Supplementary material for: The non-zero mean SIMEX: Improving estimation in the face of measurement error

Nabila Parveen
Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

Erica E. M. Moodie **erica.moodie@mcgill.ca**
Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

Bluma Brenner
Lady Davis Research Institute, Montreal, Quebec, Canada

This supplement contains details of the design of the simulation study (Table 1) and additional findings from the SPOT data (Table 2-5).

Design of the Simulation Study

Table 1: Simulation Scenarios

Scenario	Distribution of (U,V)	True δ^*	Y	Assumed δ^*_b
1	$N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & 0.447\\0.447 & 1\end{pmatrix}\right\}$	$\delta^* = \lvert\delta\rvert$ and $\delta \sim N(0, 0.25)$	$N(\eta_1^a,1)$	$\delta^*_b = \lvert\delta_b\rvert$ and $\delta_b \sim N(0, 0.25)$
2	$N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & 0.447\\0.447 & 1\end{pmatrix}\right\}$	$\delta^* = \lvert\delta\rvert$ and $\delta \sim N(0, 0.5)$	$N(\eta_1^a, 1)$	$\delta^*_b = \lvert\delta_b\rvert$ and $\delta_b \sim N(0, 0.5)$
3	$N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & 0.447\\0.447 & 1\end{pmatrix}\right\}$	$\delta^* = \lvert\delta\rvert$ and $\delta \sim N(0, 1)$	$N(\eta_1^a, 1)$	$\delta^*_b = \lvert\delta_b\rvert$ and $\delta_b \sim N(0, 1)$
4	$N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & 0.447\\0.447 & 1\end{pmatrix}\right\}$	$\delta^* = \lvert\delta\rvert$ and $\delta \sim N(0, 2)$	$N(\eta_1^a, 1)$	$\delta^*_b = \lvert\delta_b\rvert$ and $\delta_b \sim N(0, 2)$
5	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(1.5)$	$N(\eta_2^b, 1)$	$\delta^*_b \sim P(1.5)$
6	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(3)$	$N(\eta_2^b, 1)$	$\delta^*_b \sim P(3)$
7	$N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & 0.447\\0.447 & 1\end{pmatrix}\right\}$	$\delta \sim N(0, 0.5)$	$N(\eta_1^a, 1)$	$\delta_b \sim N(0, 0.25)$
8	$N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & 0.447\\0.447 & 1\end{pmatrix}\right\}$	$\delta \sim N(0, 0.5)$	$N(\eta_1^a, 1)$	$\delta_b \sim N(0, 0.75)$
9	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(1.5)$	$N(\eta_2^b, 1)$	$\delta^*_b \sim P(0.75)$
10	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(1.5)$	$N(\eta_2^b, 1)$	$\delta^*_b \sim P(2.25)$
11	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(1.5)$	$P(\exp(\eta_3^c))$	$\delta^*_b \sim P(1.5)$
12	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(3)$	$P(\exp(\eta_3^c))$	$\delta^*_b \sim P(3)$
13	$U \sim P(12), V \sim N(0, 1)$	$\delta^* \sim P(3)$	$Bernoulli(p^d, 1)$	$\delta^*_b \sim P(3)$

$\eta_1^a = -2 + 1 * U + 0.25 * V + 0.25 * UV$
$\eta_2^b = 1 + 1 * U + 1 * V + 0.5 * UV$
$\eta_3^c = 0.25 + 0.5 * U + 0.05 * V + 0.05 * UV$
$p^d = \frac{\exp(\eta_4)}{1+\exp(\eta_4)}$, where $\eta_4 = -2 + 0.25 * U - 1 * V + 0.25 * UV$

SPOT Data Analysis: Additional Results

Table 2: Results from the simple logistic regression model. SIMEX-Q is the NZM SIMEX with a quadratic fit in the extrapolation step; SIMEX-NL is the NZM SIMEX with a non-linear fit in the extrapolation step.

		Naïve			SIMEX-Q			SIMEX-NL		
Parameter	μ^*	$\hat{\beta}^{OLS}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value
Model: Relating condom use to cluster size										
β_0	3	2.619	0.772	0.001	2.198	1.104	0.012	2.197	1.105	0.012
β_1	3	-0.047	0.036	0.195	-0.013	0.083	0.567	-0.013	0.083	0.559
β_0	1				2.151	0.946	0.005	2.151	0.945	0.005
β_1	1				-0.010	0.083	0.569	-0.010	0.083	0.572
β_0	5				2.229	1.289	0.025	2.229	1.291	0.027
β_1	5				-0.013	0.085	0.564	-0.013	0.085	0.581
β_0	10				2.289	1.739	0.076	2.288	1.746	0.069
β_1	10				-0.014	0.088	0.591	-0.013	0.088	0.565
Model: Relating HIV tests in the last 24 months to cluster size										
β_0	3	1.686	0.732	0.021	1.226	0.654	0.015	1.228	0.655	0.018
β_1	3	0.038	0.067	0.573	0.044	0.054	0.515	0.044	0.054	0.482
β_0	1				1.311	0.564	0.004	1.311	0.564	0.003
β_1	1				0.046	0.057	0.499	0.046	0.057	0.535
β_0	5				1.126	0.781	0.055	1.130	0.781	0.071
β_1	5				0.045	0.058	0.527	0.044	0.058	0.448
β_0	10				0.955	1.082	0.255	0.964	1.095	0.219
β_1	10				0.042	0.058	0.486	0.041	0.059	0.559

* mean of the measurement error distribution, Poisson(μ)

Table 3: Results from log-linear model of number of sex partner on cluster size. SIMEX-Q is the NZM SIMEX with a quadratic fit in the extrapolation step; SIMEX-NL is the NZM SIMEX with a non-linear fit in the extrapolation step.

		Naïve			SIMEX-Q			SIMEX-NL		
Parameter	μ^*	$\hat{\beta}^{OLS}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value
Model: Relating number of sex partner on cluster size										
β_0	3	1.716	0.097	0.000	1.709	0.202	0.000	1.700	0.201	0.000
β_1	3	0.004	0.006	0.526	0.003	0.011	0.752	0.004	0.011	0.702
β_0	1				1.714	0.184	0.000	1.709	0.184	0.000
β_1	1				0.004	0.011	0.737	0.004	0.011	0.711
β_0	5				1.692	0.224	0.000	1.705	0.224	0.000
β_1	5				0.004	0.011	0.707	0.003	0.011	0.762
β_0	10				1.659	0.275	0.000	1.659	0.278	0.000
β_1	10				0.005	0.012	0.704	0.005	0.013	0.706

* mean of the measurement error distribution, Poisson(μ)

Table 4: Results from log-linear model of number of one night partner on cluster size. SIMEX-Q is the NZM SIMEX with a quadratic fit in the extrapolation step; SIMEX-NL is the NZM SIMEX with a non-linear fit in the extrapolation step.

		Naïve			SIMEX-Q			SIMEX-NL		
Parameter	μ^*	$\hat{\beta}^{OLS}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value
Model: Relating number of one night partner on cluster size										
β_0	3	1.411	0.112	0.000	1.399	0.296	0.000	1.389	0.296	0.000
β_1	3	0.004	0.006	0.568	0.004	0.015	0.797	0.004	0.015	0.761
β_0	1				1.407	0.267	0.000	1.403	0.267	0.000
β_1	1				0.004	0.014	0.787	0.004	0.014	0.767
β_0	5				1.390	0.314	0.000	1.391	0.313	0.000
β_1	5				0.004	0.016	0.801	0.003	0.016	0.803
β_0	10				1.354	0.408	0.001	1.345	0.414	0.001
β_1	10				0.005	0.018	0.789	0.005	0.018	0.778

* mean of the measurement error distribution, Poisson(μ)

Table 5: Results from multinomial model of number of one night partner on cluster size. SIMEX-Q is the NZM SIMEX with a quadratic fit in the extrapolation step; SIMEX-NL is the NZM SIMEX with a non-linear fit in the extrapolation step.

		Naïve			SIMEX-Q			SIMEX-NL		
Parameter	μ^*	$\hat{\beta}^{OLS}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value	$\hat{\beta}^{NZM}$	SE	p-value
Model: Relating number of one night partner on cluster size										
$\beta_{0(2-4)}$	3	-1.605	0.780	0.039	-1.727	15.143	0.909	-1.740	31.497	0.955
$\beta_{1(2-4)}$	3	0.036	0.051	0.474	0.037	0.573	0.948	0.038	1.179	0.974
$\beta_{0(5+)}$	3	-0.302	0.522	0.562	-0.427	0.760	0.574	-0.441	0.762	0.562
$\beta_{1(5+)}$	3	0.038	0.039	0.331	0.039	0.058	0.499	0.039	0.057	0.488
$\beta_{0(2-4)}$	1				-1.660	12.812	0.897	-1.630	22.317	0.942
$\beta_{1(2-4)}$	1				0.038	0.524	0.942	0.036	0.903	0.968
$\beta_{0(5+)}$	1				-0.347	0.660	0.599	-0.354	0.660	0.591
$\beta_{1(5+)}$	1				0.038	0.055	0.483	0.039	0.055	0.477
$\beta_{0(2-4)}$	5				-1.792	16.775	0.915	-1.817	38.775	0.963
$\beta_{1(2-4)}$	5				0.036	0.591	0.950	0.039	1.351	0.977
$\beta_{0(5+)}$	5				-0.499	0.879	0.569	-0.501	0.881	0.569
$\beta_{1(5+)}$	5				0.038	0.060	0.524	0.038	0.060	0.518
$\beta_{0(2-4)}$	10				-2.012	14.199	0.887	-2.01	34.997	0.954
$\beta_{1(2-4)}$	10				0.039	0.428	0.927	0.038	1.045	0.970
$\beta_{0(5+)}$	10				-0.767	1.182	0.516	-0.746	1.214	0.538
$\beta_{1(5+)}$	10				0.042	0.064	0.511	0.041	0.067	0.544

* mean of the measurement error distribution, Poisson(μ)