Measurement Error Corrections with Non-IID Auxiliary Data

Dylan Spicker*, Michael Wallace, Grace Yi

University of Waterloo

Thursday June 10, 2021

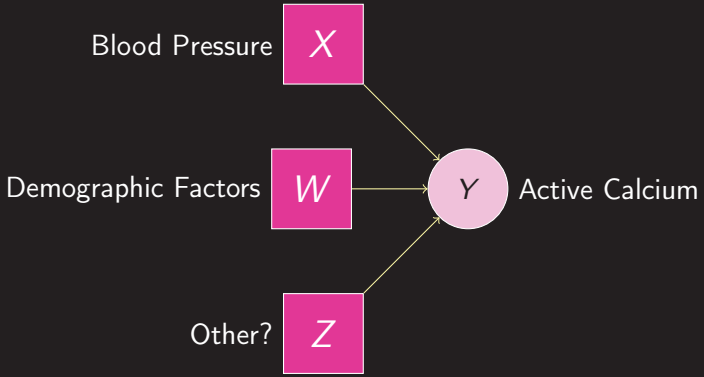
My Goals Today

1. Introduce what measurement error is.
2. Demonstrate how measurement error impacts analyses.
3. Show the common corrections used to overcome these concerns.
4. Provide an alternative method which makes more realistic assumptions.

An Illustrative Example

Can we determine active calcium in a patient from factors which are cheaper to measure?

An Illustrative Example



An Illustrative Example

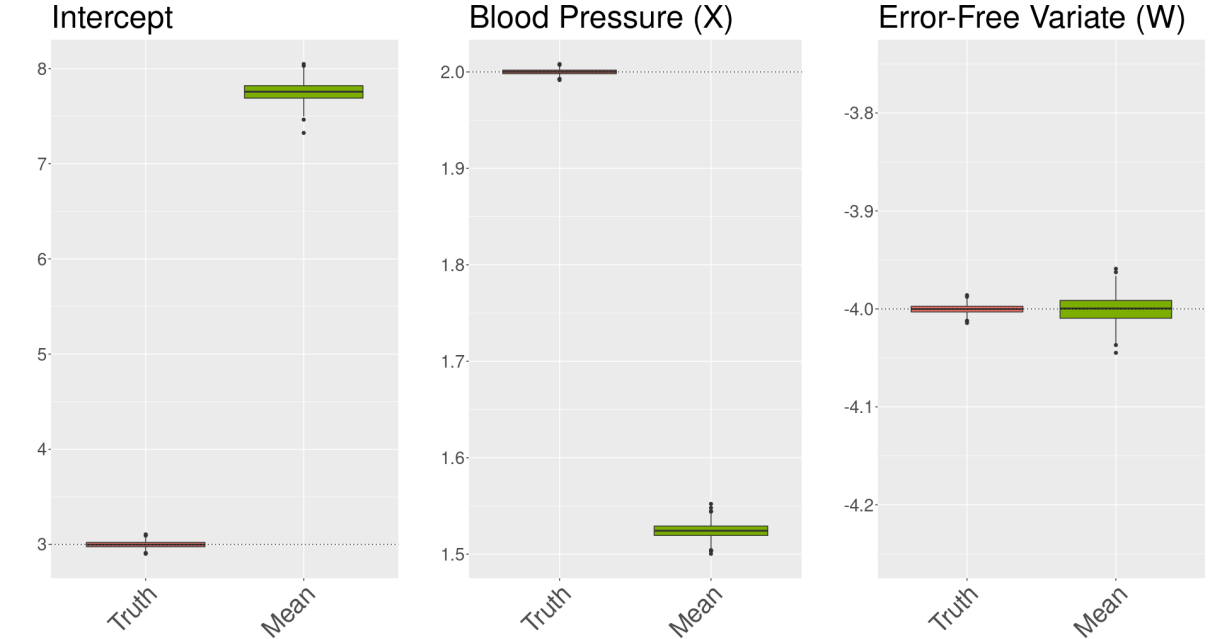
We care about long-term average blood pressure, but we can only measure current blood pressure.

$$\text{Measured Blood Pressure} = \text{True Blood Pressure} + \text{Noise}$$

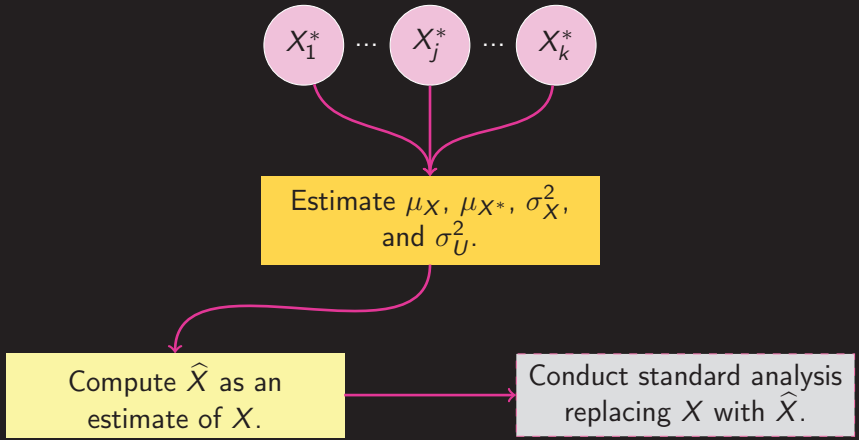
Simulated Dataset

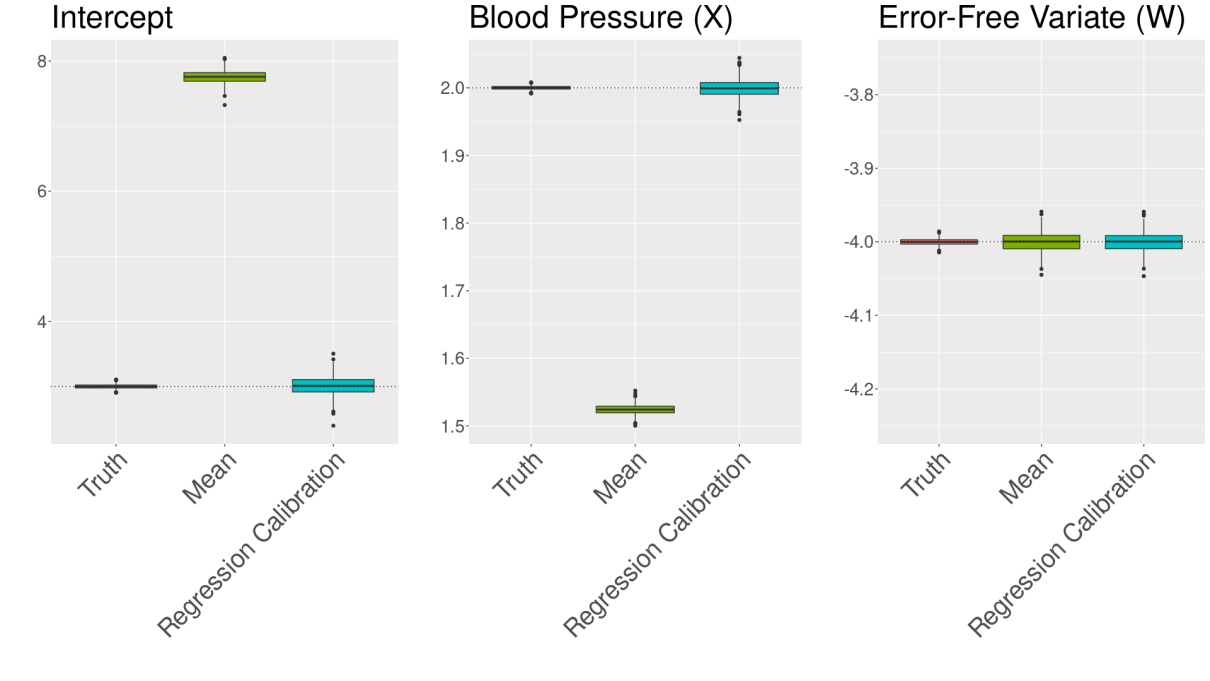
Patient #	Active Calcium	Demo. Factors	True BP	Measured BP (1) $X+U$	Measured BP (2) $X+U$	Measured BP (3) $X+U$	Measured BP (4) $X+U$
1	Y_1	W_1	X_1	X^*_{11}	X^*_{12}	X^*_{13}	X^*_{14}
2	Y_2	W_2	X_2	X^*_{21}	X^*_{22}	X^*_{23}	X^*_{24}
3	Y_3	W_3	X_3	X^*_{31}	X^*_{32}	X^*_{33}	X^*_{34}
				$\vdots$			
n	Y_n	W_n	X_n	X^*_{n1}	X^*_{n2}	X^*_{n3}	X^*_{n4}

Goal: Determine the relationship given by $E[Y|X, W]$, using linear regression.


Intercept
8
7
6
5
4
3
Truth
Mean
Blood Pressure (X)
2.0
1.9
1.8
1.7
1.6
1.5
Truth
Mean
Error-Free Variate (W)
-3.8
-3.9
-4.0
-4.1
-4.2
Truth
Mean

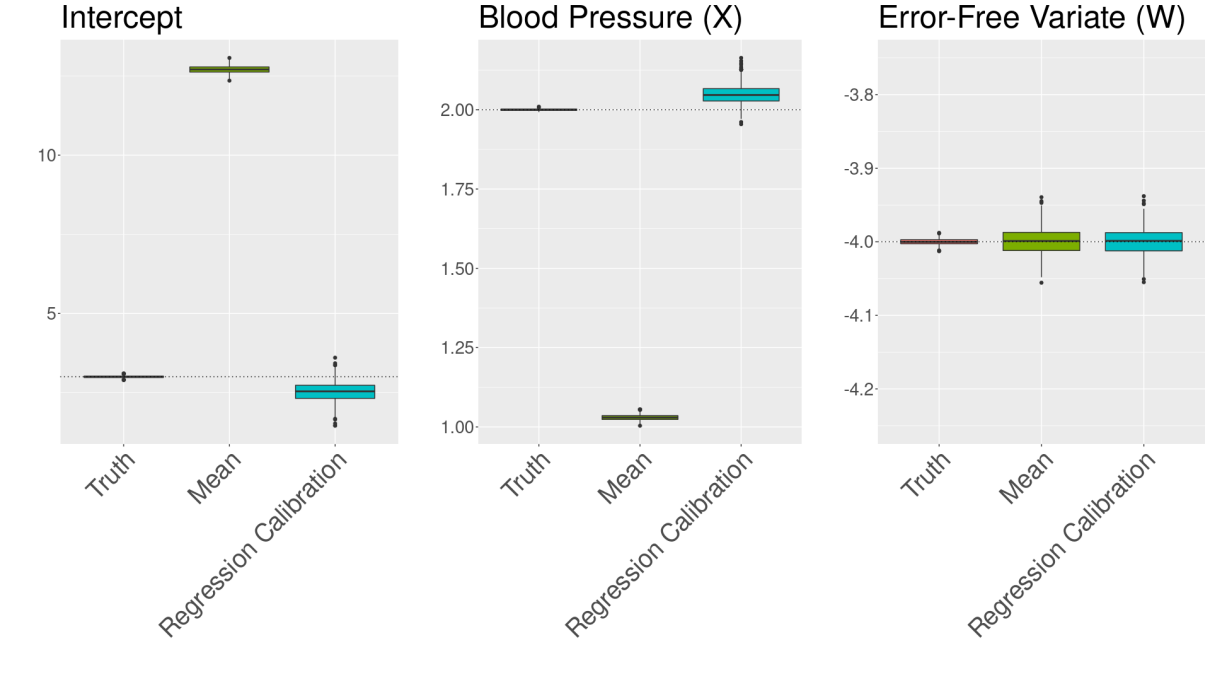
Basic Correction: Regression Calibration



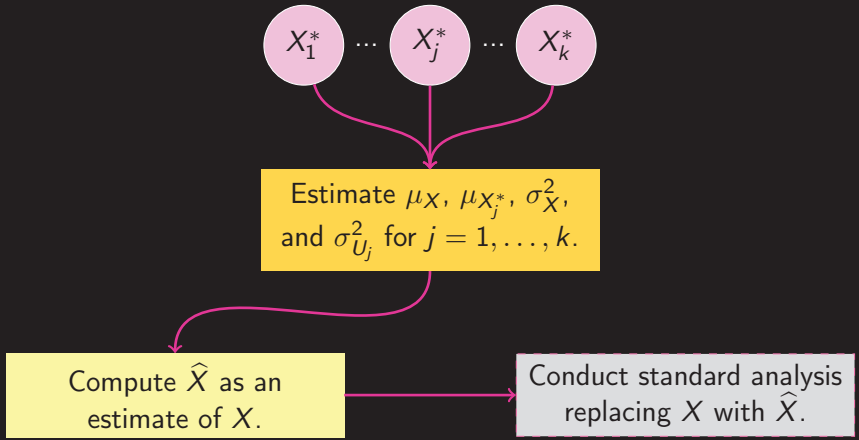

Intercept
8
6
4
Truth
Mean
Regression Calibration
Blood Pressure (X)
2.0
1.9
1.8
1.7
1.6
1.5
Truth
Mean
Regression Calibration
Error-Free Variate (W)
-3.8
-3.9
-4.0
-4.1
-4.2
Truth
Mean
Regression Calibration

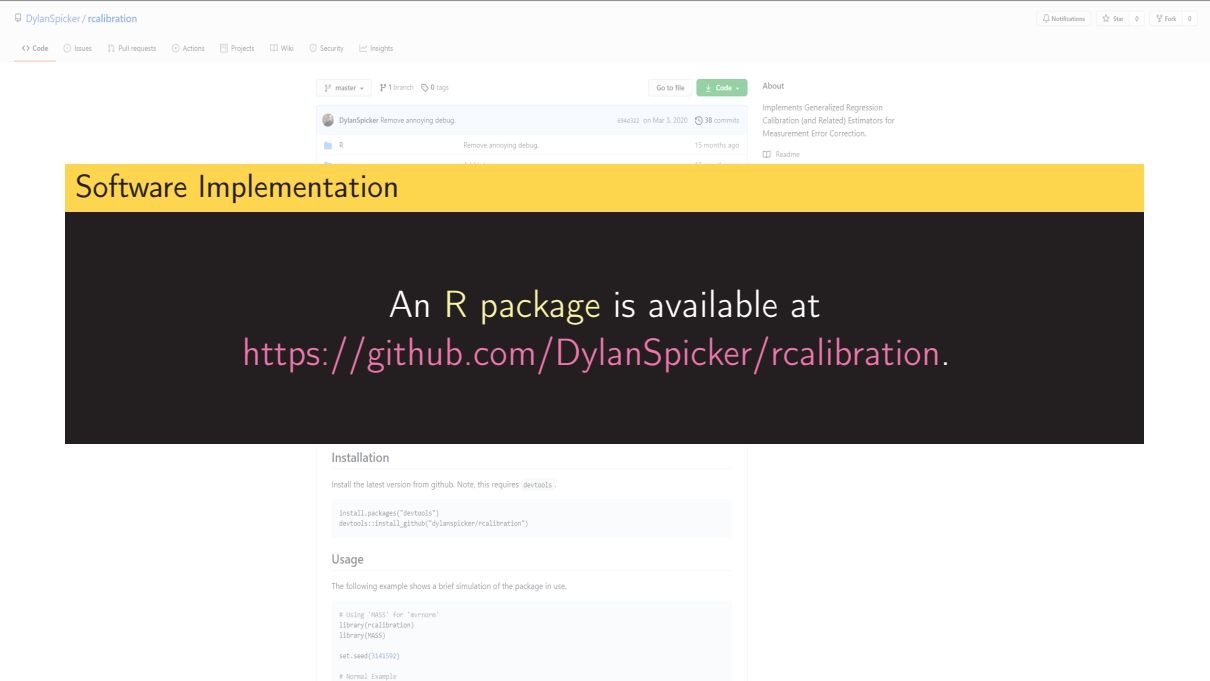
More Realistic Simulated Dataset

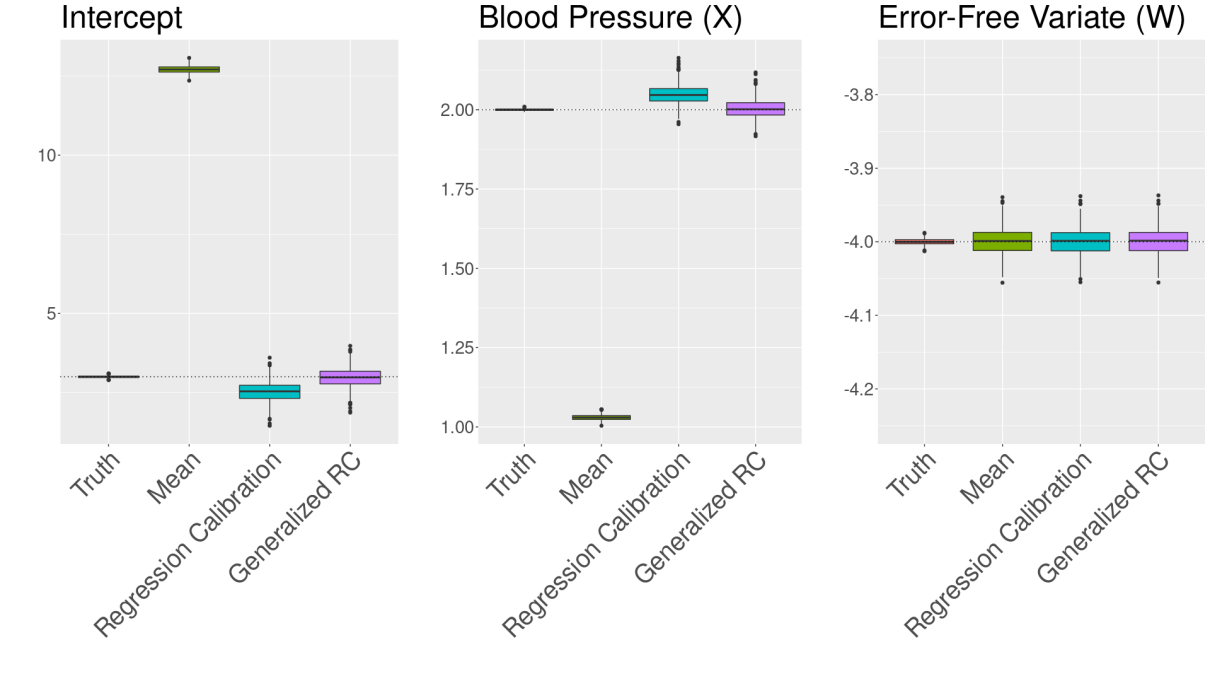
Patient #	Active Calcium	Demo. Factors	True BP	Measured BP (1) $X + U_1$	Measured BP (2) $X + U_2$	Measured BP (3) $X + U_3$	Measured BP (4) $X + U_4$
1	Y_1	W_1	X_1	X^*_{11}	X^*_{12}	X^*_{13}	–
2	Y_2	W_2	X_2	X^*_{21}	–	X^*_{23}	X^*_{24}
3	Y_3	W_3	X_3	X^*_{31}	X^*_{32}	–	–
				⋮			
n	Y_n	W_n	X_n	X^*_{n1}	–	X^*_{n3}	X^*_{n4}


Intercept
10
5
Truth
Mean
Regression Calibration
Blood Pressure (X)
2.00
1.75
1.50
1.25
1.00
Truth
Mean
Regression Calibration
Error-Free Variate (W)
-3.8
-3.9
-4.0
-4.1
-4.2
Truth
Mean
Regression Calibration

Our Suggestion: Generalized Regression Calibration




DylanSpicker / rcalibration
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
master
1 branch
0 tags
Go to file
Code
DylanSpicker Remove annoying debug.
e94d322 on Mar 3, 2020
38 commits
R
Remove annoying debug.
15 months ago
About
Implements Generalized Regression Calibration (and Related) Estimators for Measurement Error Correction.
Readme
Software Implementation
An R package is available at https://github.com/DylanSpicker/rcalibration.
Installation
Install the latest version from github. Note, this requires devtools.
install.packages("devtools")
devtools::install_github("dylanspicker/rcalibration")
Usage
The following example shows a brief simulation of the package in use.
# Using 'MASS' for 'mvrnorm'
library(rcalibration)
library(MASS)
set.seed(3141592)
# Normal Example


Intercept
10
5
Truth
Mean
Regression Calibration
Generalized RC
Blood Pressure (X)
2.00
1.75
1.50
1.25
1.00
Truth
Mean
Regression Calibration
Generalized RC
Error-Free Variate (W)
-3.8
-3.9
-4.0
-4.1
-4.2
Truth
Mean
Regression Calibration
Generalized RC

Extensions to this Work

- Works on any correction technique assuming replicates to estimate variance components.

Extensions to this Work

- Works on any correction technique assuming replicates to estimate variance components.
- Can accommodate multiplicative or additive measurement error models.

Extensions to this Work

- Works on any correction technique assuming replicates to estimate variance components.
- Can accommodate multiplicative or additive measurement error models.
- Can accommodate biased proxies.

Extensions to this Work

- Works on any correction technique assuming replicates to estimate variance components.
- Can accommodate multiplicative or additive measurement error models.
- Can accommodate biased proxies.
- Results in asymptotically normal estimators.

Conclusions

By adjusting the underlying parameter estimators we can allow for violations of the assumption that replicated measurements are identically distributed in many common measurement error correction procedures.

This is done with little additional complexity.

Thank You.

Dylan Spicker
dylan.spicker@uwaterloo.ca | www.dylanspicker.com