

POLS/CS&SS 503:
Advanced Quantitative Political Methodology

MEASUREMENT ERROR

May 5, 2015

Jeffrey B. Arnold



Measurement Error (One Variable)

$$Y = \beta_0 + \beta_1 X_1 + \epsilon$$

but estimate

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1^* + \epsilon$$
$$X_1^* = X_1 + \delta$$

- X_1^* is X_1 measured with error.
- Assumptions
 - $E(\delta) = 0$
 - Meas error: $C(\delta, X_1) = 0$. What if measurement error increases with X_1 ?
 - Meas error uncorrelated with regression components: $C(\delta, \epsilon) = 0$, $C(\delta, X_1) = 0$
 - Meas error: $C(\delta, X_1) = 0$
- Reliability: measure of measurement error

$$r = v(X_1)/v(X_1^*) = v(X_1)/(v(X_1^*) + v(\delta))$$

Example of Measurement Error

Population

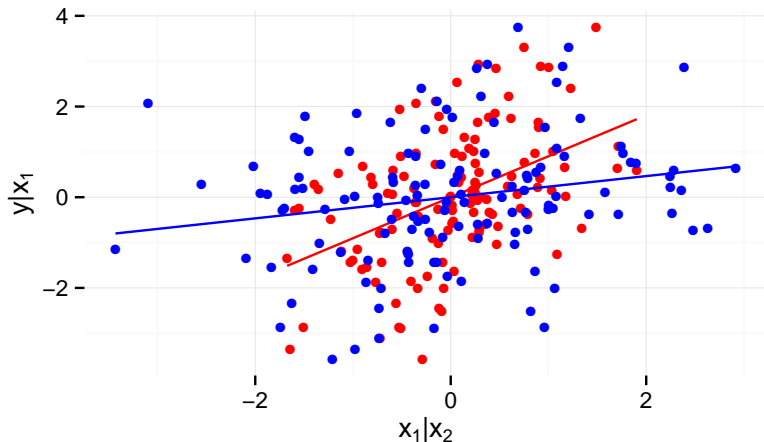
$$Y_i = X_{1,i} + X_{2,i} + \epsilon_i$$
$$X_i^* = X_{1,i} + \delta_i$$

Sample Estimate

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1,i}^* + \hat{\beta}_2 x_{2,i} + \hat{\epsilon}_i$$

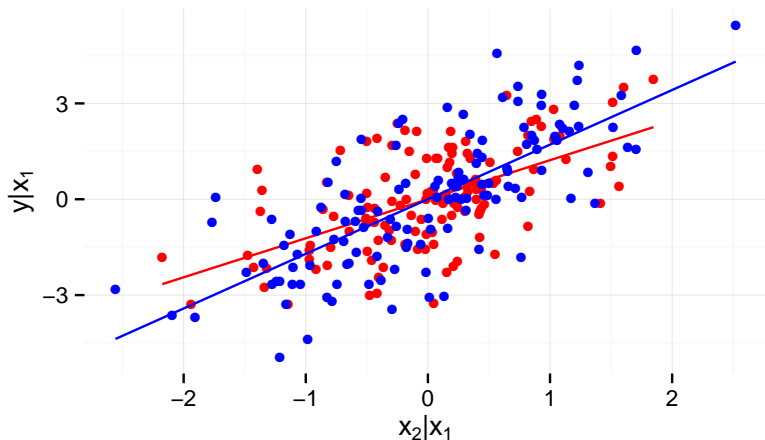
Look at cases in which $r = 0$, no measurement error in X_1^* , and $r = 0.5$, $V(\delta) = V(X_1)$.

Measurement Error, Effect on $\hat{\beta}_1$



Blue is no measurement error, $r = 1$; Red is measurement error, $r = 0.5$.

Measurement Error, Effect on $\hat{\beta}_2$



Blue is no measurement error, $r = 1$; Red is measurement error, $r = 0.5$.

What does measurement error in X do?

- attenuates (biases towards 0) coefficient of covariates with measurement error
- attenuation is **worse** as more covariates are included. Those covariates explain y 's variance, but not the measurement error in x .
- biases coefficients of other regressors towards their values in the regression without that value (omitted variable bias light)

What does measurement error in Y do?

Population

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

$$Y_i^* = Y_{1,i} + \delta_i$$

Then

$$Y_i^* = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + (\epsilon_i + \delta_i)$$

- Error variance of $E(Y|X)$ is larger: $V(\epsilon) + V(\delta)$
- Coefficients of $\hat{\beta}$ unbiased
- Coefficients have larger standard errors:

$$SE(\beta) = \sqrt{\frac{V(\epsilon) + V(\delta)}{(X'X)^{-1}}}$$

What to do about measurement error?

- Get better data or multiple measures
- Multiple imputation. See R package [Amelia](#) and Blackwell, Matthew, James Honaker, and Gary King. 10030. "A Unified Approach to Measurement Error and Missing Data: Overview." *Sociological Methods and Research*.
- Instrumental Variable (IV) models
- Bayesian latent variable models or structural equation models