Supplementary Materials for Response to Guo et al.'s Letter to the Editor

Youyi Fong, Ying Huang, Maria Lemos, Julie McElrath

A Estimated size

Because MW-MW$_2$ and SR-MW$_2$ only make comparisons within the paired subset and the unpaired subset, it is possible to perform permutation tests to obtain p-values. This is especially beneficial for MW-MW$_2$ because the variance of the Wilcoxon-Mann-Whitney statistic for paired data is not distribution-free and the need to estimate its variance under small sample sizes may lead to inflated Type 1 error rates. For example, Table 1 of Guo *and others* (2018) shows that the estimated size of MW-MW$_2$ is 0.068 when $m = 20, l = 5, n = 10, \rho = 0.8$, where m, l, n refer to the number of pairs, the number of independent x's, the number of independent y's, and ρ is the correlation between the two samples. Although using more complicated variance estimators (Fong *and others* , 2018, Section 2.1.2) can lead to smaller inflation (Table A.1-A.4), it is still better to use permutation-based methods to estimate p-values. To obtain permutation-based reference distributions for MW-MW$_2$, we independently permute the paired subset and the unpaired subset, thereby allowing us to preserve the pairing structure in the paired subset. We will refer to this variant of MW-MW$_2$ as MW-MW$_2^{perm}$. Table A.1-A.4 in the supplementary materials show that the Type 1 error rates of MW-MW$_2^{perm}$ are very close to the nominal value.

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			MW-MW$_2$			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	5.9	5.3	5.1	5.1	5.0	4.9	6.0	5.5	5.1	5.0	5.0	5.1	5.6	5.3	4.9
(10,10)	5.5	5.3	5.0	4.7	4.8	4.8	5.8	5.4	5.0	4.9	4.7	4.8	5.2	5.0	5.0
(40,5)	5.3	5.0	5.1	4.8	4.9	5.0	5.8	5.4	5.1	4.8	4.9	5.0	5.2	4.9	4.9

Table A.1: Estimated size, normal distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			MW-MW$_2$			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	5.7	5.1	5.1	5.0	4.9	4.9	6.2	5.7	5.1	4.9	4.9	4.9	5.3	5.1	5.0
(10,10)	5.9	5.3	5.0	5.3	5.0	4.8	6.6	5.7	5.0	5.5	5.1	4.9	5.7	5.3	5.0
(40,5)	5.6	5.3	5.1	5.3	5.0	5.0	6.2	5.4	5.1	5.0	5.0	4.9	5.4	5.2	4.9

Table A.2: Estimated size, logistic distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			MW-MW$_2$			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	5.9	5.3	5.1	5.1	5.0	4.9	6.0	5.5	5.1	5.0	5.0	5.1	5.6	5.3	5.1
(10,10)	5.5	5.3	5.0	4.7	4.8	4.8	5.8	5.4	5.0	5.0	5.0	4.9	5.2	5.1	5.1
(40,5)	5.3	5.0	5.1	4.8	4.9	5.0	5.8	5.4	5.1	4.7	4.8	4.8	5.3	5.0	4.7

Table A.3: Estimated size, gamma distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			MW-MW$_2$			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	5.9	5.3	5.0	5.1	4.9	4.9	6.0	5.3	4.9	5.0	4.9	5.0	5.5	5.2	5.1
(10,10)	5.5	5.1	5.2	4.7	4.8	4.7	5.8	5.4	5.0	4.8	5.0	5.1	5.3	4.9	5.1
(40,5)	5.3	5.0	5.1	4.8	4.9	4.9	5.8	5.1	4.9	4.5	4.6	4.7	5.2	5.0	4.8

Table A.4: Estimated size, lognormal distribution, $m = 20$.

B Power comparison under logistic, gamma and lognormal distributions

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			SR			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	24	34	57	23	35	62	17	31	60	22	36	61	24	35	54
(10,10)	27	37	59	25	37	63	18	31	60	25	38	63	26	38	56
(40,5)	30	41	64	25	37	64	17	31	60	24	38	63	30	42	62

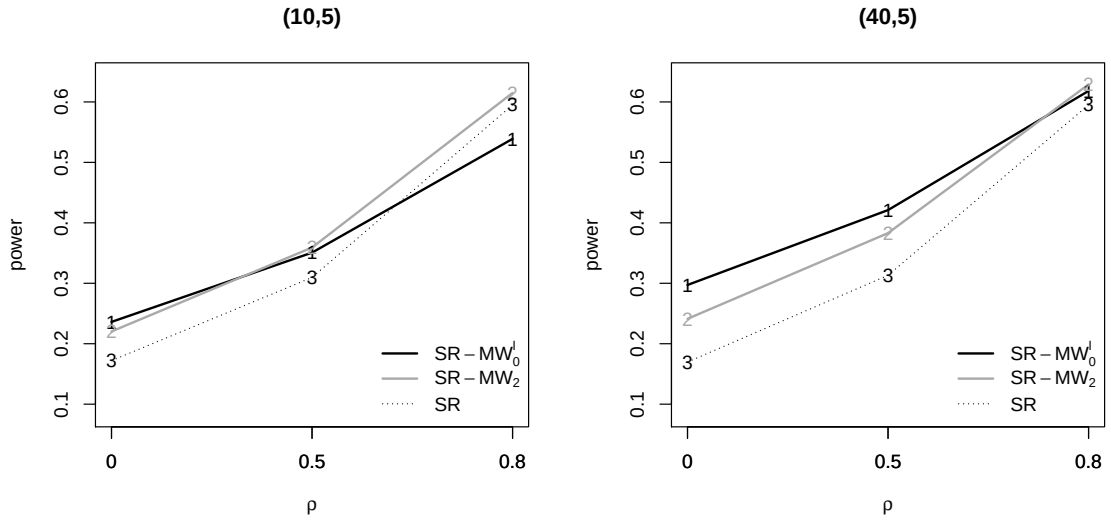
Table B.1: Estimated power, logistic distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			SR			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	14	18	32	13	18	35	9	15	31	12	17	33	13	18	29
(10,10)	15	19	34	13	19	35	9	15	31	13	18	34	14	18	30
(40,5)	16	21	35	14	19	35	9	15	31	12	18	33	16	21	33

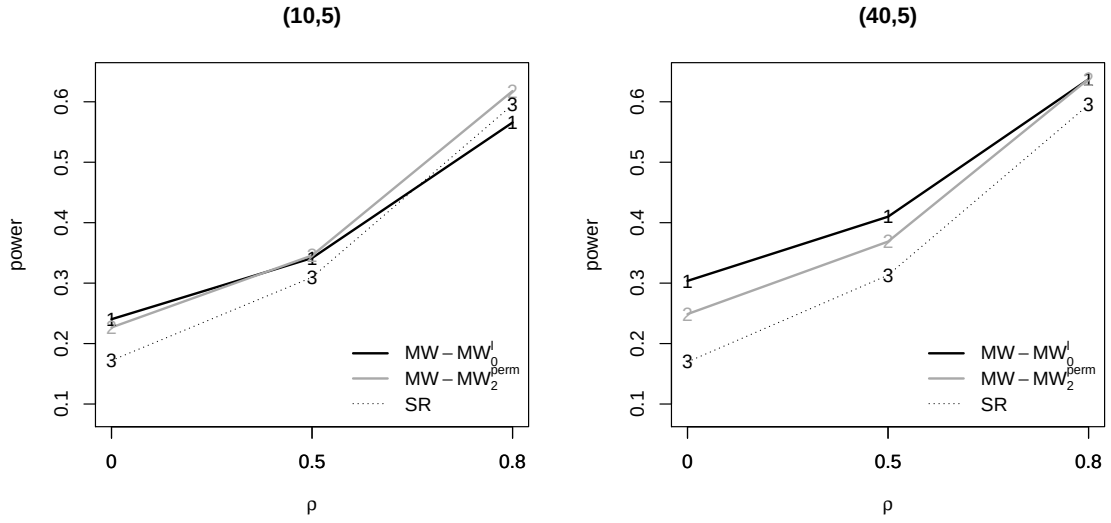
Table B.2: Estimated power, gamma distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			MW-MW$_2^{perm}$			SR			SR-MW$_2$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,5)	25	39	64	23	40	68	12	23	47	16	29	51	21	32	50
(10,10)	28	41	65	25	42	69	12	23	47	19	31	53	24	35	51
(40,5)	31	46	71	25	42	70	12	23	47	17	30	53	27	41	62

Table B.3: Estimated power, lognormal distribution, $m = 20$.

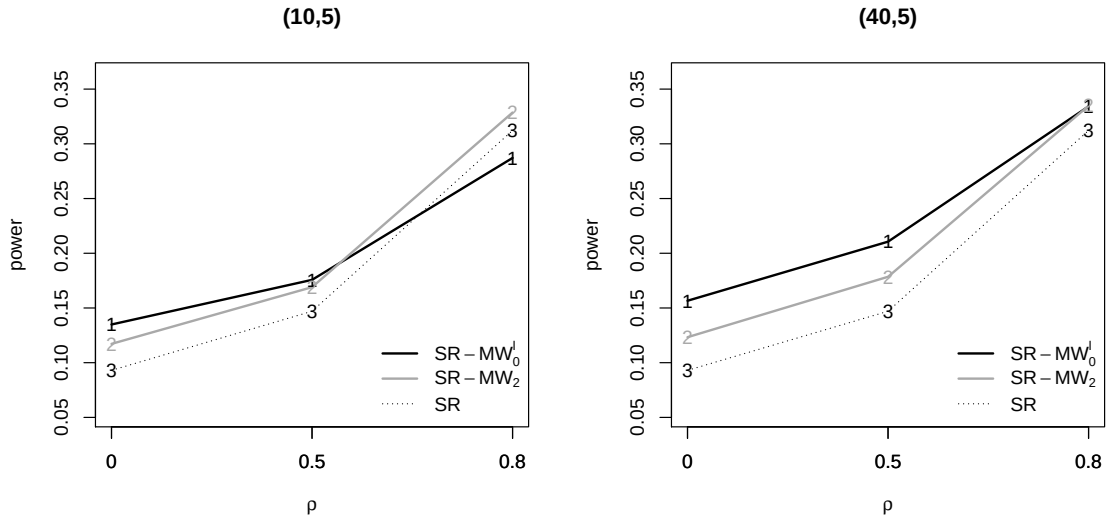


(a) Comparison between SR-MW$_0^l$, SR-MW$_2$ and SR.

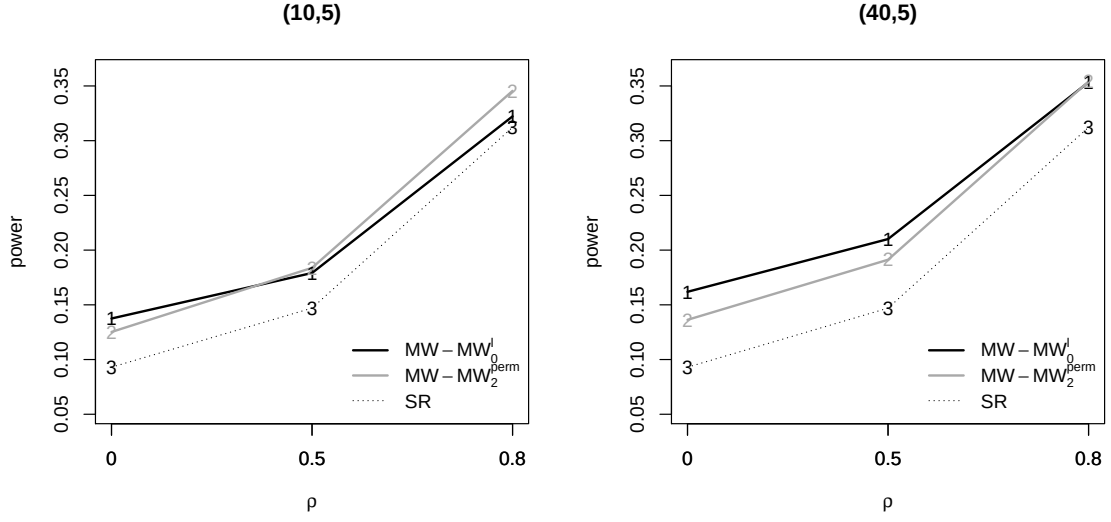


(b) Comparison between MW-MW$_0^l$, MW-MW$_2$ and SR.

Figure B.1: Power comparison when the marginal distribution is logistic. The number of paired observations, m, is 20 pairs, the numbers of unpaired data, (l, n), are given in the titles.

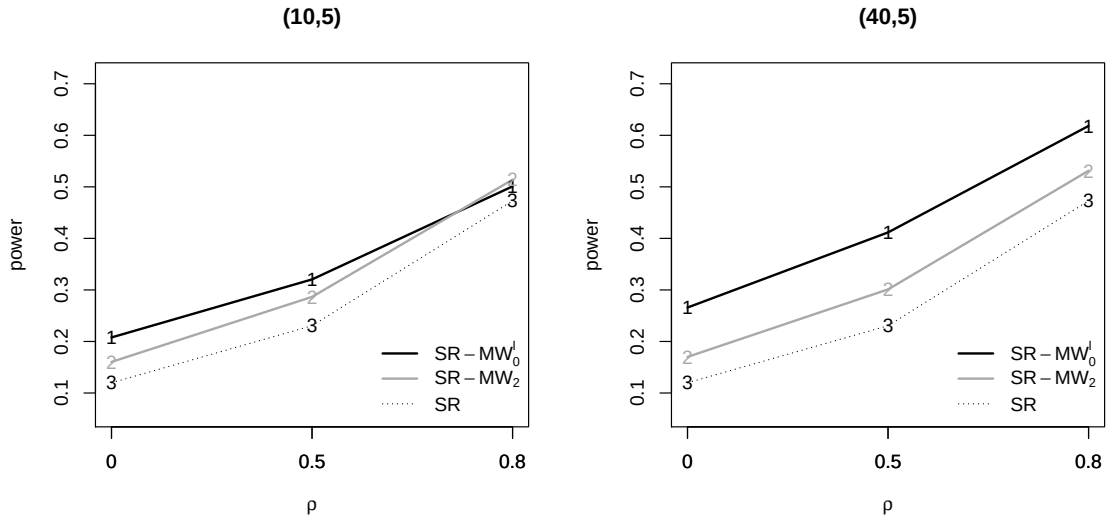


(a) Comparison between SR-MW$_0^l$, SR-MW$_2$ and SR.

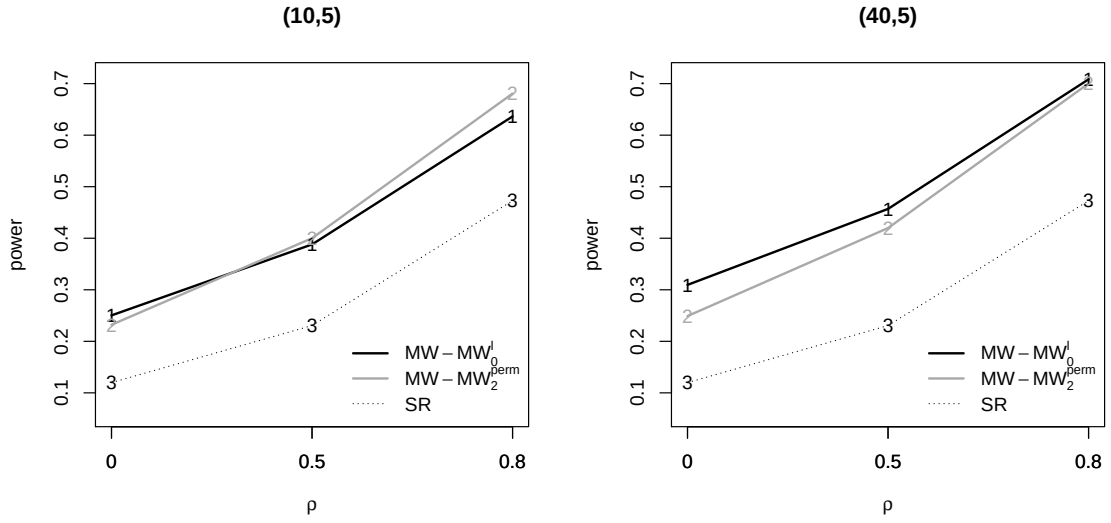


(b) Comparison between MW-MW$_0^l$, MW-MW$_2$ and SR.

Figure B.2: Power comparison when the marginal distribution is gamma. The number of paired observations, m, is 20 pairs, the numbers of unpaired data, (l, n), are given in the titles.



(a) Comparison between SR-MW$_0^l$, SR-MW$_2$ and SR.



(b) Comparison between MW-MW$_0^l$, MW-MW$_2^{perm}$ and SR.

Figure B.3: Power comparison when the marginal distribution is lognormal. Sample sizes: $m = 20$ and (l, n) are given in the titles.

C Comparison of tests when all data are paired

ρ	0	0.5	0.8
MW	0.18	0.36	0.66
SR	0.12	0.23	0.47

Table C.1: Estimated power, lognormal distribution. Paired data only, $m = 20$. For MW, p values are obtained from permutation-based reference distributions.

ρ	0	0.5	0.8
MW	0.14	0.23	0.47
SR	0.14	0.23	0.51

Table C.2: Estimated power, normal distribution. Paired data only, $m = 20$. For MW, p values are obtained from permutation-based reference distributions.

D Comparison of tests when there are only unpaired observations from one sample

(l,n)	MW-MW$_0^l$			SR			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,0)	18	26	48	14	23	51	16	21	32
(40,0)	21	29	50	14	23	51	20	29	46

Table D.1: Estimated power, normal distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			SR			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,0)	23	34	59	17	31	60	19	26	39
(40,0)	27	37	62	17	31	60	26	37	56

Table D.2: Estimated power, logistic distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			SR			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8	0	0.5	0.8
(10,0)	13	18	33	9	15	31	11	14	21
(40,0)	15	20	35	9	15	31	13	18	30

Table D.3: Estimated power, gamma distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8
(10,0)	5.8	5.5	5.0	5.1	5.1	5.1
(40,0)	5.4	5.1	5.1	5.0	5.0	5.1

Table D.4: Estimated size, normal distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8
(10,0)	5.7	5.6	5.0	4.9	4.8	5.1
(40,0)	5.7	5.1	5.1	5.2	5.0	5.1

Table D.5: Estimated size, logistic distribution, $m = 20$.

(l,n)	MW-MW$_0^l$			SR-MW$_0^l$		
	0	0.5	0.8	0	0.5	0.8
(10,0)	5.8	5.5	5.0	5.0	5.1	5.2
(40,0)	5.4	5.1	5.1	5.1	5.1	5.1

Table D.6: Estimated size, gamma distribution, $m = 20$.

References

Fong, Youyi, Huang, Ying, Lemos, Maria P, & Mcelrath, M Juliana. (2018). Rank-based two-sample tests for paired data with missing values. *Biostatistics*, **19**, 281–294.

Guo, Xu, Gao, Yujing, Niu, Cuizhen, & Zhang, Shumei. (2018). Letter to the editor. *Biostatistics*, kxy047.