Biometrika (2011), **xx**, x, *pp*. 1–5

Printed in Great Britain

Supplementary materials for multiple imputation in quantile regression

BY YING WEI

Department of Biostatistics, Columbia University, 722 West 168th St., New York, NY 10032
ying.wei@columbia.edu

YANYUAN MA AND RAYMOND J. CARROLL

Department of Statistics, Texas A&M University, TAMU 3143, College Station, TX 77843-3143
ma@stat.tamu.edu carroll@stat.tamu.edu

We present the supplementary simulation results in the following settings:

Setting 1: Observation (Y, X, Z) follows Models (4) and (5) in the section of Simulation respectively. The (X, Z) correlation is 0.5. A heavier missing rate is introduced, i.e., 35% Xs are missing. The estimated coefficients and their standard errors from the "complete-data-estimator", multiple imputation estimator and shrinkage estimator are presented in Table 1.

Setting 2: Observation (Y, X, Z) follows Models (4) and (5), with 25% x missing. A weaker (X, Z) correlation is used, i.e., (X, Z) correlation is 0.3. The estimated coefficients and their standard errors from complete-data-estimator, multiple imputation estimator and shrinkage estimator are presented in Table 2.

Setting 3: Observation (Y, X, Z) follows Models (4) and (5), with 25% x missing. A stronger (X, Z) correlation is used, i.e., (X, Z) correlation is 0.7. The estimated coefficients and their standard errors from complete-data-estimator, multiple imputation estimator and shrinkage estimator are presented in Table 3.

			$\tau = 0.1$			$\tau = 0.5$			$\tau = 0.9$		
			True	Mean	S.E.	True	Mean	S.E.	True	Mean	S.E.
	intercept	$\widehat{\beta}$	-0.28	-0.33	0.88	1.00	0.98	0.71	2.28	2.22	0.96
M		$\widetilde{\beta}$	-0.28	-0.13	0.71	1.00	1.13	0.55	2.28	2.48	0.81
O		$\widehat{\beta}^{(s)}$	-0.28	-0.33	0.75	1.00	1.03	0.62	2.28	2.36	0.83
D	x	$\widehat{\beta}$	1.00	0.99	0.16	1.00	1.00	0.12	1.00	1.00	0.16
E		$\widetilde{\beta}$	1.00	1.00	0.17	1.00	1.00	0.14	1.00	1.01	0.16
L		$\widehat{\beta}^{(s)}$	1.00	0.91	0.18	1.00	0.93	0.15	1.00	0.90	0.19
(4)	z	$\widehat{\beta}$	1.00	1.01	0.20	1.00	1.00	0.15	1.00	1.00	0.21
		$\widetilde{\beta}$	1.00	1.05	0.17	1.00	1.04	0.12	1.00	1.03	0.17
		$\widehat{\beta}^{(s)}$	1.00	1.03	0.18	1.00	1.02	0.13	1.00	1.02	0.19
	intercept	$\widehat{\beta}$	1.00	0.79	3.88	1.00	0.99	2.69	1.00	0.72	4.07
M		$\widetilde{\beta}$	1.00	1.10	2.20	1.00	1.11	1.79	1.00	1.74	2.25
O		$\widehat{\beta}^{(s)}$	1.00	0.85	3.16	1.00	1.03	2.24	1.00	1.34	3.25
D	x	$\widehat{\beta}$	0.36	0.37	0.82	1.00	1.01	0.55	1.64	1.68	0.69
E		$\widetilde{\beta}$	0.36	0.34	0.64	1.00	0.85	0.49	1.64	1.34	0.60
L		$\widehat{\beta}^{(s)}$	0.36	0.37	0.76	1.00	0.98	0.53	1.64	1.58	0.67
(5)	z	$\widehat{\beta}$	0.36	0.41	0.87	1.00	1.00	0.59	1.64	1.67	1.03
		$\widetilde{\beta}$	0.36	0.37	0.54	1.00	1.11	0.38	1.64	1.70	0.60
		$\widehat{\beta}^{(s)}$	0.36	0.41	0.70	1.00	1.02	0.49	1.64	1.67	0.85

Table 1. *Means and standard errors of the estimated coefficients at quantile levels 0.1, 0.5 and 0.9 from 500 simulations in Models (4) and (5), with X missing 35% of the time. Here $\widehat{\beta}$ stands for the estimated coefficients using the completely observed data only, $\widetilde{\beta}$ is multiple imputation estimator with 10 imputations; and $\widehat{\beta}^{(s)}$ is the shrinkage estimator. The true coefficients are listed under the "True"columns.*

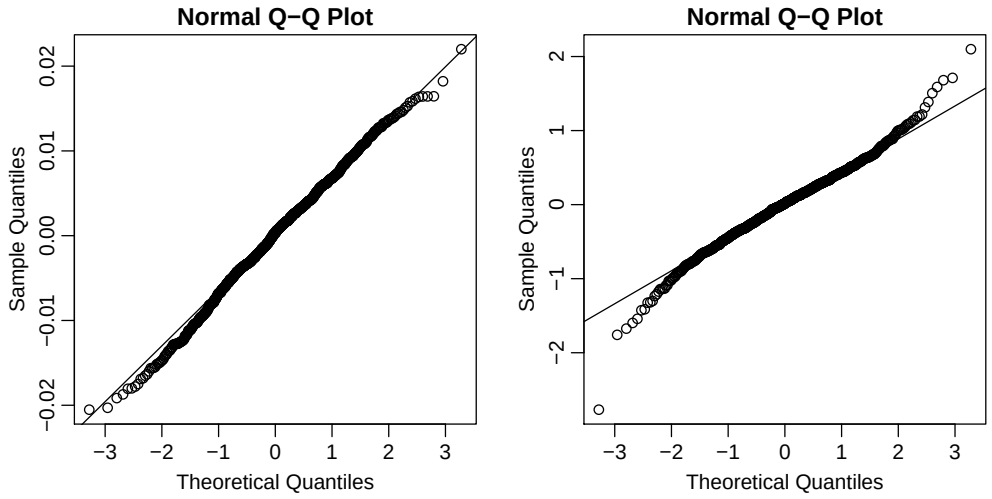
			$\tau = 0.1$			$\tau = 0.5$			$\tau = 0.9$		
			True	Mean	S.E.	True	Mean	S.E.	True	Mean	S.E.
	intercept	$\widehat{\beta}$	-0.28	-0.31	0.89	1.00	1.01	0.62	2.28	2.17	0.84
M		$\widetilde{\beta}$	-0.28	-0.14	0.71	1.00	1.04	0.54	2.28	2.70	0.73
O		$\widehat{\beta}^{(s)}$	-0.28	-0.28	0.25	1.00	1.02	0.56	2.28	2.32	0.82
D	x	$\widehat{\beta}$	1.00	0.99	0.16	1.00	0.99	0.11	1.00	1.00	0.14
E		$\widetilde{\beta}$	1.00	0.93	0.14	1.00	0.96	0.11	1.00	0.96	0.16
L		$\widehat{\beta}^{(s)}$	1.00	0.98	0.15	1.00	0.99	0.11	1.00	0.99	0.15
(4)	z	$\widehat{\beta}$	1.00	1.01	0.17	1.00	1.00	0.12	1.00	1.02	0.18
		$\widetilde{\beta}$	1.00	1.02	0.14	1.00	1.01	0.10	1.00	0.98	0.16
		$\widehat{\beta}^{(s)}$	1.00	1.02	0.15	1.00	1.01	0.10	1.00	1.00	0.17
	intercept	$\widehat{\beta}$	1.00	0.96	3.73	1.00	1.12	2.33	1.00	0.87	4.02
M		$\widetilde{\beta}$	1.00	1.06	2.41	1.00	1.01	1.69	1.00	1.55	2.53
O		$\widehat{\beta}^{(s)}$	1.00	0.84	3.20	1.00	1.08	1.99	1.00	1.21	3.44
D	x	$\widehat{\beta}$	0.36	0.43	0.62	1.00	1.01	0.51	1.64	1.68	0.62
E		$\widetilde{\beta}$	0.36	0.39	0.52	1.00	0.92	0.46	1.64	1.45	0.53
L		$\widehat{\beta}^{(s)}$	0.36	0.42	0.58	1.00	0.98	0.49	1.64	1.62	0.60
(5)	z	$\widehat{\beta}$	0.36	0.37	0.64	1.00	1.04	0.44	1.64	1.61	0.58
		$\widetilde{\beta}$	0.36	0.33	0.52	1.00	0.95	0.41	1.64	1.45	0.55
		$\widehat{\beta}^{(s)}$	0.36	0.36	0.61	1.00	1.01	0.42	1.64	1.57	0.56

Table 2. *Means and standard errors of the estimated coefficients at quantile levels 0.1, 0.5 and 0.9 from 500 simulations in Models (4) and (5) with (x, z) correlation 0.3. Here $\widehat{\beta}$ stands for the estimated coefficient using the completely observed data only, $\widetilde{\beta}$ is multiple imputation estimator with 10 imputations; and $\widehat{\beta}^{(s)}$ is the shrinkage estimator. The true coefficients are listed under the "True"columns.*

			$\tau = 0.1$			$\tau = 0.5$			$\tau = 0.9$		
			True	Mean	S.E.	True	Mean	S.E.	True	Mean	S.E.
	intercept	$\widehat{\beta}$	-0.28	-0.27	0.78	1.00	1.02	0.61	2.28	2.32	0.88
M		$\widetilde{\beta}$	-0.28	-0.25	0.61	1.00	1.10	0.46	2.28	2.55	0.69
O		$\widehat{\beta}^{(s)}$	-0.28	-0.27	0.68	1.00	1.04	0.53	2.28	2.40	0.73
D	x	$\widehat{\beta}$	1.00	0.99	0.20	1.00	0.98	0.15	1.00	0.98	0.19
E		$\widetilde{\beta}$	1.00	0.94	0.18	1.00	0.94	0.14	1.00	0.90	0.20
L		$\widehat{\beta}^{(s)}$	1.00	0.98	0.19	1.00	0.98	0.14	1.00	0.97	0.19
(4)	z	$\widehat{\beta}$	1.00	1.01	0.23	1.00	1.01	0.17	1.00	1.00	0.23
		$\widetilde{\beta}$	1.00	1.06	0.20	1.00	1.04	0.14	1.00	1.03	0.19
		$\widehat{\beta}^{(s)}$	1.00	1.03	0.21	1.00	1.02	0.15	1.00	1.01	0.20
	intercept	$\widehat{\beta}$	1.00	0.86	3.49	1.00	1.09	2.27	1.00	0.91	3.66
M		$\widetilde{\beta}$	1.00	0.99	1.92	1.00	1.00	1.34	1.00	1.06	2.09
O		$\widehat{\beta}^{(s)}$	1.00	0.80	2.92	1.00	1.08	1.79	1.00	1.11	3.03
D	x	$\widehat{\beta}$	0.36	0.40	0.84	1.00	1.06	0.63	1.64	1.68	0.78
E		$\widetilde{\beta}$	0.36	0.33	0.67	1.00	0.93	0.57	1.64	1.47	0.68
L		$\widehat{\beta}^{(s)}$	0.36	0.39	0.80	1.00	1.03	0.60	1.64	1.62	0.75
(5)	z	$\widehat{\beta}$	0.36	0.36	0.89	0.36	0.93	0.66	0.36	1.60	1.04
		$\widetilde{\beta}$	0.36	0.41	0.62	0.36	1.06	0.48	0.36	1.78	0.71
		$\widehat{\beta}^{(s)}$	0.36	0.40	0.77	0.36	0.96	0.58	0.36	1.63	0.92

Table 3. *Means and standard errors of the estimated coefficients at quantile levels 0.1, 0.5 and 0.9 from 500 simulations in Models (4) and (5) with (x, z) correlation 0.7. Here $\widehat{\beta}$ stands for the estimated coefficient using the completely observed data only, $\widetilde{\beta}$ is multiple imputation estimator with 10 imputations; and $\widehat{\beta}^{(s)}$ is the shrinkage estimator. The true coefficients are listed under the "True"columns.*

SUPPLEMENTARY PLOT FOR SECTION 5: APPLICATION TO THE EATING AT AMERICAN'S TABLES STUDY



(a) bmi (b) carbohydrate intake

Fig. 1. Quantile-quantile plots. The left figure is quantile-quantile plot of the residuals from the model of carbohydrate intake, while the right one is that of body mass index.