1 Measurement error, e, in X

$$y = \beta X + u,\ X^{*} = X + e$$

$$\begin{aligned}
\widehat{\beta}_{OLS} &= (X^{*\prime}X^{*})^{-1}X^{*\prime}y = \\
&= (X^{*\prime}X^{*})^{-1}X^{*\prime}(\beta X + u) = \\
&= (X^{*\prime}X^{*})^{-1}\left[\beta(X^{*\prime}X) + X^{*\prime}u\right] = \\
&= (X^{*\prime}X^{*})^{-1}[\beta(X^{*\prime}X^{*} + X^{*\prime}e) + X^{*\prime}u] = \\
&= \beta + \beta(X^{*\prime}X^{*})^{-1}X^{*\prime}e + (X^{*\prime}X^{*})^{-1}X^{*\prime}u = \qquad (1)\\
&= \beta + \beta(X^{*\prime}X^{*})^{-1}(X'e + e'e) + (X^{*\prime}X^{*})^{-1}(X'u + e'u) = \\
p\lim\widehat{\beta}_{OLS} &= \beta + \beta\frac{VAR(e) + COV(X,e)}{VAR(X^{*})} + \frac{COV(X,u) + COV(e,u)}{VAR(X^{*})} = \\
&= \beta + \beta\frac{VAR(e) + COV(X,e)}{VAR(X) + VAR(e) + 2COV(X,e)} + \frac{COV(X,u) + COV(e,u)}{VAR(X) + VAR(e) + 2COV(X,e)}
\end{aligned}$$

$$\begin{aligned}
X^{*\prime}X &= X^{*\prime}(X^{*} - e) = X^{*\prime}X^{*} - X^{*\prime}e \\
X^{*\prime}e &= (X + e)^{\prime}e = X'e + e'e \\
X^{*\prime}u &= (X + e)^{\prime}u = X'u + e'u \\
X^{*\prime}X^{*} &= (X + e)^{\prime}(X + e) = X'X + e'e + 2X'e
\end{aligned}$$

1.0.1 Pure endogeneity

$VAR(e) = 0; \qquad COV(X,u) \lesseqgtr 0$

$$\begin{aligned}
p\lim\widehat{\beta}_{OLS} &= \beta - \beta\frac{0 + 0}{VAR(X) + 0 + 0} + \frac{COV(X,u) + 0}{VAR(X) + 0 + 0} = \\
&= \beta + \frac{COV(X,u)}{VAR(X)} \lesseqgtr \beta
\end{aligned}$$

Example: $COV(X,u) > 0$ due to ability bias because $X = f(ability)$ and u also contains ability and therefore $COV(X,u) > 0$.

1.0.2 Pure measurement error

$COV(X,u) = 0; \qquad COV(e,u) = 0; \qquad COV(X,e) = 0$

$$\begin{aligned} p\lim \widehat{\beta}_{OLS} &= \beta + \beta \frac{VAR(e) + 0}{VAR(X) + VAR(e) + 2 * 0} + \frac{0 + 0}{VAR(X) + VAR(e) + 2 * 0} = \\ &= \beta + \beta \frac{VAR(e)}{VAR(X) + VAR(e)} = \beta \left[1 - \frac{VAR(e)}{VAR(X) + VAR(e)} \right] < \beta \end{aligned}$$

Noting that

$$0 < \frac{VAR(e)}{VAR(X) + VAR(e)} < 1$$

1.0.3 Systematic measurement error- type 1

$COV(X, u) = 0; \qquad COV(e, u) = 0; \qquad COV(X, e) \lesseqgtr 0$

Measurement error is correlated with actual values of X (example: low educated people tend to overstate years of education).

$$p\lim \widehat{\beta}_{OLS} = \beta + \beta \frac{VAR(e) + COV(X, e)}{VAR(X) + VAR(e) + 2 * COV(X, e)}$$

1.0.4 Systematic measurement error- type 2

$COV(X, u) = 0; \qquad COV(X, e) = 0 \qquad ; COV(e, u) \lesseqgtr 0$

Measurement error e, correlated with unexplained component of LHS variable, u.(example: people who over/under-state education also over/understate earnings).

$$\begin{aligned} p\lim \widehat{\beta}_{OLS} &= \beta + \beta \frac{VAR(e) + 0}{VAR(X) + VAR(e) + 2 * 0} + \frac{0 + COV(e, u)}{VAR(X) + VAR(e) + 2 * 0} = \\ &= \beta + \beta \frac{VAR(e)}{VAR(X) + VAR(e)} + \frac{COV(e, u)}{VAR(X) + VAR(e)} \end{aligned}$$