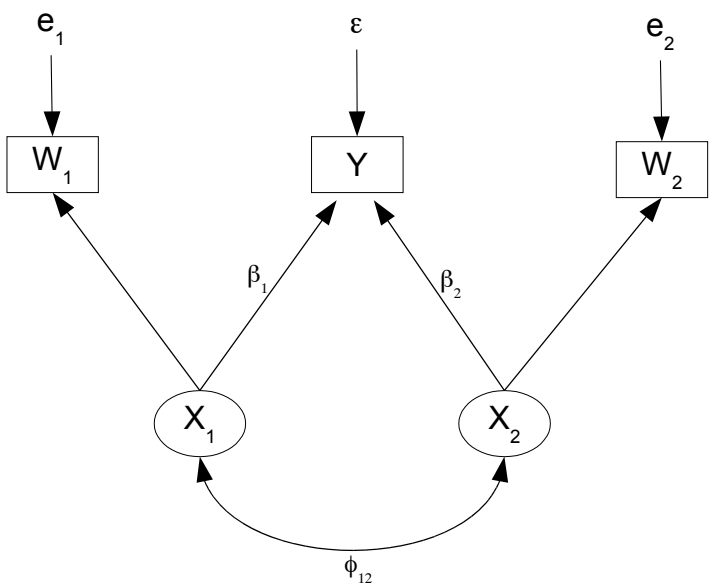
True Model versus Naive Model

Independently for $i = 1, \ldots, n$

True model:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i \\ W_{i,1} &= X_{i,1} + e_{i,1} \\ W_{i,2} &= X_{i,2} + e_{i,2}, \end{aligned}$$

Naive model: $Y_i = \beta_0 + \beta_1 W_{i,1} + \beta_2 W_{i,2} + \epsilon_i$



Copyright Information