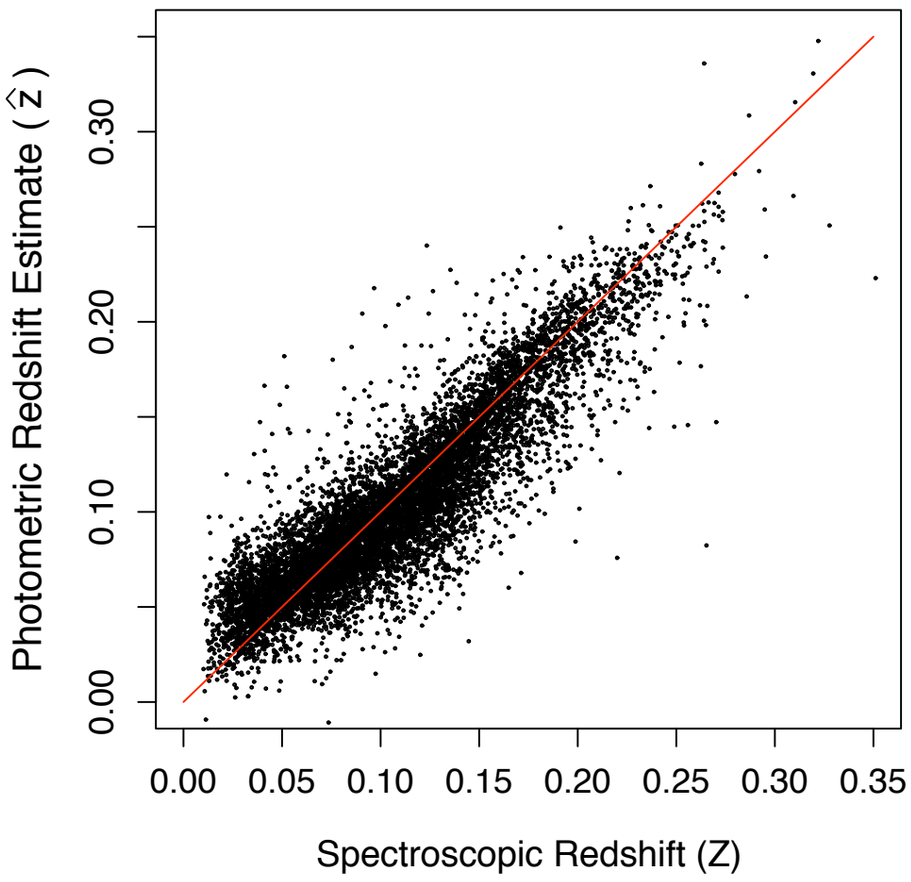
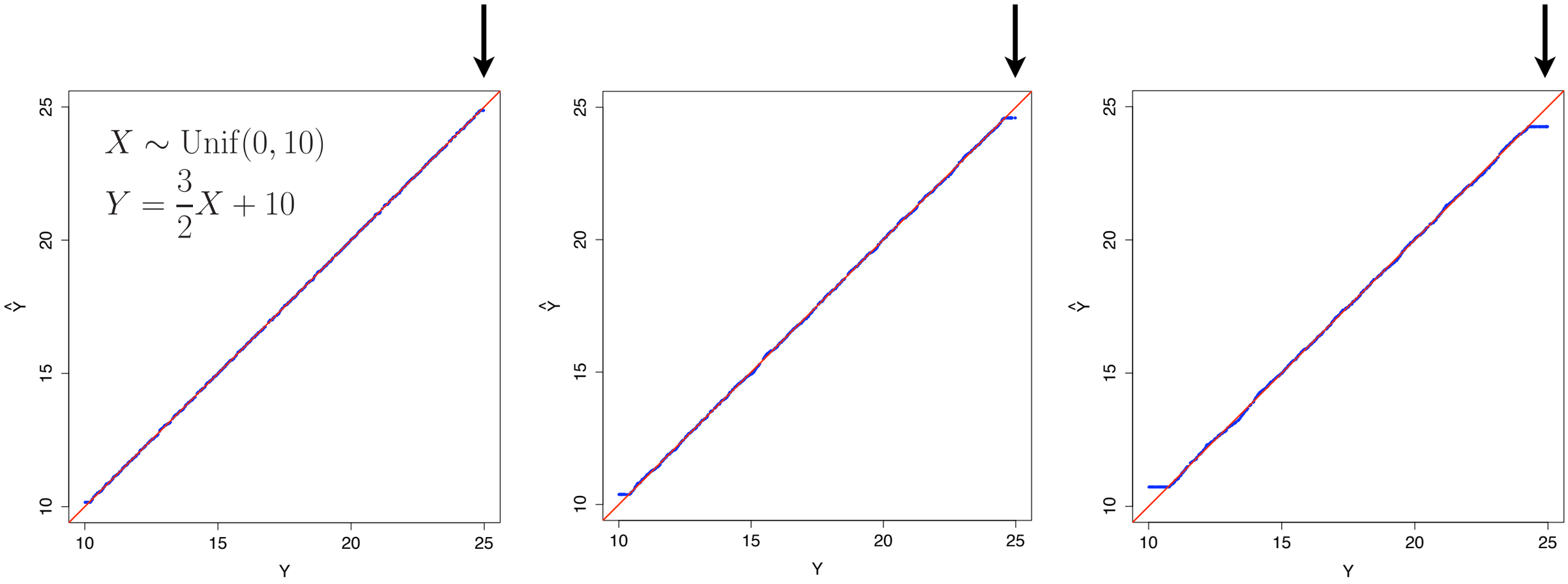
Measurement Error and Estimator Bias

Peter E. Freeman
Department of Statistics - Carnegie Mellon University
(in association with: Ann Lee, Joey Richards, Chad Schafer, and Jeff Newman)



A ten-minute (or so) musing motivated by our work on photometric redshift estimation: arXiv: 0906.0995

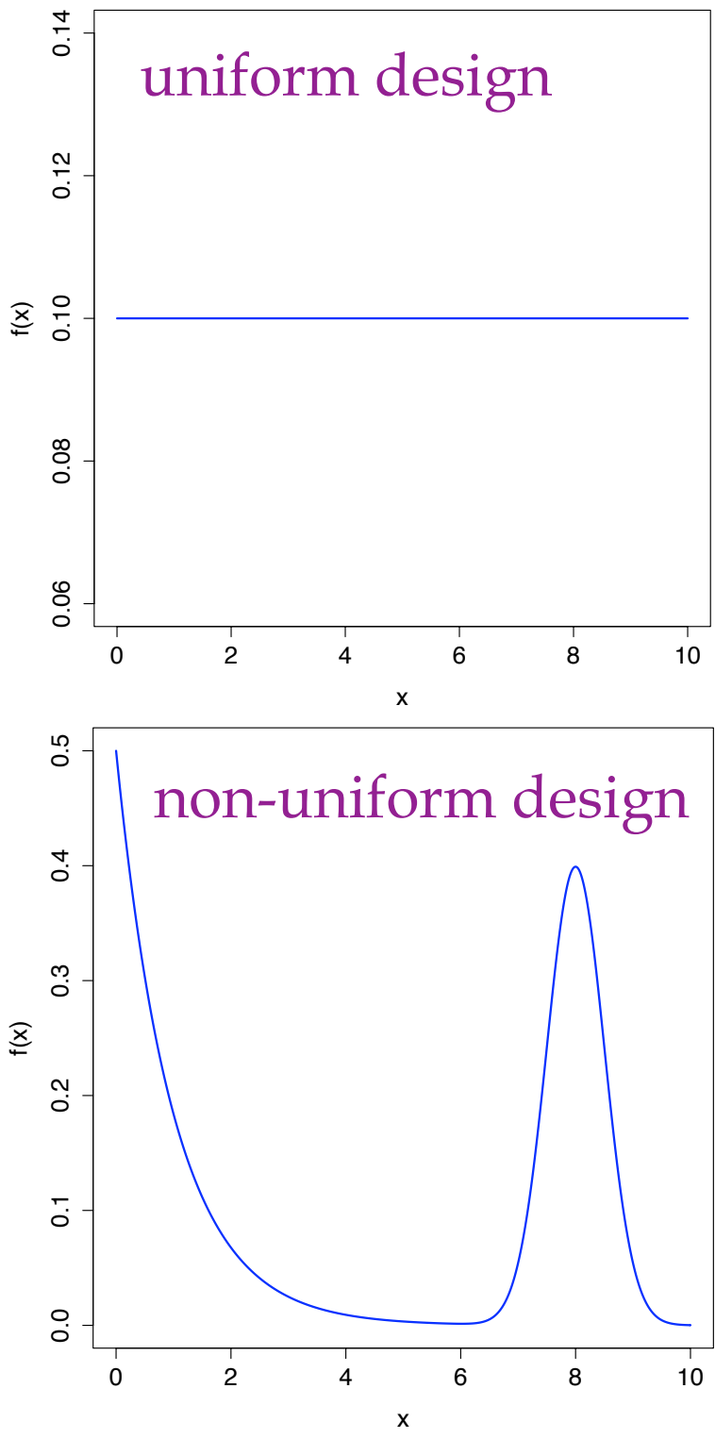
Boundary Bias (kNN example)

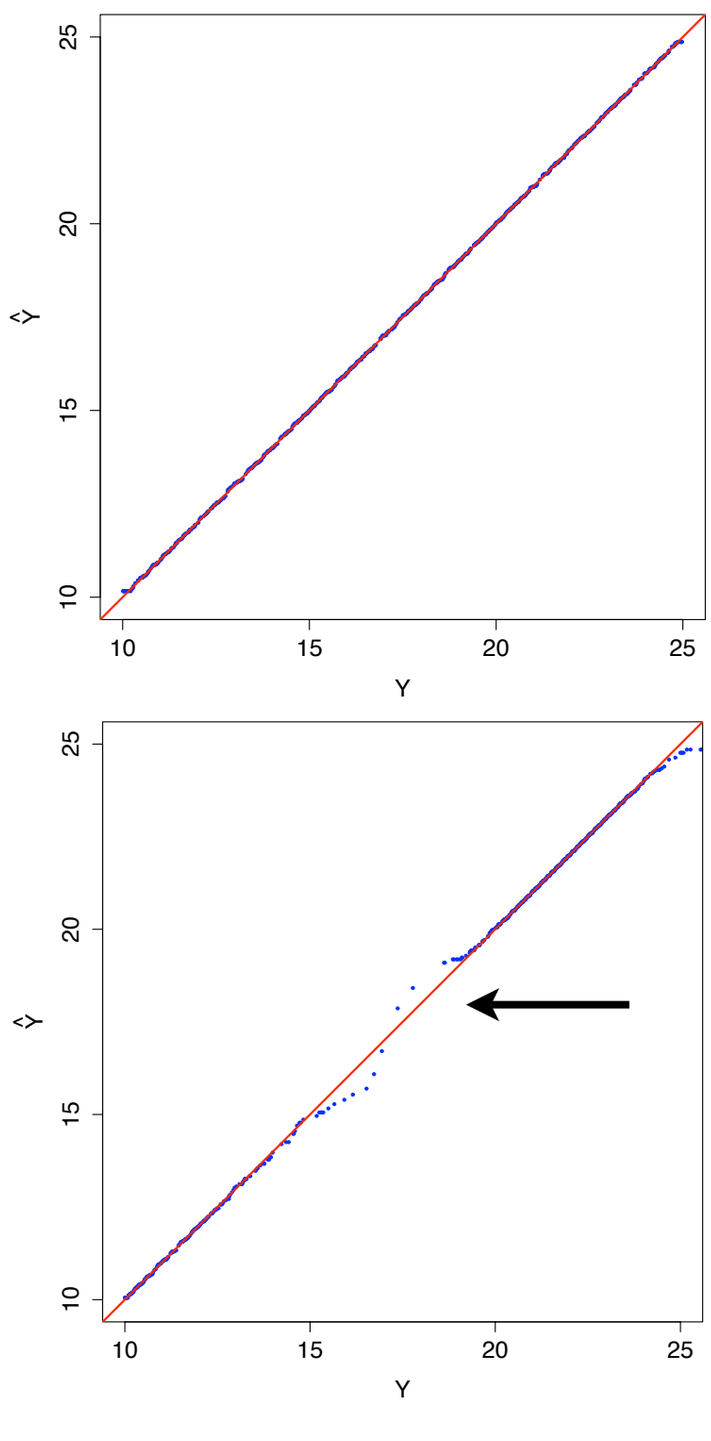


k = 20 k = 50 k = 100

general issue with nonparametric estimators; mitigated via, e.g., local linear regression

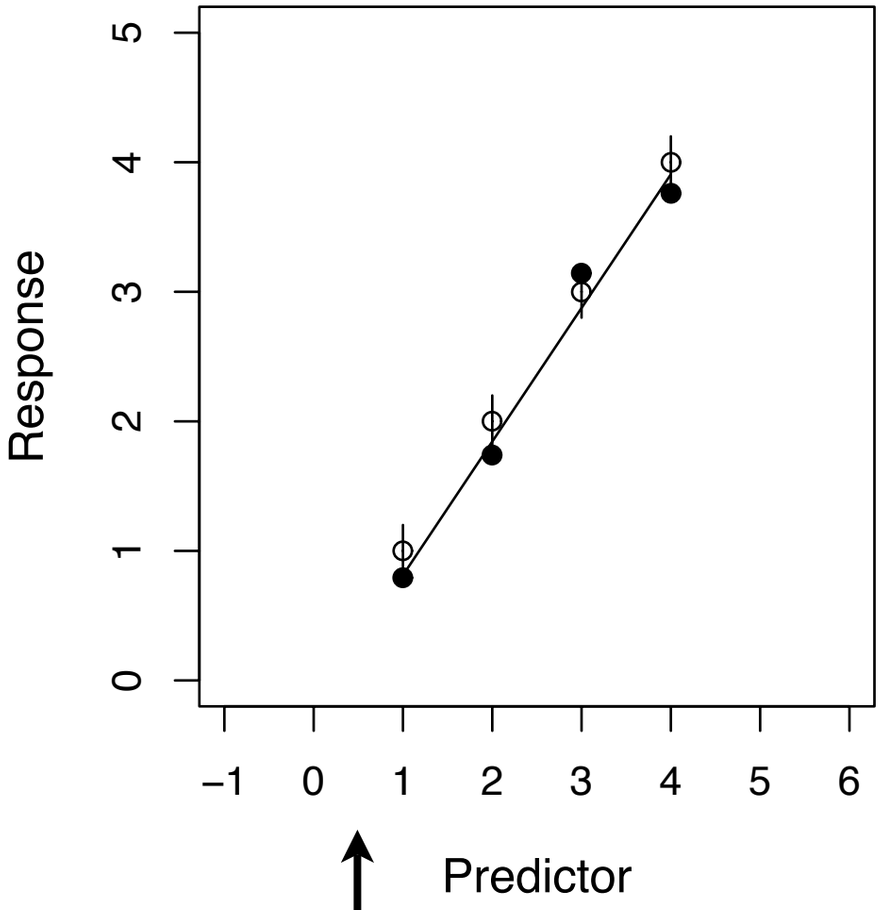
Design Bias (kNN example)





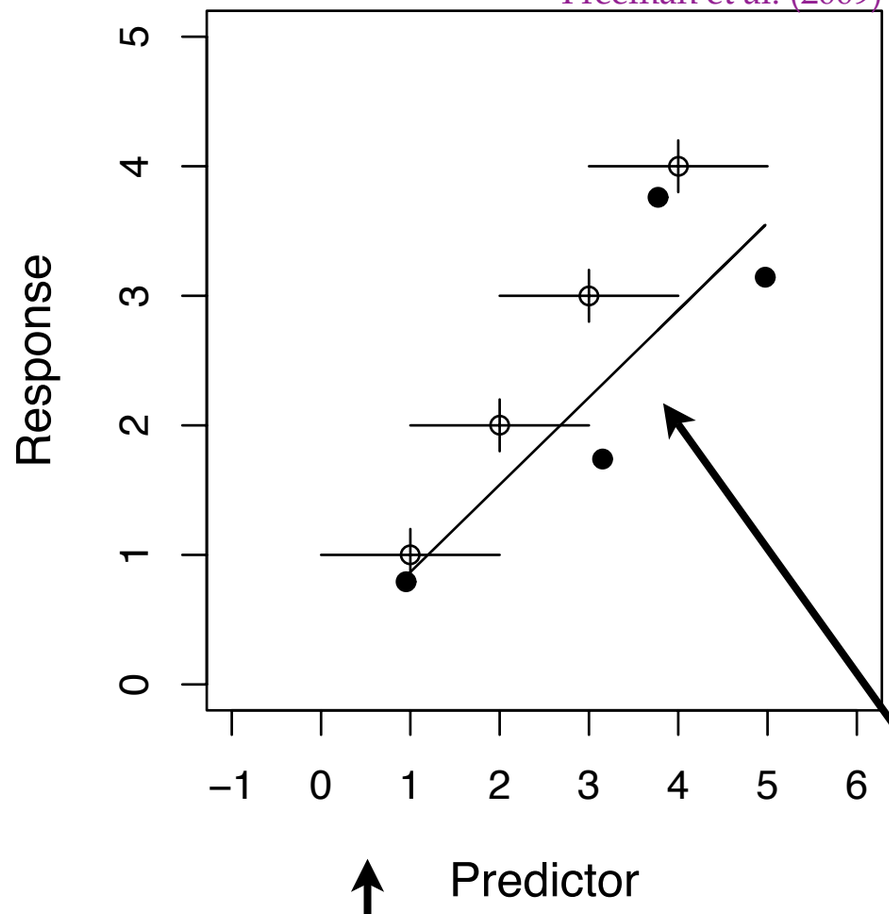
also general issue with nonparametric estimators

Measurement Error (LR example)



Predictor (e.g., flux)
measured without error.

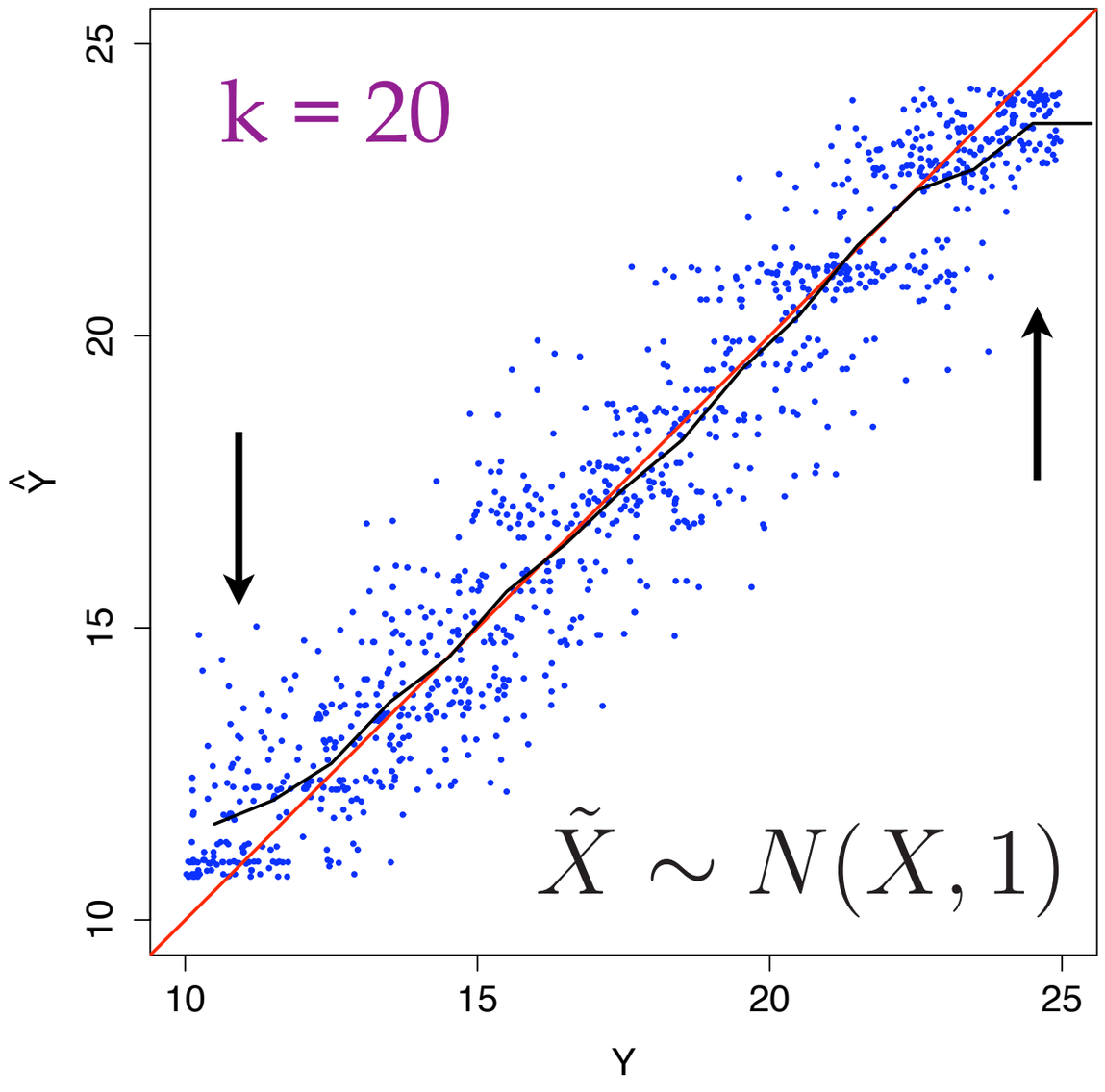
Freeman et al. (2009)

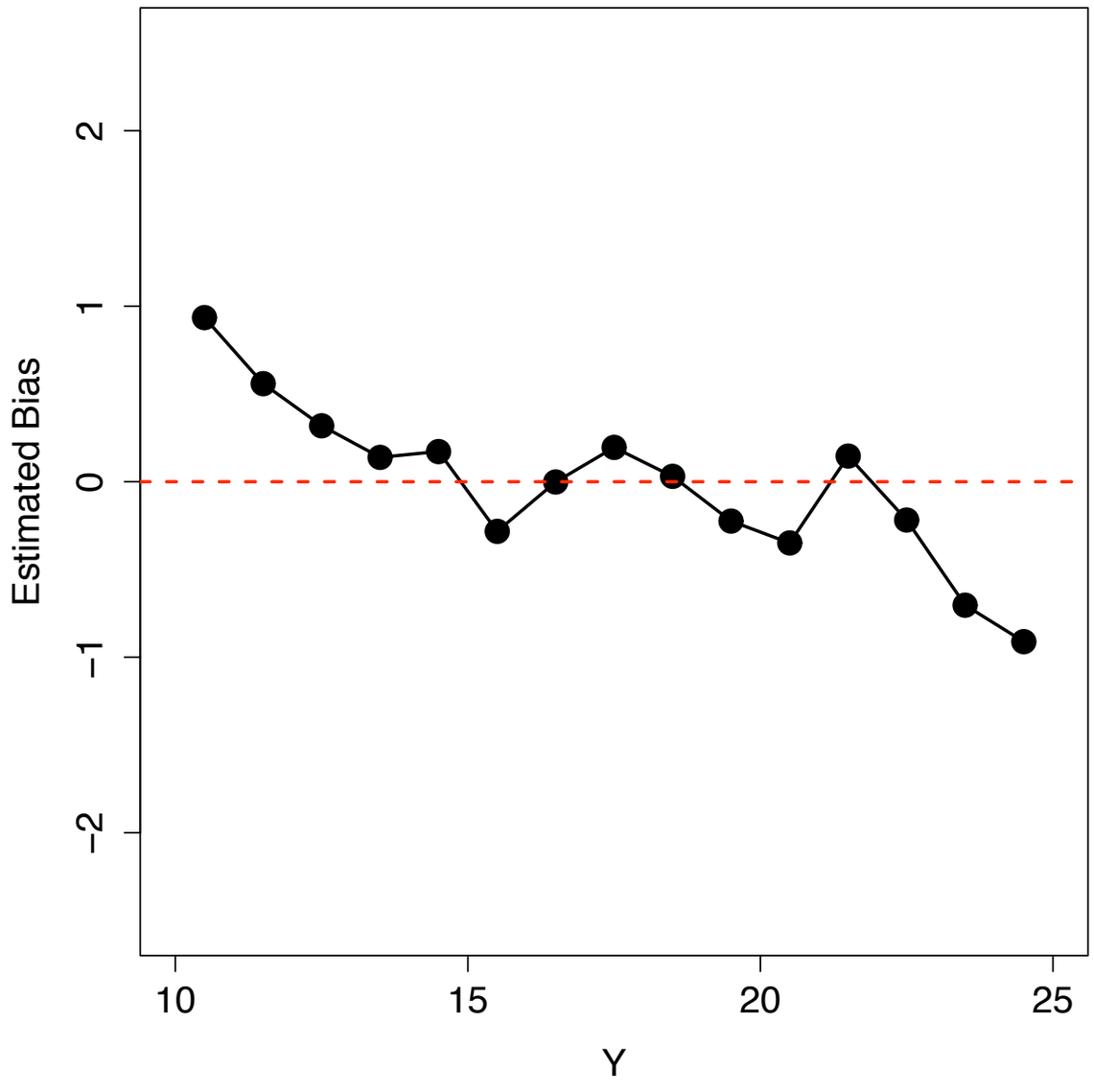


Predictor (e.g., flux)
measured *with* error.

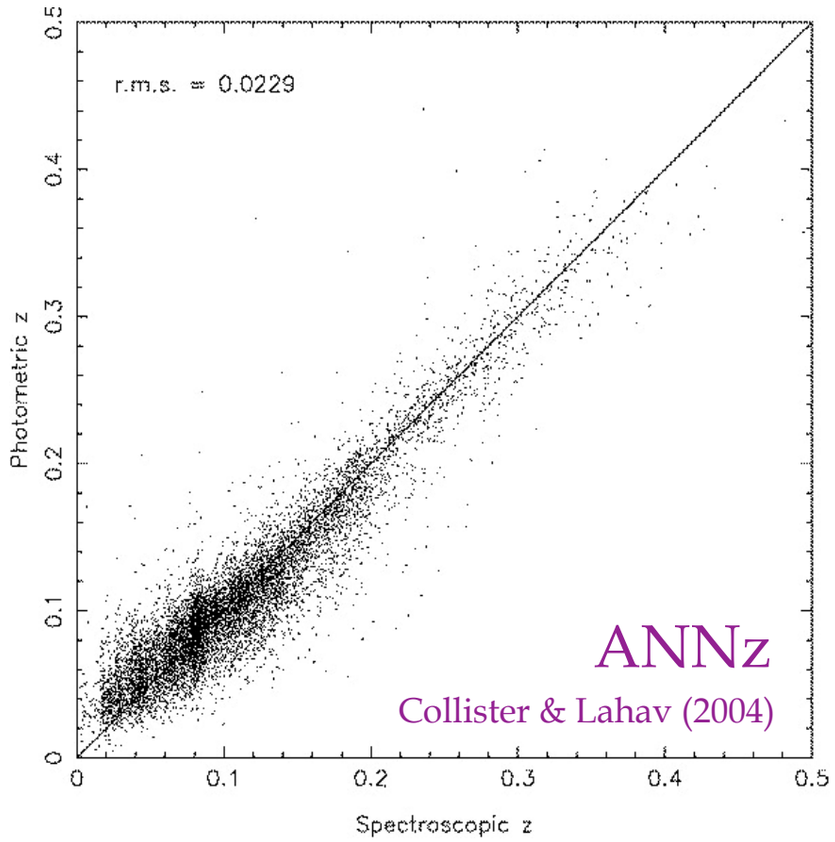
Slope of inferred
regression line
reduced:
attenuation bias.

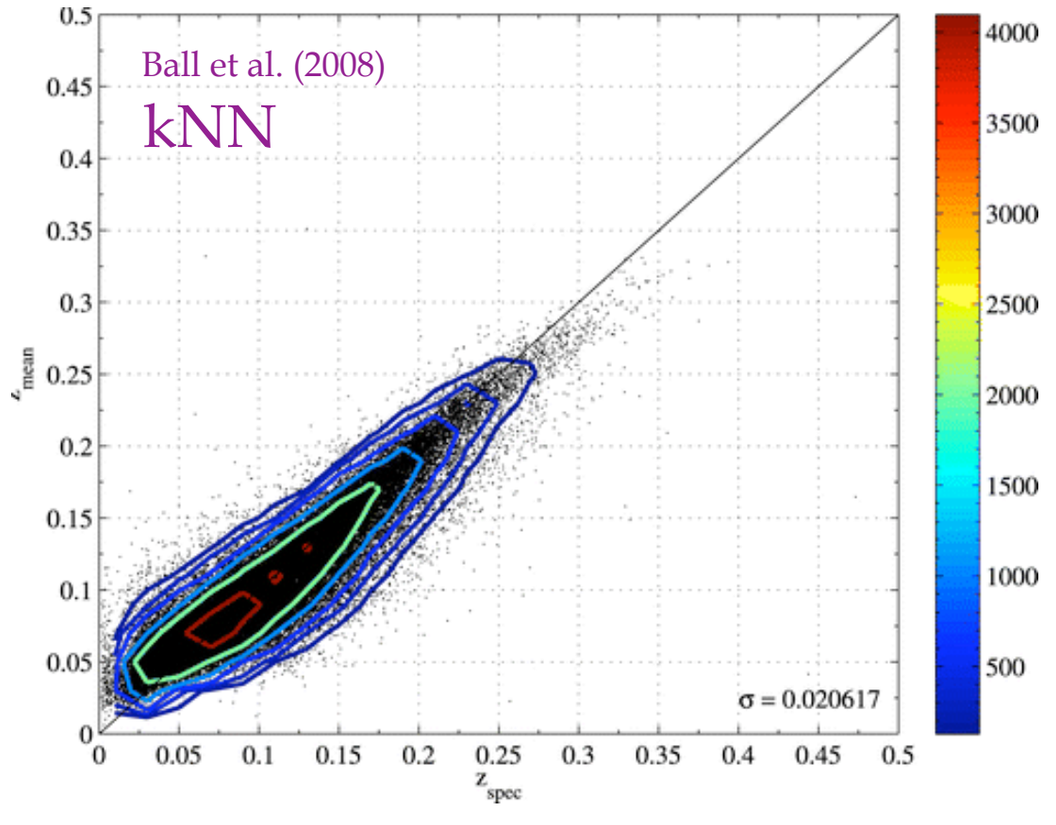
Measurement Error (kNN example)

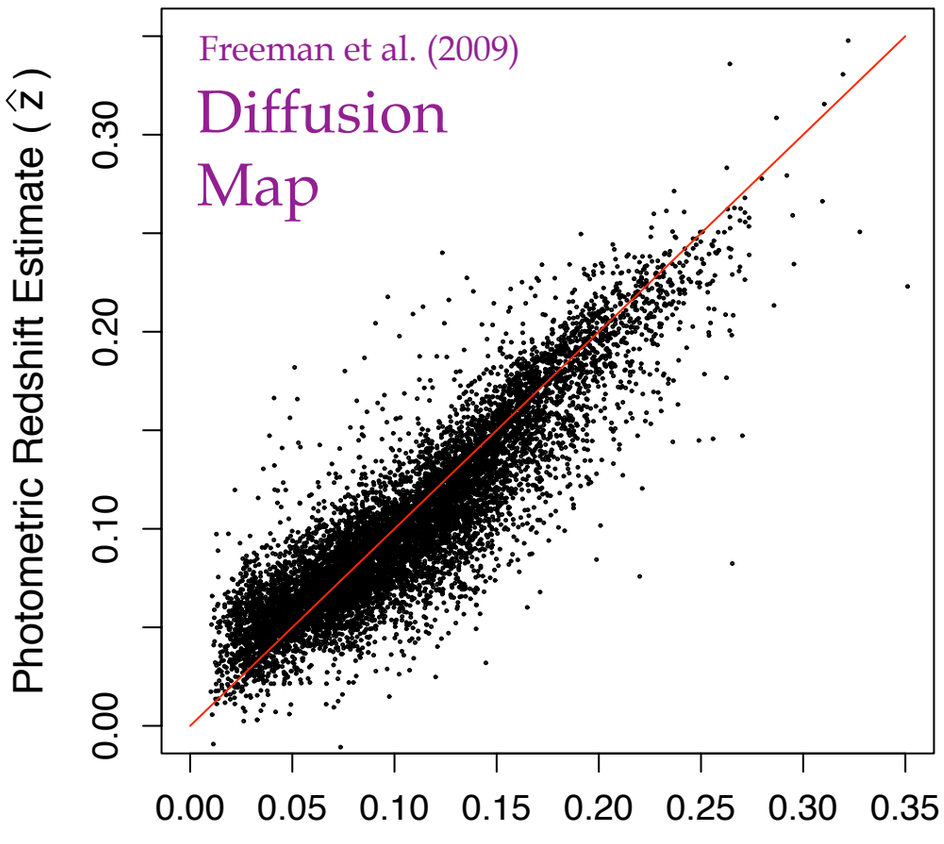


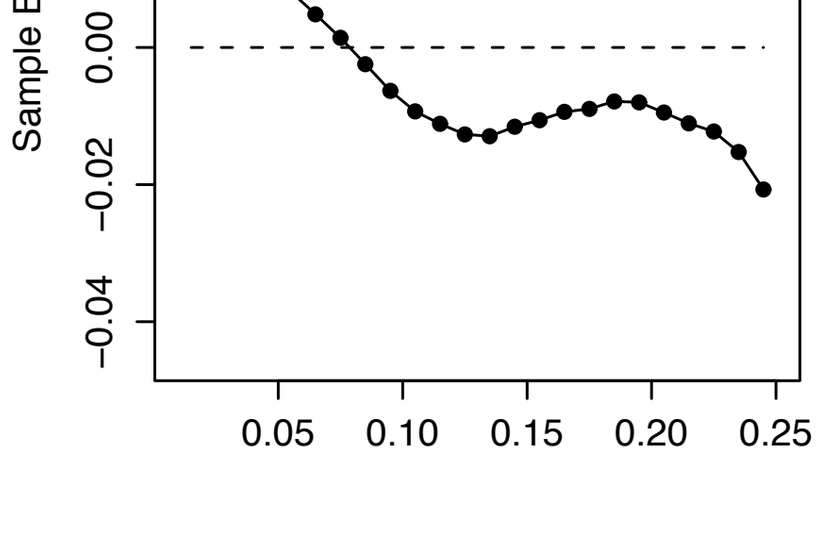


Measurement Error (photo-z)









Take Home Point:

If you desire a (very-)low-bias estimator,
you will have to deal with measurement error!

Simulation-Extrapolation (SIMEX)

(Algorithm for simple linear regression. Similar ones for nonparametric estimators exist.)

Observed Data:

$$\tilde{X} = X + U$$

Observed Slope:

$$\hat{\beta}_1 \to \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2}\beta_1$$

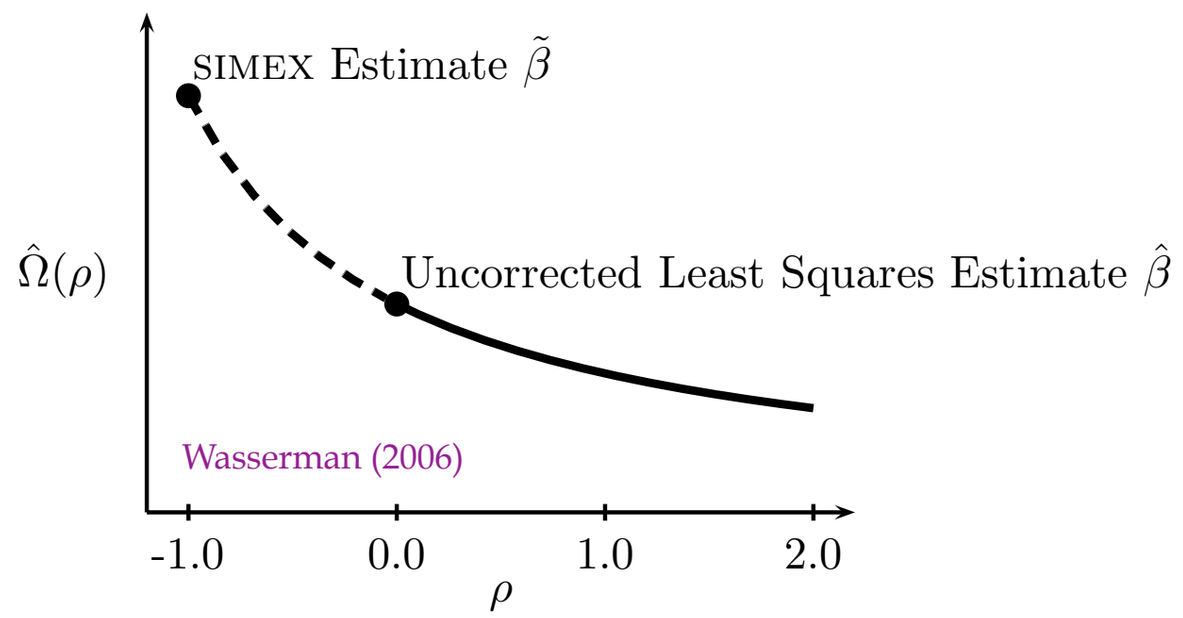
Resample and Fit Data:

$$\tilde{X}' = \tilde{X} + \sqrt{\rho}U \to \hat{\beta}_1(\rho)$$

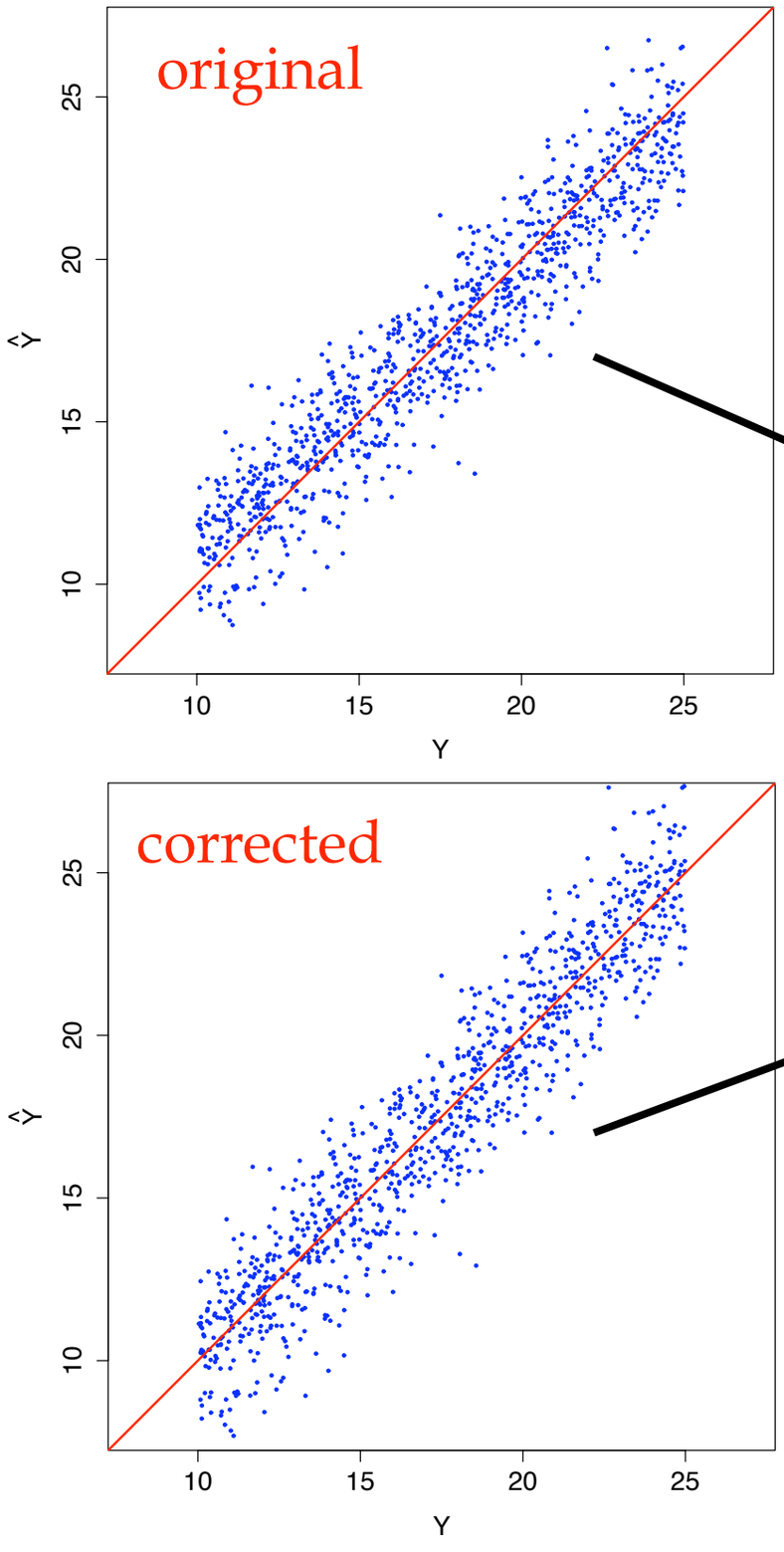
$$\Omega(\rho) = \frac{\sigma_x^2}{\sigma_x^2 + (1+\rho)\sigma_u^2}\beta_1$$

$$\hat{\Omega}(\rho) = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_{1,i}(\rho)$$

$$\hat{\Omega}(-1) \to \beta_1$$

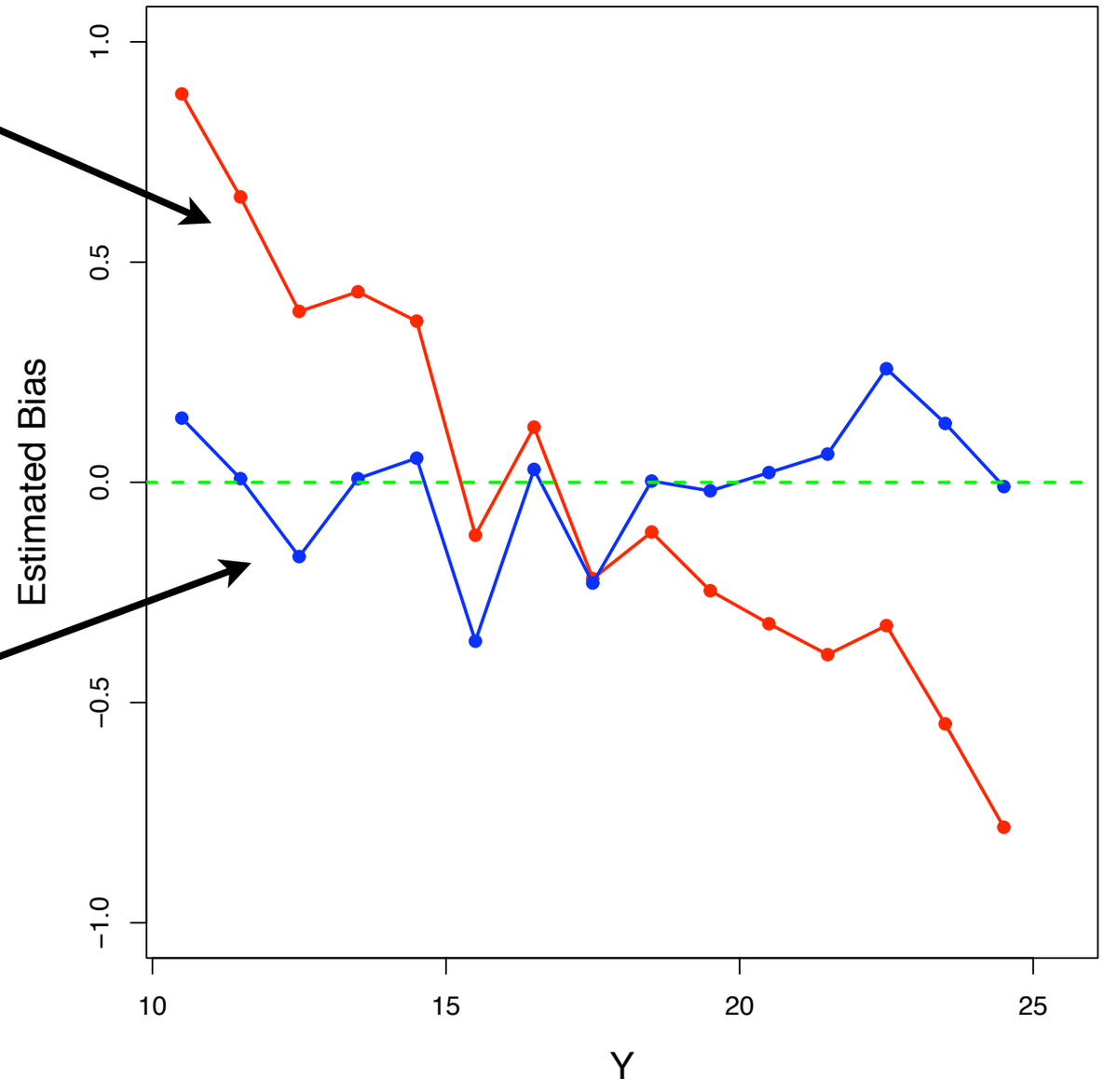


Simulation-Extrapolation (SIMEX)



Result: reduced bias
but (slightly) higher variance.

(You can't have everything...)

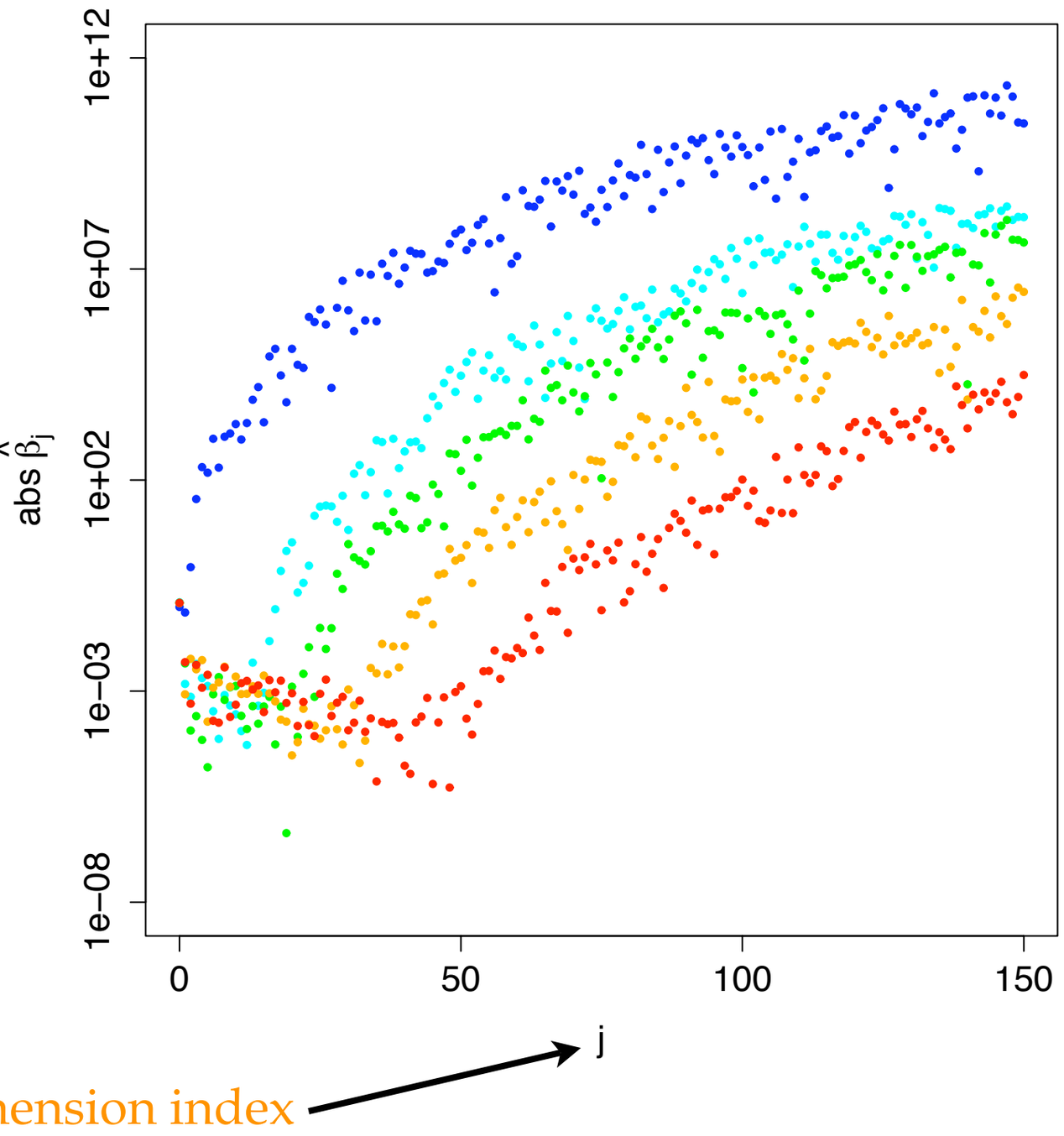


Simulation-Extrapolation (SIMEX)

Real-life example: naive application to heteroskedastic SDSS MSG data via flux resampling.

Forecast: accurate extrapolation to $\rho = -1$ may be difficult.

Results: TBD



$\rho = 0$

$\rho = 0.1$

$\rho = 0.25$

$\rho = 0.5$

$\rho = 1$