Dealing with Artificially Dichotomized Variables in Meta-Analytic Structural Equation Modeling

H. de Jonge, S. Jak, & K.J. Kan

Electronic Supplementary Material 3

Tables with additional results of simulation study 1 (full mediation)

Table 1

Simulation study 1 (full mediation): Percentages estimation bias in the direct effects and their standard errors when the predictor variable X was not dichotomized at all and the Pearson product-moment correlation was used

Condition							
DICH	CO	ES	Converged[1]	Bias in β_{MX}	Bias in SE of β_{MX}	Bias in β_{YM}	Bias in SE of β_{YM}
25	.1	.16	2000	-0.298	-0.956	0.032	-4.144
		.23	2000	-0.228	-2.831	-0.450	-2.584
		.33	2000	-0.085	-2.939	0.034	-3.214
	.5	.16	2000	-0.243	-1.132	-0.042	-1.377
		.23	2000	-0.227	-1.826	-0.052	-3.248
		.33	2000	-0.228	-0.014	-0.017	-2.387
75	.1	.16	2000	0.013	-4.899	-0.210	-1.846
		.23	2000	0.018	-2.124	-0.036	-3.169
		.33	2000	-0.233	-1.571	-0.206	-2.702
	.5	.16	2000	0.049	0.307	-0.212	-2.241
		.23	2000	-0.019	-1.922	-0.383	-0.154
		.33	2000	-0.002	-2.256	-0.219	-2.966

[1] The number of converged datasets between the different tables can differ because we ran the simulation twice. Once to get the standard errors and associated confidence intervals based on the Wald test and once to get the likelihood-based confidence intervals (see R-scripts).

Table 1 (Continued)

| Condition | | | | | | | |
DICH	CO	ES	Converged	Bias in β_{MX}	Bias in *SE* of β_{MX}	Bias in β_{YM}	Bias in *SE* of β_{YM}
100	.1	.16	2000	-0.052	-2.229	-0.329	-3.037
		.23	2000	0.146	-2.676	-0.335	-2.956
		.33	2000	-0.174	0.856	-0.187	-1.457
	.5	.16	2000	-0.077	-1.906	0.061	-3.578
		.23	2000	-0.064	-4.793	-0.191	-0.747
		.33	2000	-0.110	-1.771	0.127	-0.521

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; Bias in β_{MX} = relative percentage bias in the path coefficient between X and M; Bias in *SE* of β_{MX} = relative percentage bias in the standard error of the path coefficient between X and M; Bias in β_{YM} = relative percentage bias in the path coefficient between M and Y; Bias in *SE* of β_{YM} = relative percentage bias in the standard error of the path coefficient between M and Y.

Note 2. DICH and CO are only labels because in this case X was not dichotomized at all.

Table 2

Simulation study 1 (full mediation): Percentages estimation bias in the indirect effect and the coverage percentages of the 95% likelihood-based confidence interval when the predictor variable X was not dichotomized at all and the Pearson product-moment correlation was used

Condition					
DICH	CO	ES	Converged[1]	Bias in indirect	Coverage
25	.1	.16	2000	-0.249	94.250
		.23	2000	-0.679	93.950
		.33	2000	-0.068	93.750
	.5	.16	2000	-0.311	95.050
		.23	1999	-0.290	94.447
		.33	2000	-0.268	94.400
75	.1	.16	2000	-0.217	93.450
		.23	2000	-0.050	94.500
		.33	1998	-0.476	94.595
	.5	.16	2000	-0.169	94.650
		.23	2000	-0.436	95.050
		.33	1999	-0.241	94.097
100	.1	.16	2000	-0.386	93.750
		.23	2000	-0.207	94.500
		.33	2000	-0.383	94.950
	.5	.16	2000	-0.032	94.200
		.23	2000	-0.285	95.150
		.33	1999	-0.003	94.547

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; Bias in indirect = relative percentage bias in the indirect effect of X on Y ($\beta_{MX} * \beta_{YM}$); Coverage = percentage of confidence intervals that includes the population parameter of the indirect effect of X on Y.
Note 2. DICH and CO are only labels because in this case X was not dichotomized at all.

Table 3

Simulation study 1 (full mediation): Coverage percentages of the 95% likelihood-based confidence intervals of the direct effects

Condition			Converged[1]			Coverage β_{MX}			Coverage β_{YM}		
DICH	CO	ES	r	r_{pb}	r_b	r	r_{pb}	r_b	r	r_{pb}	r_b
25	.1	.16	2000	2000	2000	94.150	83.350	93.250	93.800	91.800	92.200
		.23	2000	2000	2000	94.300	76.600	92.850	93.900	91.900	92.300
		.33	2000	2000	2000	93.750	60.950	93.350	93.350	92.800	92.700
	.5	.16	2000	2000	1999	94.900	92.100	94.797	93.950	92.350	92.246
		.23	1999	2000	2000	94.197	89.200	94.050	93.897	91.650	91.650
		.33	2000	2000	2000	94.700	83.650	93.350	93.800	92.150	92.200
75	.1	.16	2000	1999	2000	93.600	15.408	92.450	94.000	93.297	93.050
		.23	2000	2000	1999	93.150	1.950	93.597	93.350	92.850	92.546
		.33	1998	2000	2000	94.294	0.050	93.850	94.244	93.300	93.250
	.5	.16	2000	2000	1999	94.400	71.650	93.697	93.550	92.300	92.346
		.23	2000	2000	2000	93.100	50.550	92.750	94.850	93.100	93.150
		.33	1999	2000	2000	93.847	21.900	94.300	93.547	92.800	92.600
100	.1	.16	2000	1994	2000	93.850	0.401	93.400	93.850	92.277	92.200
		.23	2000	1995	2000	93.450	0.000	93.400	93.750	93.383	93.050
		.33	2000	1990	1999	94.800	0.000	93.897	93.700	92.915	92.896
	.5	.16	2000	2000	2000	94.350	51.050	93.500	93.350	91.150	91.250
		.23	2000	1999	1999	93.400	22.461	92.696	94.850	91.846	91.846
		.33	1999	2000	2000	93.347	2.000	93.400	94.597	92.500	92.600

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; Coverage β_{MX} = percentage of confidence intervals that includes the population parameter of the direct effect between X and M; Coverage β_{YM} = percentage of confidence intervals that includes the population parameter of the direct effect between M and Y; r = Pearson product-moment correlation; r_{pb} = point-biserial correlation; r_b = biserial correlation.
Note 2. For r, DICH and CO are only labels because in this case X was not dichotomized at all.

Table 4

Simulation study 1 (full mediation): Coverage percentages of the 95% Wald confidence intervals of the direct effects

Condition			Converged[1]			Coverage β_{MX}			Coverage β_{YM}		
DICH	CO	ES	r	r_{pb}	r_b	r	r_{pb}	r_b	r	r_{pb}	r_b
25	.1	.16	2000	2000	2000	94.150	83.350	93.250	93.800	91.800	92.200
		.23	2000	2000	2000	94.300	76.550	92.800	93.900	91.950	92.250
		.33	2000	2000	2000	93.750	60.900	93.300	93.350	92.800	92.700
	.5	.16	2000	2000	1999	94.900	92.100	94.847	93.950	92.350	92.246
		.23	2000	2000	2000	94.200	89.200	94.050	93.900	91.650	91.650
		.33	2000	2000	2000	94.700	83.650	93.350	93.800	92.150	92.200
75	.1	.16	2000	1999	2000	93.600	15.408	92.450	94.000	93.297	93.050
		.23	2000	2000	1999	93.150	1.900	93.647	93.350	92.900	92.596
		.33	2000	2000	2000	94.250	0.050	93.900	94.250	93.250	93.250
	.5	.16	2000	2000	1999	94.400	71.600	93.697	93.600	92.250	92.346
		.23	2000	2000	2000	93.100	50.450	92.750	94.850	93.050	93.200
		.33	2000	2000	2000	93.850	21.800	94.300	93.500	92.800	92.600
100	.1	.16	2000	1994	2000	93.800	0.401	93.450	93.900	92.277	92.200
		.23	2000	1995	2000	93.450	0.000	93.400	93.750	93.383	93.050
		.33	2000	1990	1999	94.800	0.000	93.897	93.800	92.864	92.946
	.5	.16	2000	2000	2000	94.350	51.050	93.500	93.350	91.150	91.150
		.23	2000	1999	1999	93.400	22.411	92.696	94.850	91.896	91.846
		.33	2000	2000	2000	93.350	2.000	93.400	94.600	92.500	92.600

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; Coverage β_{MX} = percentage of confidence intervals that includes the population parameter of the direct effect between X and M; Coverage β_{YM} = percentage of confidence intervals that includes the population parameter of the direct effect between M and Y; r = Pearson product-moment correlation; r_{pb} = point-biserial correlation; r_b = biserial correlation.
Note 2. For r, DICH and CO are only labels because in this case X was not dichotomized at all.

Table 5

Simulation study 1 (full mediation): Rejection rates of chi-square test of model fit at Stage 2 of the random-effects TSSEM

Condition			Converged[1]			Rejection rate		
DICH	CO	ES	r	r_{pb}	r_b	r	r_{pb}	r_b
25	.1	.16	2000	2000	2000	**.065**	**.062**	.060
		.23	2000	2000	2000	**.065**	**.072**	**.066**
		.33	2000	2000	2000	**.063**	.056	.058
	.5	.16	2000	2000	1999	**.062**	.056	.060
		.23	2000	2000	2000	**.063**	.058	.058
		.33	2000	2000	2000	.060	.058	.058
75	.1	.16	2000	1999	2000	**.062**	.060	**.066**
		.23	2000	2000	1999	.047	.050	.049
		.33	2000	2000	2000	.051	.044	.052
	.5	.16	2000	2000	1999	**.065**	**.068**	**.067**
		.23	2000	2000	2000	.052	.046	.046
		.33	2000	2000	2000	.056	.054	.052
100	.1	.16	2000	1994	2000	.058	.057	.054
		.23	2000	1995	2000	**.060**	.058	.050
		.33	2000	1990	1999	**.062**	**.070**	**.068**
	.5	.16	2000	2000	2000	.050	.055	.054
		.23	2000	1999	1999	.058	**.069**	**.067**
		.33	2000	2000	2000	.056	.048	.048

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; r = Pearson product-moment correlation; r_{pb} = point-biserial correlation; r_b = biserial correlation.
Note 2. Bold figures indicate a significant difference of the rejection rate and the nominal α-level (.05) indicated by the proportion test (using α = .05).
Note 3. For r, DICH and CO are only labels because in this case X was not dichotomized at all.

Table 6

Simulation study 1 (full mediation): Results of the Kolmogorov-Smirnov test

Condition			Converged[1]			D statistic			p value		
DICH	CO	ES	r	r_{pb}	r_b	r	r_{pb}	r_b	r	r_{pb}	r_b
25	.1	.16	2000	2000	2000	.017	.035	.030	.590	**.014**	.056
		.23	2000	2000	2000	.027	.028	.024	.098	.094	.208
		.33	2000	2000	2000	.027	.022	.024	.098	.307	.204
	.5	.16	2000	2000	1999	.019	.032	.032	.460	**.037**	**.030**
		.23	2000	2000	2000	.025	.032	.033	.168	**.031**	**.024**
		.33	2000	2000	2000	.013	.015	.022	.907	.746	.290
75	.1	.16	2000	1999	2000	.021	.024	.018	.334	.214	.546
		.23	2000	2000	1999	.017	.015	.014	.621	.726	.854
		.33	2000	2000	2000	.023	.032	.022	.259	**.030**	.305
	.5	.16	2000	2000	1999	.027	.032	.032	.119	**.032**	**.037**
		.23	2000	2000	2000	.012	.019	.025	.932	.452	.155
		.33	2000	2000	2000	.012	.020	.021	.937	.378	.321
100	.1	.16	2000	1994	2000	.023	.031	.020	.230	**.045**	.407
		.23	2000	1995	2000	.028	.022	.027	.094	.288	.107
		.33	2000	1990	1999	.038	.039	.038	**.006**	**.005**	**.006**
	.5	.16	2000	2000	2000	.023	.015	.018	.222	.739	.561
		.23	2000	1999	1999	.040	.036	.037	**.003**	**.011**	**.009**
		.33	2000	2000	2000	.020	.024	.023	.430	.218	.231

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; r = Pearson product-moment correlation; r_{pb} = point-biserial correlation; r_b = biserial correlation.
Note 2. Bold figures indicate a significant difference between the theoretical chi-square distribution with $df = 1$ with the empirical chi-square distribution indicated by the Kolmogorov-Smirnov test (using $\alpha = .05$).
Note 3. For r, DICH and CO are only labels because in this case X was not dichotomized at all.

Table 7

Simulation study 1 (full mediation): Percentage estimation bias in the pooled correlations[2] at

Stage 1 of the random-effects TSSEM

Condition			Converged[1]			Bias in pooled ES$_{MX}$			Bias in pooled ES$_{YM}$		
DICH	CO	ES	r_{pb}	r_b	r	r_{pb}	r_b	r	r_{pb}	r_b	r
25	.1	.16	2000	2000	2000	-10.546	-0.127	-0.274	-0.276	-0.277	0.030
		.23	2000	2000	2000	-10.478	-0.130	-0.231	-0.451	-0.451	-0.442
		.33	2000	2000	2000	-10.675	-0.325	-0.150	0.098	0.096	-0.103
	.5	.16	2000	1999	2000	-5.068	0.027	-0.216	0.016	0.017	-0.051
		.23	2000	2000	2000	-5.258	-0.219	-0.270	0.009	0.010	-0.074
		.33	2000	2000	2000	-5.400	-0.339	-0.251	-0.167	-0.169	-0.055
75	.1	.16	1999	2000	2000	-31.456	-0.358	-0.058	0.044	0.041	-0.245
		.23	2000	1999	2000	-31.286	-0.138	0.034	0.403	0.403	-0.021
		.33	2000	2000	2000	-31.202	-0.066	-0.266	0.096	0.101	-0.247
	.5	.16	2000	1999	2000	-15.062	0.085	0.028	-0.373	-0.371	-0.235
		.23	2000	2000	2000	-15.218	-0.090	-0.039	0.044	0.044	-0.416
		.33	2000	2000	2000	-15.198	-0.011	-0.011	-0.442	-0.441	-0.232
100	.1	.16	1994	2000	2000	-41.327	0.281	-0.034	-0.518	-0.529	-0.323
		.23	1995	2000	2000	-41.504	-0.093	0.113	-0.336	-0.339	-0.372
		.33	1990	1999	2000	-41.714	-0.365	-0.207	-0.119	-0.133	-0.234
	.5	.16	2000	2000	2000	-20.032	0.202	-0.047	0.123	0.123	0.073
		.23	1999	1999	2000	-20.196	0.014	-0.049	-0.153	-0.153	-0.178
		.33	2000	2000	2000	-20.492	-0.361	-0.102	0.027	0.027	0.154

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; r_{pb} = point-biserial correlation; r_b = biserial correlation; r = Pearson product-moment correlation; Bias in pooled ES$_{MX}$ = relative percentage bias in the pooled correlation between X and M at Stage 1; Bias in pooled ES$_{YM}$ = relative percentage bias in the pooled correlation between M and Y at Stage 1. *Note 2.* For r, DICH and CO are only labels because in this case X was not dichotomized at all.

---

[2] In simulation study 1, we focused on the relationships between X and M and between M and Y. Therefore, we only reported the bias in the pooled ES$_{MX}$, the pooled ES$_{YM}$, and their standard errors.

Table 8

Simulation study 1 (full mediation): Percentage estimation bias in the standard errors of the

pooled correlations[2] at Stage 1 of the random-effects TSSEM

Condition			Converged[1]			Bias in SE of pooled ES_{MX}			Bias in SE of pooled ES_{YM}		
DICH	CO	ES	r_{pb}	r_b	r	r_{pb}	r_b	r	r_{pb}	r_b	r
25	.1	.16	2000	2000	2000	0.507	-0.977	0.290	-5.738	-5.726	-3.917
		.23	2000	2000	2000	0.864	-4.442	-3.239	-4.683	-4.699	-2.522
		.33	2000	2000	2000	5.714	-2.701	-2.822	-4.404	-4.369	-3.031
	.5	.16	2000	1999	2000	-0.464	-0.930	0.072	-6.024	-6.022	-1.387
		.23	2000	2000	2000	-0.539	-2.058	-1.425	-7.485	-7.478	-2.466
		.33	2000	2000	2000	1.344	-0.485	0.735	-5.850	-5.826	-1.959
75	.1	.16	1999	2000	2000	-2.319	-4.821	-4.844	-3.981	-3.960	-1.997
		.23	2000	1999	2000	2.533	-2.775	-1.453	-3.436	-3.461	-2.908
		.33	2000	2000	2000	11.627	-1.929	-1.554	-2.948	-2.946	-1.446
	.5	.16	2000	1999	2000	-1.586	-2.180	0.093	-6.330	-6.327	-2.215
		.23	2000	2000	2000	-1.085	-2.861	-1.035	-2.660	-2.664	-0.683
		.33	2000	2000	2000	0.365	-1.165	-1.897	-5.082	-5.085	-2.206
100	.1	.16	1994	2000	2000	-2.565	-2.898	-1.869	-4.372	-4.534	-2.923
		.23	1995	2000	2000	-3.075	-2.647	-2.085	-4.204	-4.438	-2.343
		.33	1990	1999	2000	-0.696	-0.681	0.790	-4.971	-5.078	-0.222
	.5	.16	2000	2000	2000	-1.631	-1.437	-1.513	-8.166	-8.166	-2.822
		.23	1999	1999	2000	-5.820	-5.484	-4.774	-6.398	-6.400	-0.228
		.33	2000	2000	2000	-2.838	-2.791	-1.049	-4.512	-4.511	-0.273

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; r_{pb} = point-biserial; r_b = biserial correlation; r = Pearson product-moment correlation; Bias in SE of pooled ES_{MX}= relative percentage bias in the standard error of the pooled correlation between X and M at Stage 1; Bias in SE of pooled ES_{YM}= relative percentage bias in the standard error of the pooled correlation between M and Y at Stage 1.
Note 2. For r, DICH and CO are only labels because in this case X was not dichotomized at all.

Table 9

Simulation study 1 (full mediation): Percentage estimation bias in the between-studies variances[3] at Stage 1 of the random-effects TSSEM

Condition			Converged[1]			Bias in τ^2_{MX}			Bias in τ^2_{YM}		
DICH	CO	ES	r_{pb}	r_b	r	r_{pb}	r_b	r	r_{pb}	r_b	r
25	.1	.16	2000	2000	2000	-13.250	7.252	-2.995	-9.994	-9.999	-3.569
		.23	2000	2000	2000	-6.233	5.043	-4.133	-9.820	-9.835	-3.425
		.33	2000	2000	2000	11.176	3.725	-4.054	-9.307	-9.285	-2.727
	.5	.16	2000	1999	2000	-13.281	-2.803	-3.427	-10.077	-10.036	-3.957
		.23	2000	2000	2000	-10.636	-2.183	-3.273	-9.112	-9.114	-2.598
		.33	2000	2000	2000	-8.442	-4.103	-2.913	-10.040	-10.030	-3.159
75	.1	.16	1999	2000	2000	-45.042	33.865	-2.644	-7.818	-7.820	-2.586
		.23	2000	1999	2000	-36.791	31.849	-2.334	-9.799	-9.789	-3.590
		.33	2000	2000	2000	-20.357	28.234	-2.716	-8.998	-8.968	-2.571
	.5	.16	2000	1999	2000	-31.007	3.956	-3.579	-9.457	-9.411	-3.866
		.23	2000	2000	2000	-29.262	3.541	-3.217	-10.603	-10.591	-3.462
		.33	2000	2000	2000	-26.525	1.627	-3.789	-10.456	-10.453	-4.035
100	.1	.16	1994	2000	2000	-69.375	43.167	-3.857	-8.965	-9.017	-2.430
		.23	1995	2000	2000	-69.270	43.454	-3.534	-10.416	-10.475	-3.708
		.33	1990	1999	2000	-70.674	38.693	-4.071	-9.806	-9.880	-3.629
	.5	.16	2000	2000	2000	-40.694	9.090	-3.938	-11.641	-11.639	-4.767
		.23	1999	1999	2000	-41.788	7.351	-4.086	-9.168	-9.168	-2.827
		.33	2000	2000	2000	-42.358	6.183	-2.681	-10.098	-10.096	-2.921

Note 1. DICH = percentage of primary studies in which X was artificially dichotomized; CO = cut-off point at which X was artificially dichotomized; ES = size of the systematically varied (standardized) path coefficient between X and M; Converged = number of datasets that converged in Stage 1 and Stage 2 of the random-effects TSSEM; r_{pb} = point-biserial correlation; r_b = biserial correlation; r = Pearson product-moment correlation; Bias in τ^2_{MX} = relative percentage bias in the between-study variance of the correlation coefficient between X and M at Stage 1; Bias in τ^2_{YM} = relative percentage bias in the between-study variance of correlation coefficient between M and Y at Stage 1.

Note 2. For r, DICH and CO are only labels because in this case X was not dichotomized at al.

---

[3] In simulation study 1, we focused on the relationships between X and M and between M and Y. Therefore, we only reported the bias in τ^2_{MX} and τ^2_{YM}, and not in τ^2_{YX}.