

Online Supplement to the *Estimation with Pairwise Observations* paper

Felix Chan¹, László Mátyás², and Kristóf Reizinger²

¹Curtin University, Perth, Australia

²Central European University, Budapest, Hungary and Vienna, Austria

May 18, 2023

Acknowledgement: Contribution by Gyorgy Ruzicska in the early stages of this project is kindly acknowledged. Providing the initial versions of the code files is also highly appreciated.

The Online Supplement presents additional summary tables and figures related to the full-pairwise Monte Carlo exercises. Namely, the results provide a more comprehensive overview about the coefficient estimates, the test statistics and the empirical distribution of the test statistics considering Normal and Uniform data generating processes, and Δx and $|\Delta x|$ weighting.

The related codes were written by the authors and are available on [GitHub](#).

Appendix A: Monte Carlo Simulations Setups and Simulation Results for the EwPO Estimation

This Online Supplement presents additional Monte Carlo (MC) simulation results to assess the properties of the EwPO estimator with selected weights.

The data generating process for the MC simulations is based on the model

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

and the MC experiments consider two possible distributions for u_i , namely

1. $u_i \sim N(0, 1)$,
2. $u_i \sim$ skewed normal distribution,

where the skewed normal distribution is generated as

$$u_i = \xi + \lambda|v_i| + z_i,$$

with $\xi = -\lambda\sqrt{\frac{2}{\pi}}$, $v_i \sim N(0, 1)$ and $z_i \sim N(0, \sigma^2)$ such that v_i and w_i are independently distributed.

The MC experiments consider uniform distribution $U(-10, 10)$ for the regressor x_i .

The parameter vector, presented here (for purposes of robustness checking), is $(\beta_0, \beta_1) = (1, 1.5)$. The number of MC replications is 1000.

Sorted MC	Estimates and MC standard errors			
n=50	Parameter	Estimate/S.e.	OLS	pairwise
	$\hat{\beta}_0$	Estimate	0.9866	0.9866
		S.e.	0.1609	0.1609
	$\hat{\beta}_1$	Estimate	1.4999	1.4999
		S.e.	0.2928	0.2928
n = 500	$\hat{\beta}_0$	Estimate	1.0001	1.0001
		S.e.	0.0504	0.0504
	$\hat{\beta}_1$	Estimate	1.4971	1.4971
		S.e.	0.0948	0.0948
n = 5000	$\hat{\beta}_0$	Estimate	0.9992	0.9992
		S.e.	0.0169	0.0169
	$\hat{\beta}_1$	Estimate	1.4997	1.4997
		S.e.	0.0281	0.0281

Table 1: Sorted – full-pairwise MC, $x_i \sim U(-10, 10)$, $u_i \sim$ skewed normal, Δx weighted estimator

Sorted MC	Estimates and MC standard errors			
	Parameter	Estimate/S.e.	OLS	pairwise
n=50	$\hat{\beta}_0$	Estimate	0.9866	0.9866
		S.e.	0.1609	0.1609
	$\hat{\beta}_1$	Estimate	1.4999	1.4999
		S.e.	0.2928	0.2928
n = 500	$\hat{\beta}_0$	Estimate	1.0001	1.0001
		S.e.	0.0504	0.0504
	$\hat{\beta}_1$	Estimate	1.4971	1.4971
		S.e.	0.0948	0.0948
n = 5000	$\hat{\beta}_0$	Estimate	0.9992	0.9992
		S.e.	0.0169	0.0169
	$\hat{\beta}_1$	Estimate	1.4997	1.4997
		S.e.	0.0281	0.0281

Table 2: Non-sorted full-pairwise MC, $x_i \sim U(-10, 10)$, Δx weighted estimator

Note 1: It is no mistake, the sorted and non-sorted results are identical here.

Note 2: In general, the standard errors are much larger for β_0 than β_1 , but when the distribution of the x_i -s is ‘informative’, they are in fact quite close to the OLS ones.

Non-sorted MC	Estimates and MC standard errors			
	Parameter	Estimate/S.e.	OLS	pairwise
n=50	$\hat{\beta}_0$	Estimate	1.0009	0.9967
		S.e.	0.1408	0.1997
	$\hat{\beta}_1$	Estimate	1.5007	1.5018
		S.e.	0.0251	0.0291
n = 500	$\hat{\beta}_0$	Estimate	0.9967	0.9966
		S.e.	0.0446	0.0640
	$\hat{\beta}_1$	Estimate	1.5001	1.4998
		S.e.	0.0081	0.0095
n = 5000	$\hat{\beta}_0$	Estimate	1.001	1.0004
		S.e.	0.0144	0.0192
	$\hat{\beta}_1$	Estimate	1.4999	1.4999
		S.e.	0.0025	0.0031

Table 3: Non-sorted adjacent MC, $x_i \sim U(-10, 10)$, $u_i \sim N(0, 1)$, $|\Delta x|$ weighted estimator

Full-pairwise MC	Estimates and MC standard errors			
n=50	Parameter	Estimate/S.e.	OLS	pairwise
	$\hat{\beta}_0$	Estimate	1.0009	1.0009
		S.e.	0.1408	0.1408
	$\hat{\beta}_1$	Estimate	1.5007	1.5007
		S.e.	0.0251	0.0251
n = 500	$\hat{\beta}_0$	Estimate	0.9967	0.9967
		S.e.	0.0446	0.0446
	$\hat{\beta}_1$	Estimate	1.5001	1.5001
		S.e.	0.0081	0.0081
n = 5000	$\hat{\beta}_0$	Estimate	1.0001	1.001
		S.e.	0.0144	0.0144
	$\hat{\beta}_1$	Estimate	1.4999	1.4999
		S.e.	0.0025	0.0025

Table 4: Sorted full-pairwise MC, $x_i \sim U(-10, 10)$, $u_i \sim N(0, 1)$, Δx weighted estimator

Appendix B: Monte Carlo Simulations Setups and Simulation Results for the Test

The MC setup considers sample size $n = 50$, 500 , and 5000 with 1000 replications.

Step 1. Generate model the model with one explanatory variable namely,

$$y_i = \alpha + x_i\beta + u_i$$

with $\alpha = 1$, $\beta = 0.5$ to start with, and u_i is generated as $N(0, 1)$. Finally, x should be generated as $N(0, 1)$ and also $U(-5, 5)$.

The simulation of x_i and u_i is conducted under four different correlations namely $\rho = 0$ (benchmark ideal case), $\rho = 0.2$ (small), $\rho = 0.5$ (medium), and $\rho = 0.8$ (large).

Step 2. Estimate the model with EwPO with $w_{ij} = \Delta x_{ij}$ and $w_{ij} = |\Delta x_{ij}|$. In each case, calculate the test statistics as defined in Equation (B.1).

$$S(\mathbf{w}) = n^{-2} \sum_{p=2}^n \sum_{q=1}^{p-1} \Delta x_{pq} \Delta \hat{u}_{pq}. \quad (\text{B.1})$$

Full-pairwise MC		Average test statistics			
	Parameter	Pairwise	Standard deviation	Skewness	Kurtosis
n=50	Exogen	0.0204	0.6551	-0.0043	3.2285
	$\rho = 0.2$	-0.4338	0.7196	0.0537	2.7248
	$\rho = 0.5$	-0.6800	0.5816	0.0045	3.0386
	$\rho = 0.8$	-2.2359	0.4627	0.0591	2.9205
n = 500	Exogen	0.0014	0.2100	-0.0992	3.1302
	$\rho = 0.2$	-0.4109	0.2199	-0.0596	3.1559
	$\rho = 0.5$	-0.9357	0.1888	0.0415	2.9658
	$\rho = 0.8$	-1.4518	0.1335	0.1038	3.1749
n = 5000	Exogen	0.0001	0.0671	0.0696	2.8285
	$\rho = 0.2$	-0.3970	0.0680	-0.0949	2.8970
	$\rho = 0.5$	-1.0143	0.0594	-0.2178	3.1821
	$\rho = 0.8$	-1.5325	0.0407	-0.0089	3.0313

Table 5: Average test statistics, full-pairwise MC, $x_i \sim N(0, 5)$, $|\Delta x|$ weighted estimator

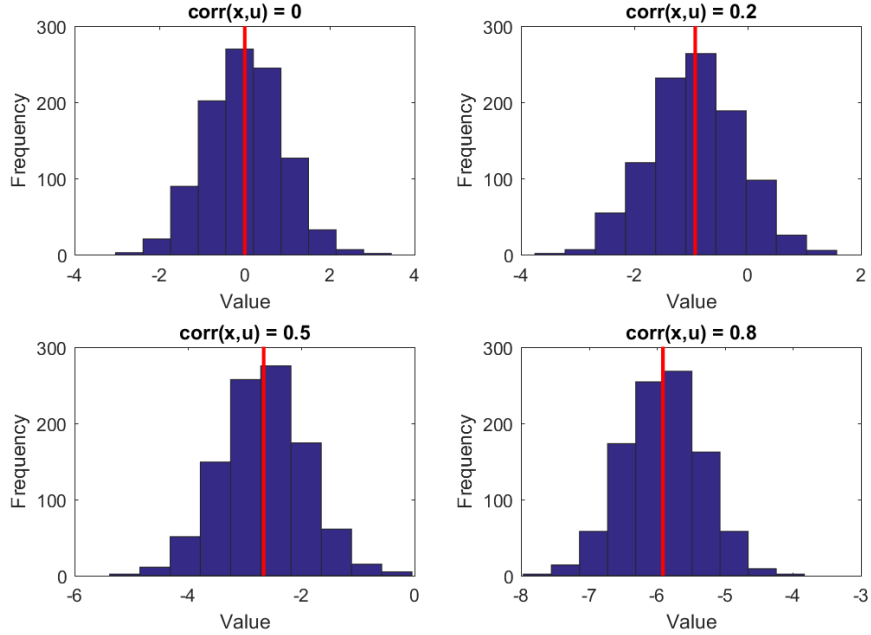


Figure 1: Distribution of the test-statistics, $x_i \sim \text{Uniform}(-5, 5)$, Δx weighted full-pairwise estimator, $n = 50$

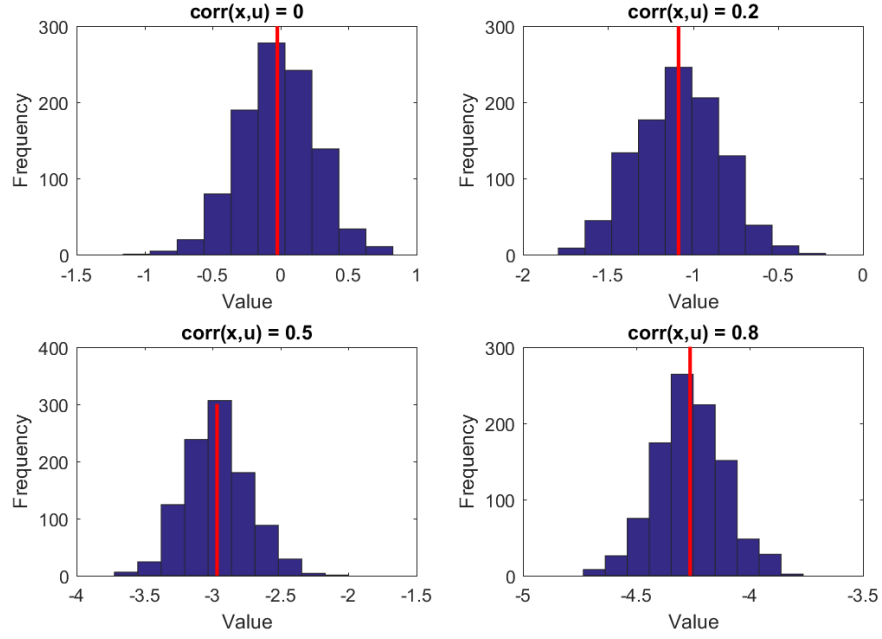


Figure 2: Distribution of the test-statistics, $x_i \sim \text{Uniform}(-5,5)$, Δx weighted full-pairwise estimator, $n = 500$

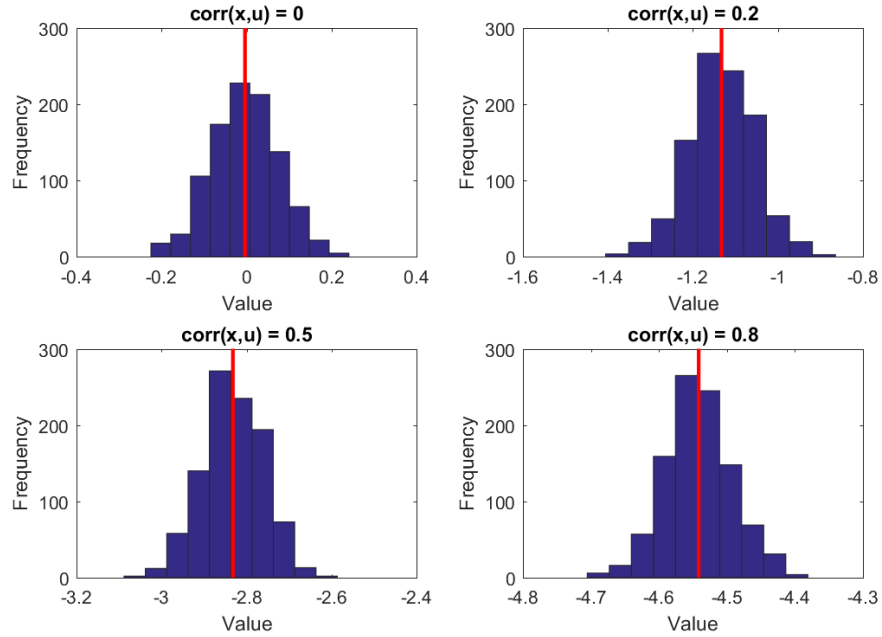


Figure 3: Distribution of the test-statistics, $x_i \sim \text{Uniform}(-5,5)$, Δx weighted full-pairwise estimator, $n = 5000$

Full-Pairwise MC	Estimates and MC standard errors			
	Parameter	Estimate/S.e.	OLS	pairwise
n=50	Exogen	Estimate	0.4997	0.4999
		S.e.	0.0602	0.0638
	$\rho = 0.2$	Estimate	0.5378	0.5408
		S.e.	0.0573	0.0607
	$\rho = 0.5$	Estimate	0.5852	0.5781
		S.e.	0.1609	0.1801
	$\rho = 0.8$	Estimate	0.6277	0.6348
		S.e.	0.0395	0.0408
n=500	Exogen	Estimate	0.5007	0.5006
		S.e.	0.0202	0.0207
	$\rho = 0.2$	Estimate	0.5426	0.5430
		S.e.	0.0190	0.0196
	$\rho = 0.5$	Estimate	0.6027	0.6000
		S.e.	0.0162	0.0165
	$\rho = 0.8$	Estimate	0.6586	0.6525
		S.e.	0.0113	0.0116
n=5000	Exogen	Estimate	0.5003	0.5003
		S.e.	0.0062	0.0063
	$\rho = 0.2$	Estimate	0.5394	0.5393
		S.e.	0.0060	0.0062
	$\rho = 0.5$	Estimate	0.5962	0.5963
		S.e.	0.0055	0.0056
	$\rho = 0.8$	Estimate	0.6571	0.6563
		S.e.	0.0038	0.0039

Table 6: β_1 coefficient estimates - Δx weighted full-pairwise MC, $x_i \sim N(0,5)$

Full-Pairwise MC	Estimates and MC standard errors			
	Parameter	Estimate/S.e.	OLS	pairwise
n=50	Exogen	Estimate	0.5014	0.5014
		S.e.	0.0156	0.0156
	$\rho = 0.2$	Estimate	0.5662	0.5663
		S.e.	0.0150	0.0151
	$\rho = 0.5$	Estimate	0.6677	0.6684
		S.e.	0.0131	0.0131
	$\rho = 0.8$	Estimate	0.7720	0.7707
		S.e.	0.0094	0.0094
n=500	Exogen	Estimate	0.4996	0.4997
		S.e.	0.0216	0.0221
	$\rho = 0.2$	Estimate	0.5406	0.5399
		S.e.	0.0191	0.0196
	$\rho = 0.5$	Estimate	0.6047	0.6038
		S.e.	0.0184	0.0192
	$\rho = 0.8$	Estimate	0.6456	0.6474
		S.e.	0.0122	0.0126
n=5000	Exogen	Estimate	0.5001	0.5000
		S.e.	0.0061	0.0062
	$\rho = 0.2$	Estimate	0.5404	0.5401
		S.e.	0.0061	0.0063
	$\rho = 0.5$	Estimate	0.6017	0.6016
		S.e.	0.0055	0.0056
	$\rho = 0.8$	Estimate	0.6539	0.6546
		S.e.	0.0038	0.0039

Table 7: β_1 coefficient estimates - $|\Delta x|$ weighted full-pairwise MC, $x_i \sim N(0,5)$

Full-Pairwise MC	Estimates and MC standard errors			
	Parameter	Estimate/S.e.	OLS	pairwise
n=50	Exogen	Estimate	0.4999	0.4998
		S.e.	0.0464	0.0467
	$\rho = 0.2$	Estimate	0.5625	0.5629
		S.e.	0.0526	0.0529
	$\rho = 0.5$	Estimate	0.6613	0.6622
		S.e.	0.0445	0.0448
	$\rho = 0.8$	Estimate	0.7841	0.7875
		S.e.	0.0277	0.0278
n=500	Exogen	Estimate	0.5014	0.5014
		S.e.	0.0156	0.0156
	$\rho = 0.2$	Estimate	0.5662	0.5663
		S.e.	0.0150	0.0151
	$\rho = 0.5$	Estimate	0.6677	0.6684
		S.e.	0.0131	0.0131
	$\rho = 0.8$	Estimate	0.7720	0.7707
		S.e.	0.0094	0.0094
n=5000	Exogen	Estimate	0.5002	0.5002
		S.e.	0.0048	0.0048
	$\rho = 0.2$	Estimate	0.5681	0.5681
		S.e.	0.0047	0.0047
	$\rho = 0.5$	Estimate	0.6701	0.6700
		S.e.	0.0043	0.0043
	$\rho = 0.8$	Estimate	0.7717	0.7717
		S.e.	0.0030	0.0030

Table 8: β_1 coefficient estimates - Δx weighted full-pairwise MC, $x_i \sim U(-5,5)$

Full-Pairwise MC	Estimates and MC standard errors			
	Parameter	Estimate/S.e.	OLS	pairwise
n=50	Exogen	Estimate	0.4986	0.4988
		S.e.	0.0492	0.0494
	$\rho = 0.2$	Estimate	0.5664	0.5673
		S.e.	0.0440	0.0442
	$\rho = 0.5$	Estimate	0.6738	0.6726
		S.e.	0.0437	0.0440
	$\rho = 0.8$	Estimate	0.7910	0.7970
		S.e.	0.0262	0.0265
n=500	Exogen	Estimate	0.5000	0.5000
		S.e.	0.0157	0.0157
	$\rho = 0.2$	Estimate	0.5690	0.5692
		S.e.	0.0145	0.0145
	$\rho = 0.5$	Estimate	0.6708	0.6705
		S.e.	0.0136	0.0136
	$\rho = 0.8$	Estimate	0.7726	0.7715
		S.e.	0.0092	0.0092
n=5000	Exogen	Estimate	0.5000	0.5000
		S.e.	0.0047	0.0047
	$\rho = 0.2$	Estimate	0.5679	0.5678
		S.e.	0.0048	0.0048
	$\rho = 0.5$	Estimate	0.6691	0.6691
		S.e.	0.0042	0.0042
	$\rho = 0.8$	Estimate	0.7688	0.7690
		S.e.	0.0029	0.0029

Table 9: β_1 coefficient estimates - $|\Delta x|$ weighted full-pairwise MC, $x_i \sim U(-5,5)$