Supporting Information for

Estimating Correlation with Multiply Censored Data Arising from the Adjustment of Singly Censored Data

by

Elizabeth Newton
Ruthann Rudel

Pages S1-S42: Box plots of estimates of correlation.
Page S43: Scatter plots of simulated data.
Pages S44-S45 Example of calculation of Kendall's tau-b.

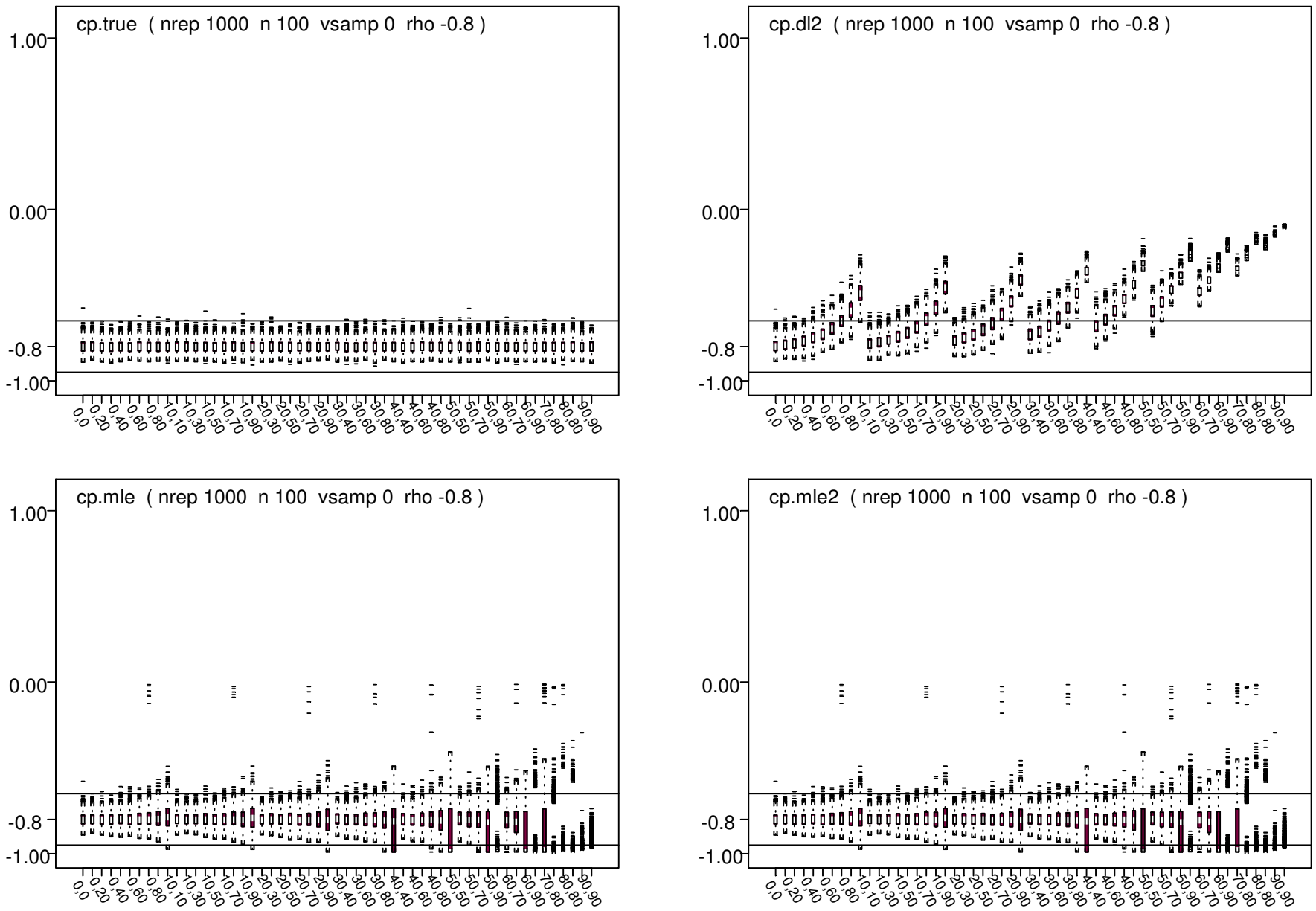

Figure S1. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

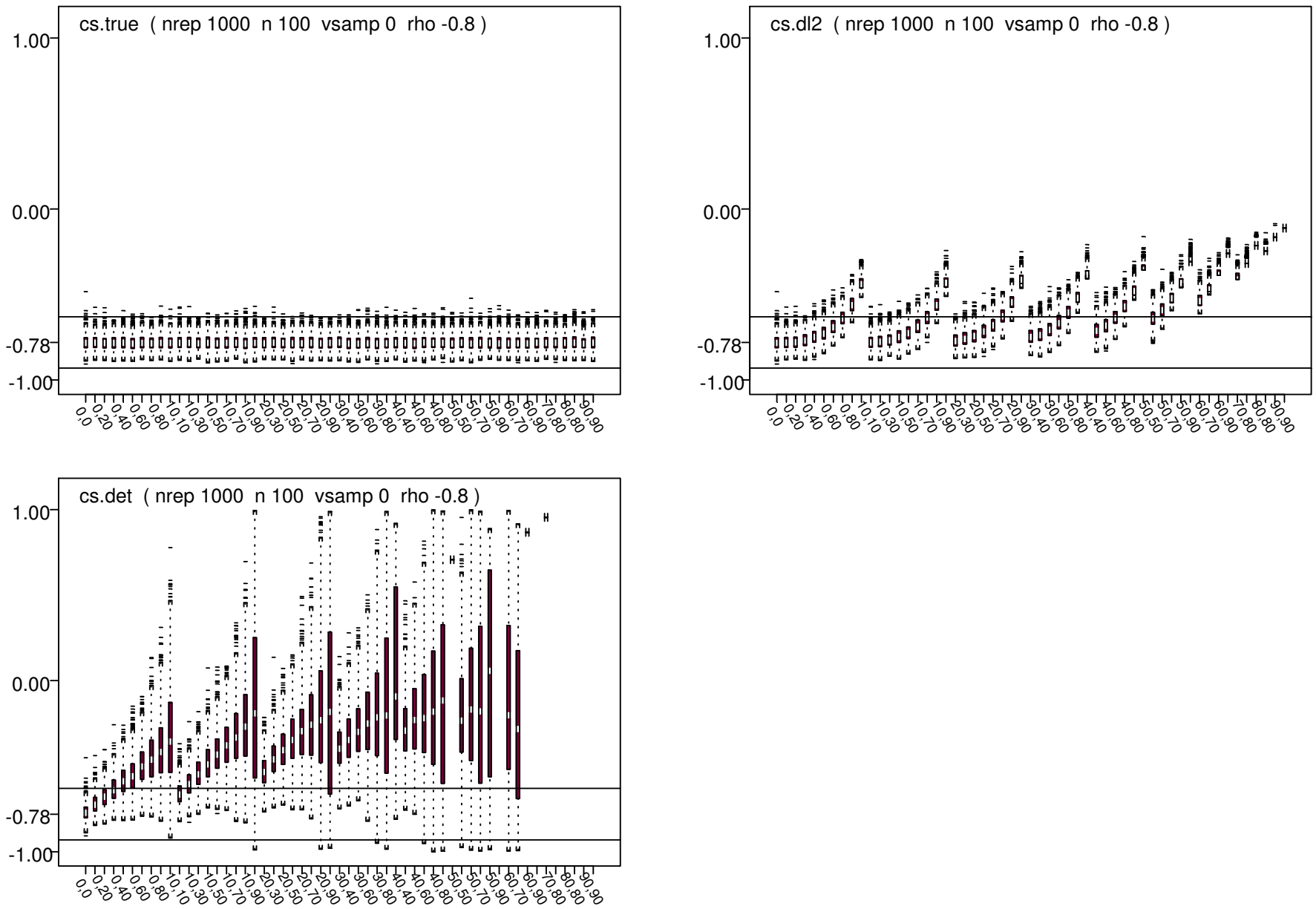

Figure S2. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

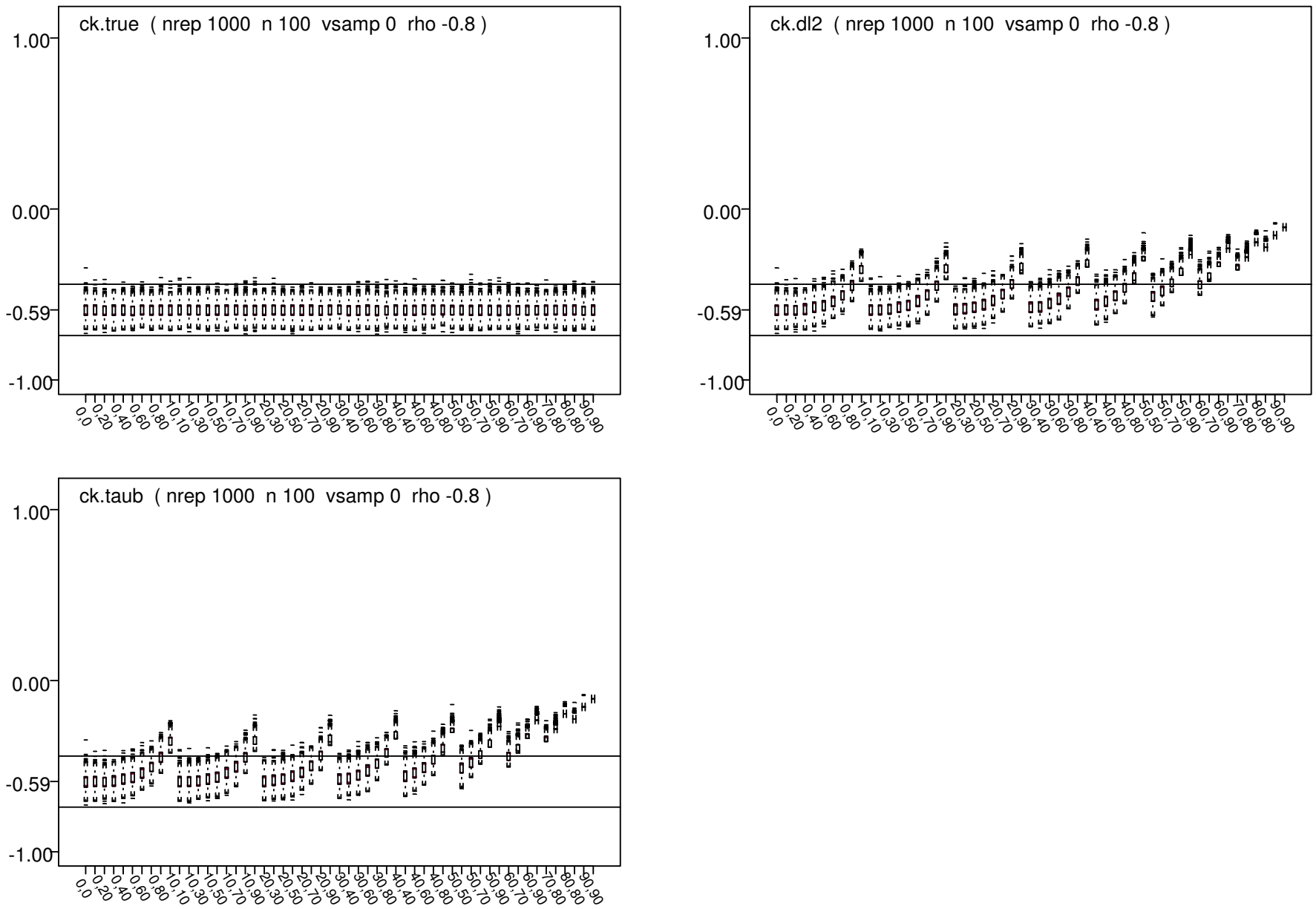

Figure S3. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

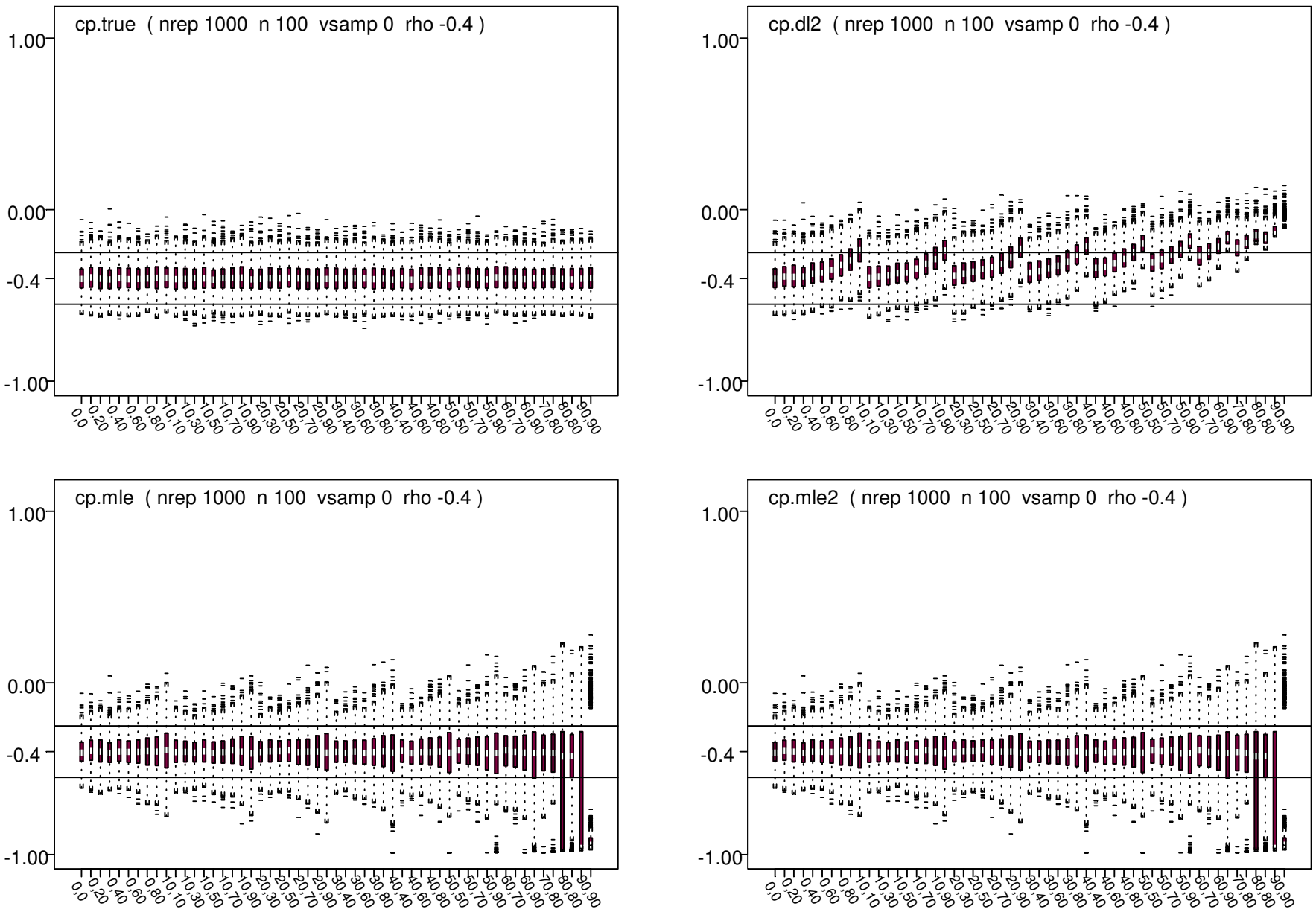

Figure S4. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

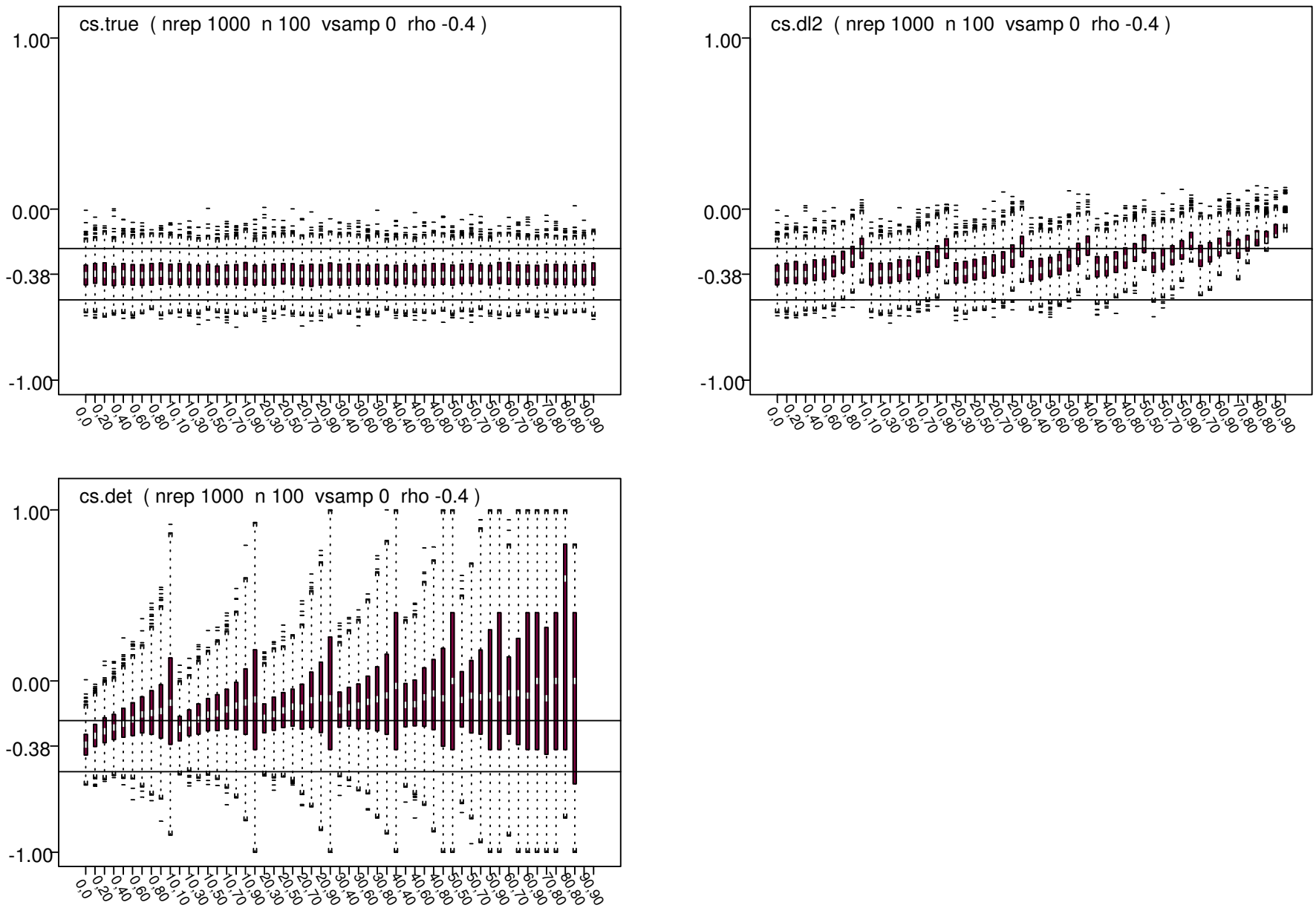

Figure S5. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

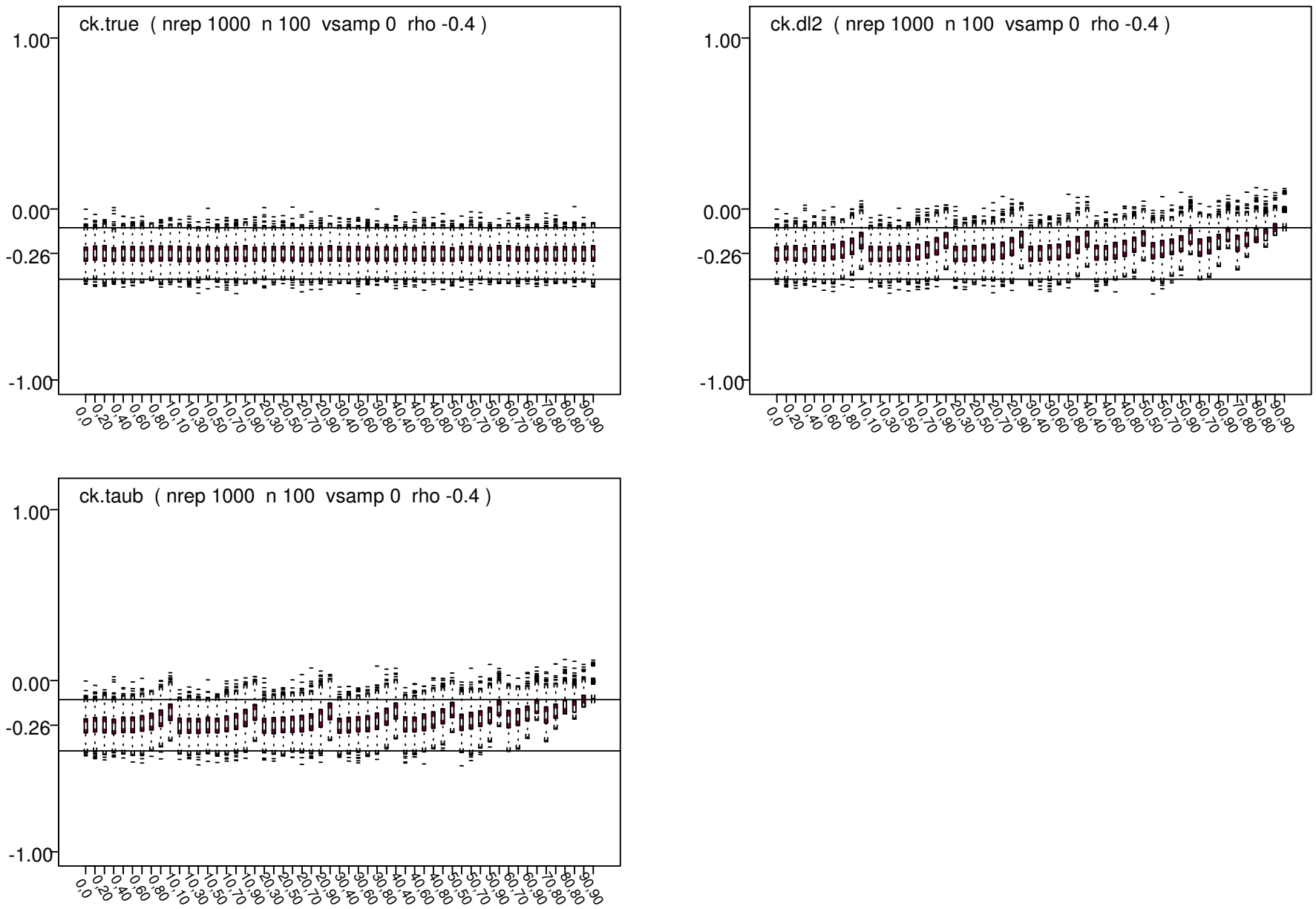

Figure S6. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

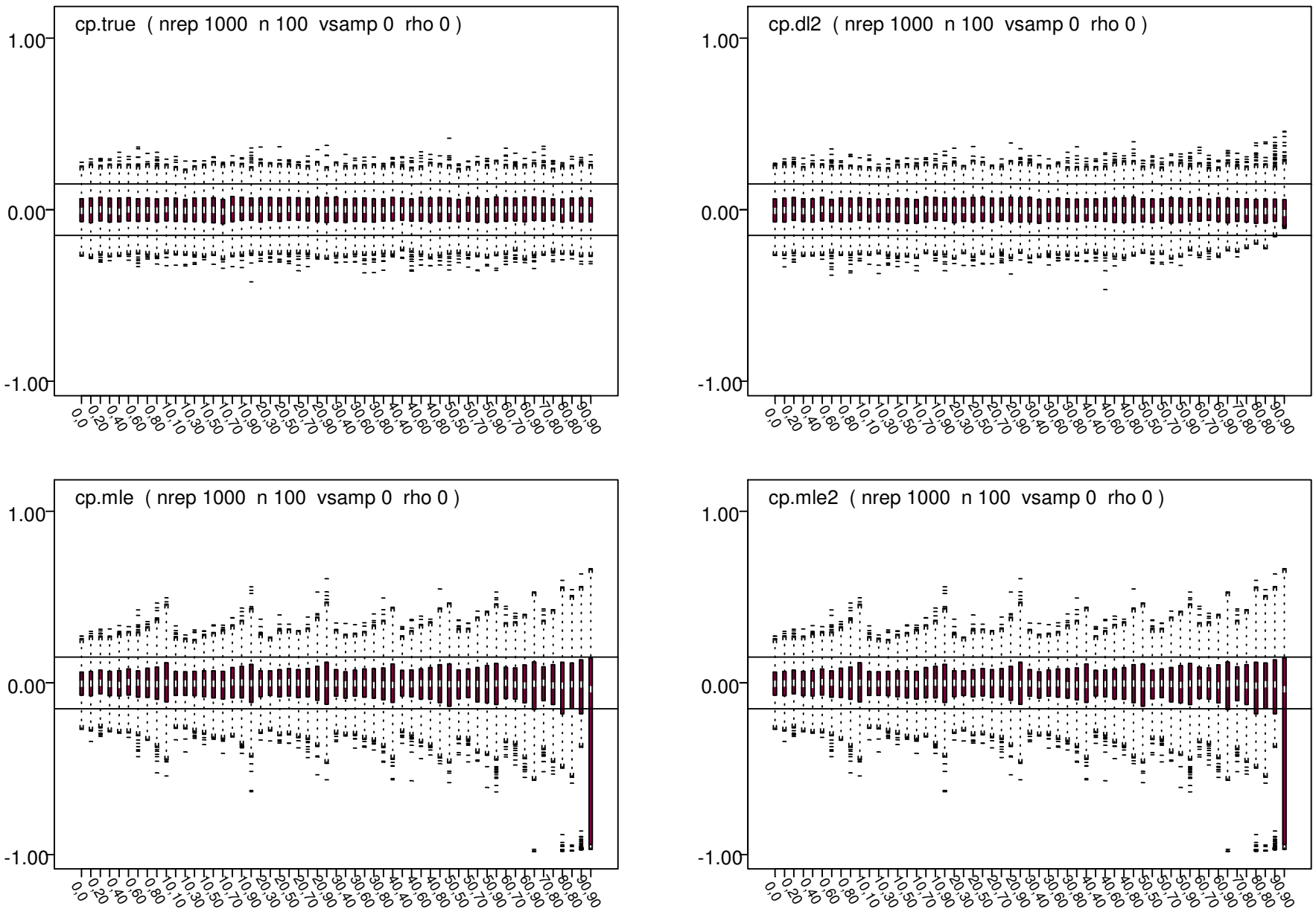

Figure S7. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

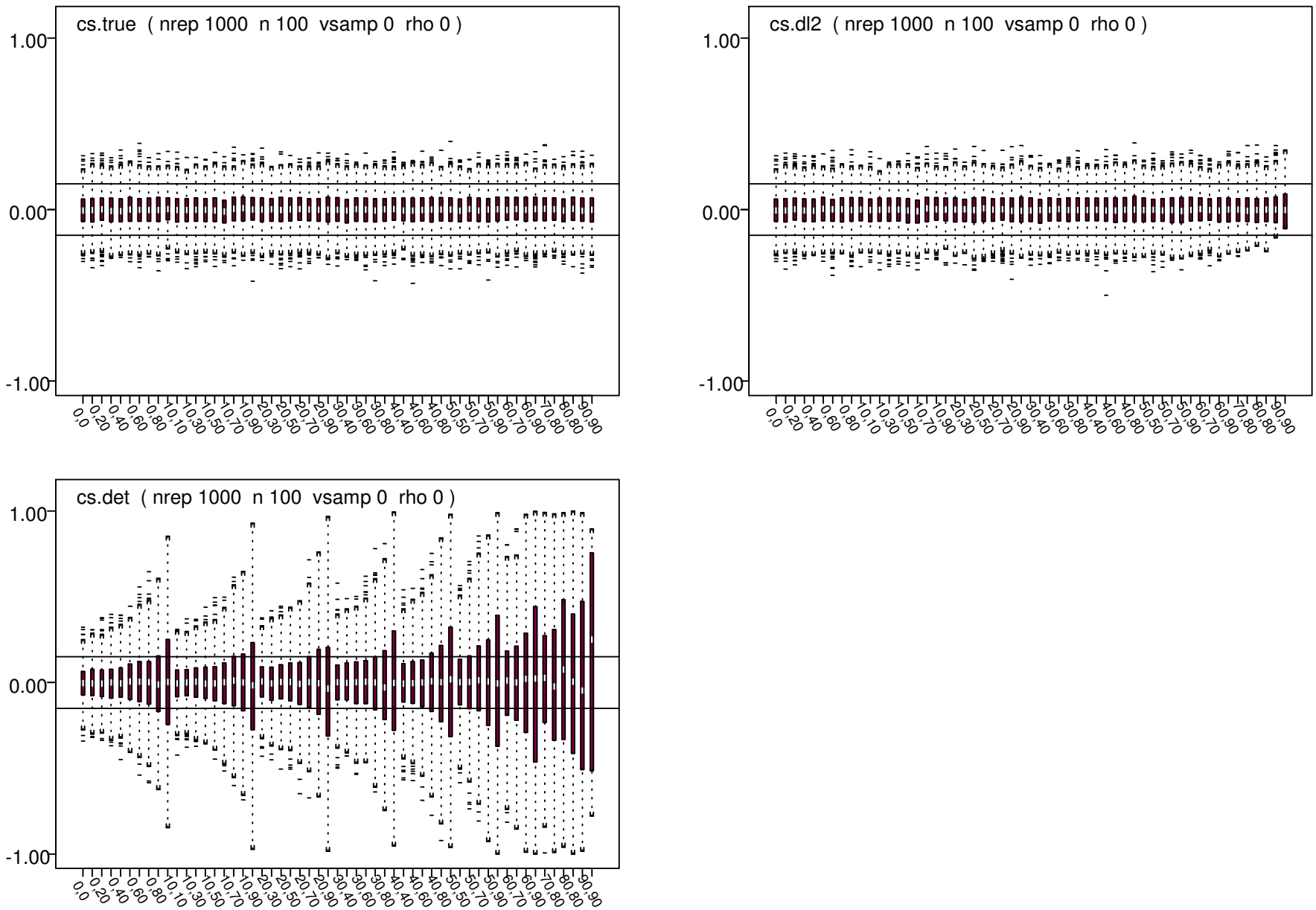

Figure S8. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

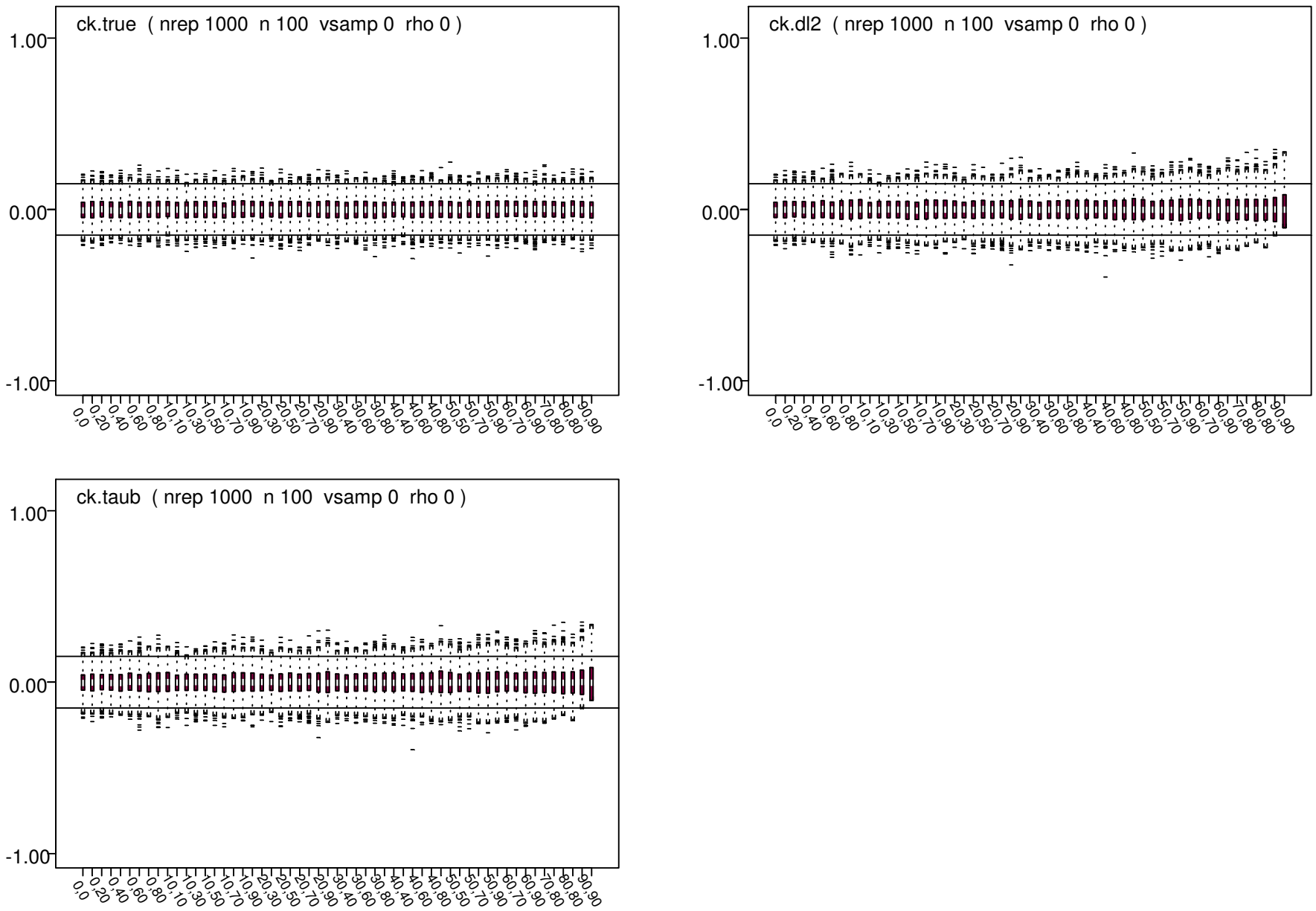

Figure S9. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

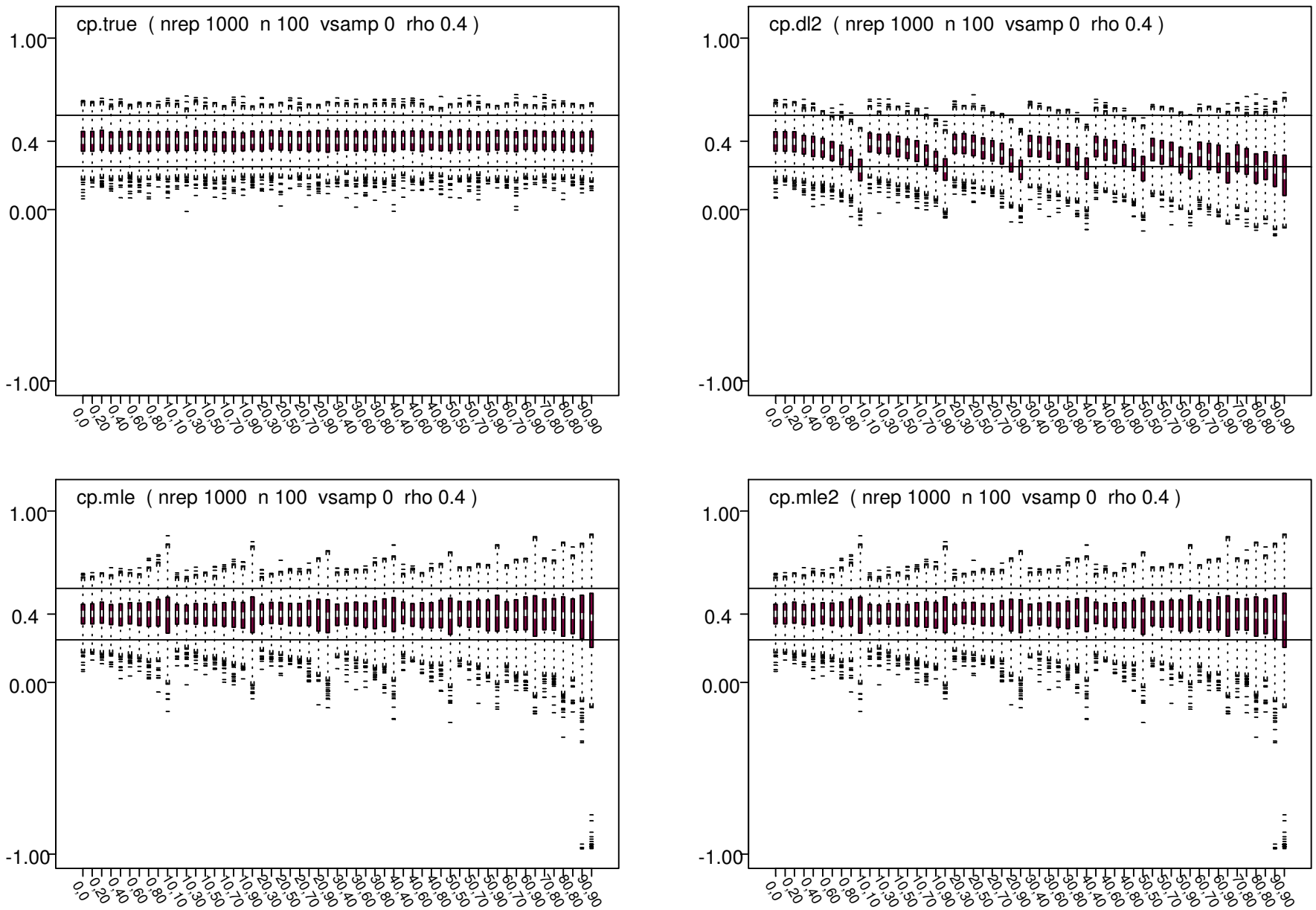

Figure S10. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

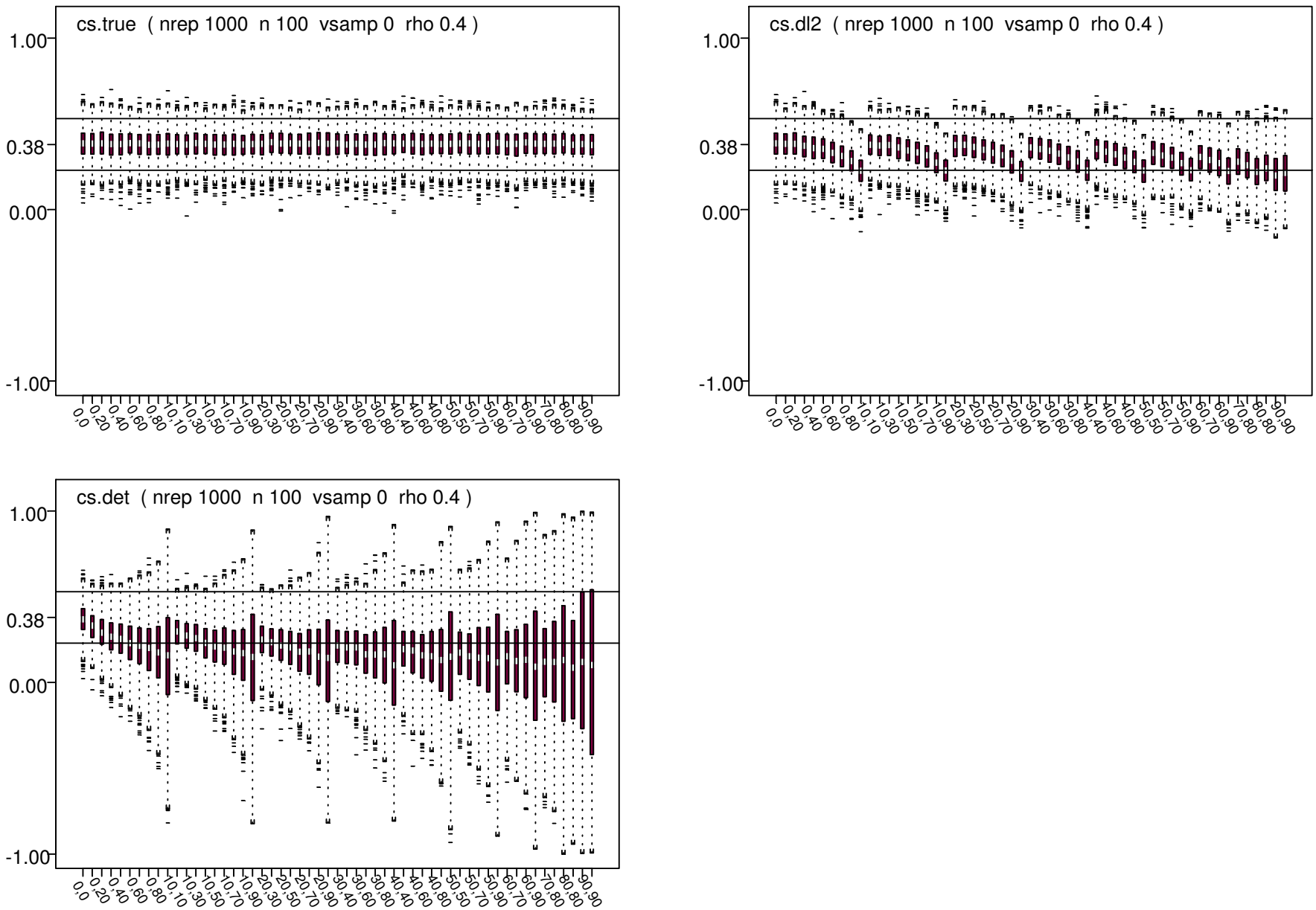

Figure S11. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

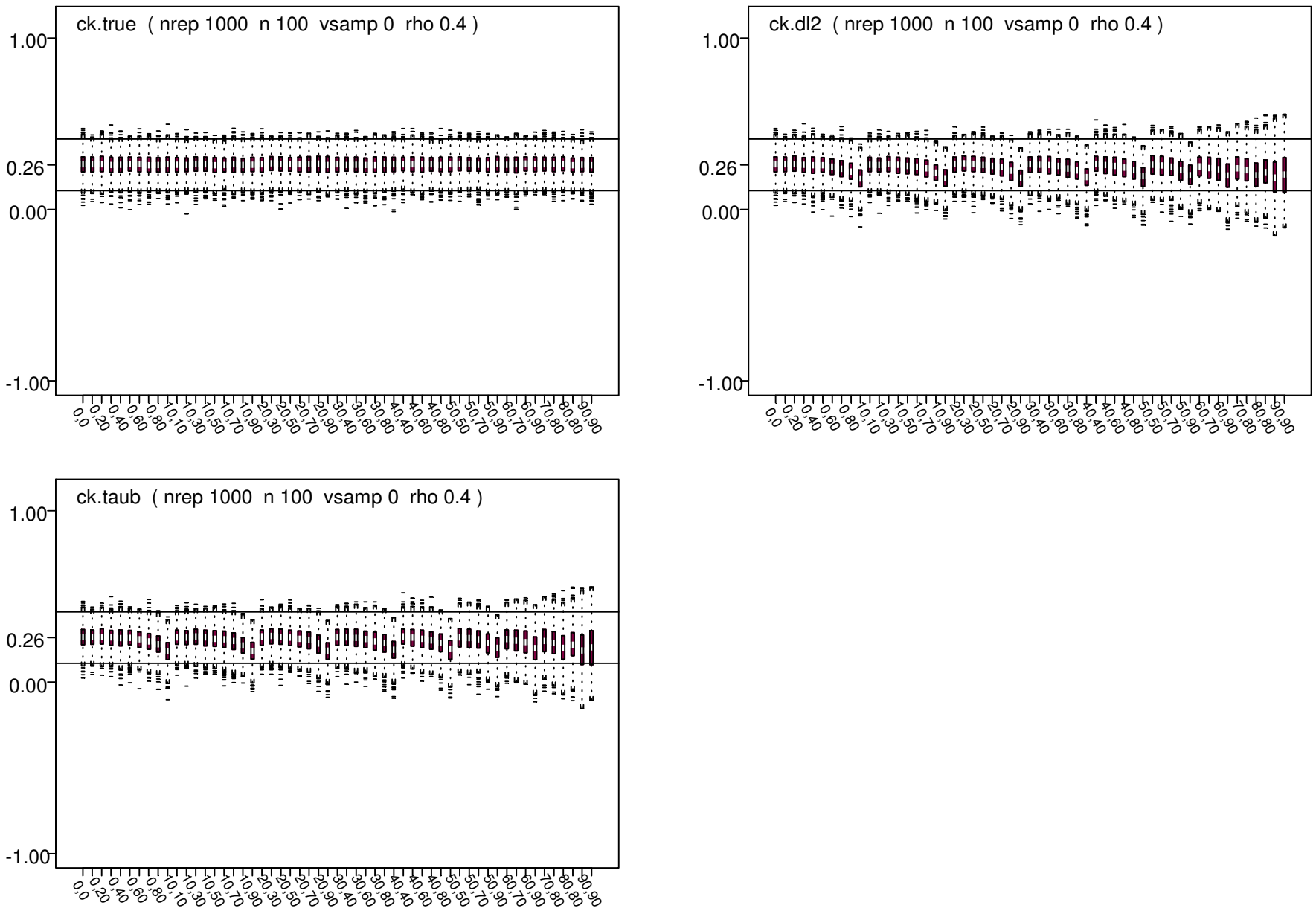

Figure S12. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

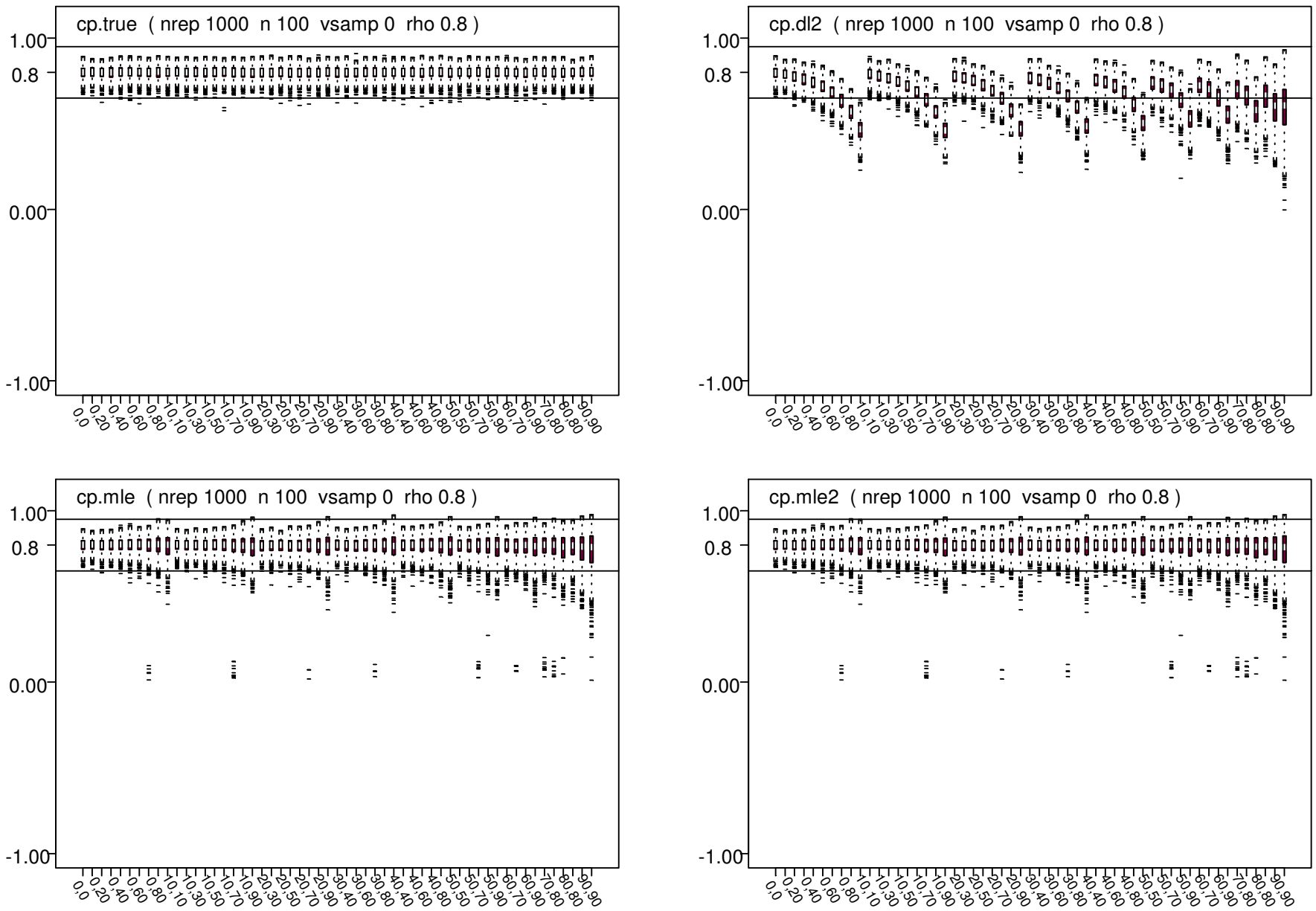

Figure S13. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

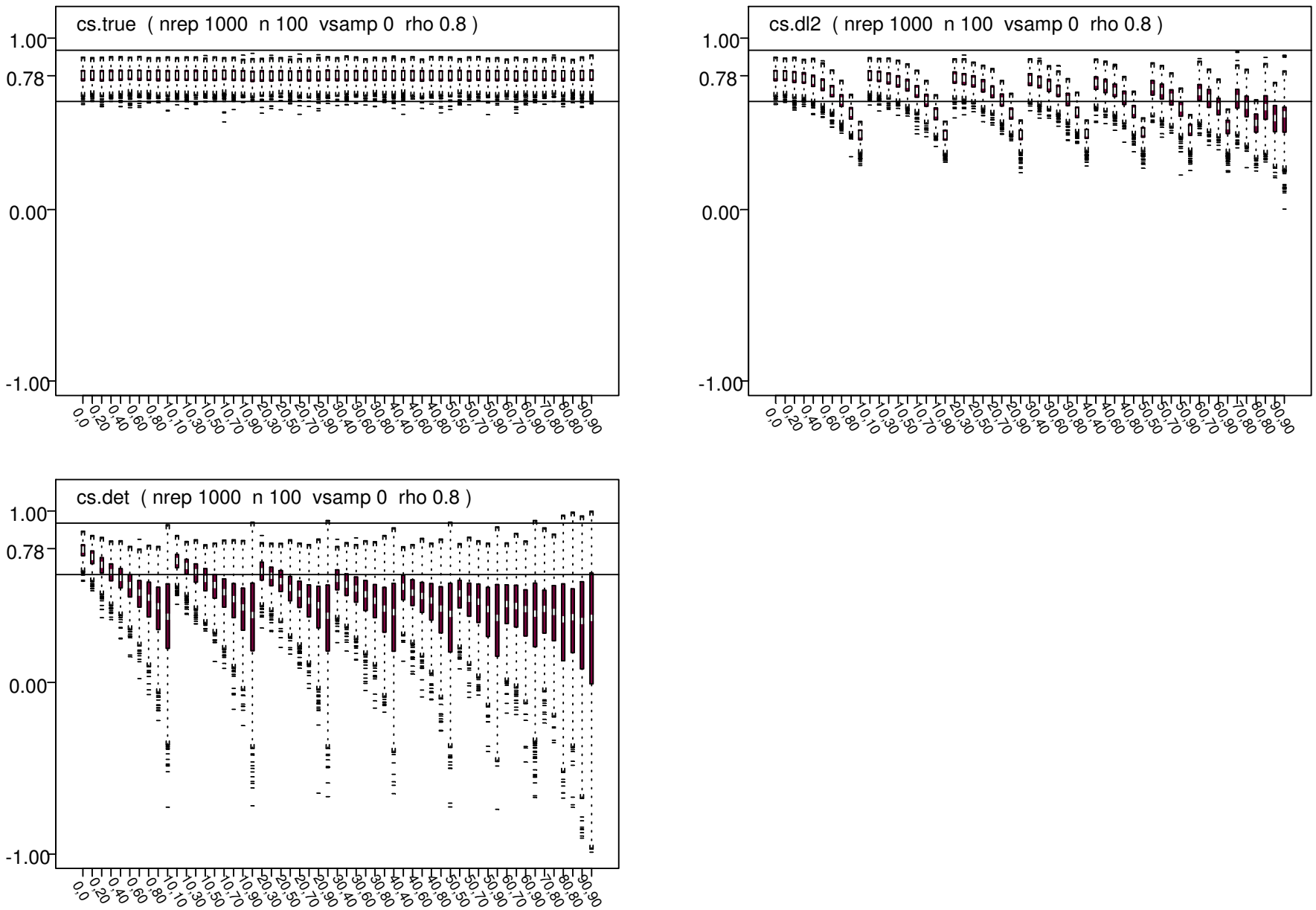

Figure S14. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

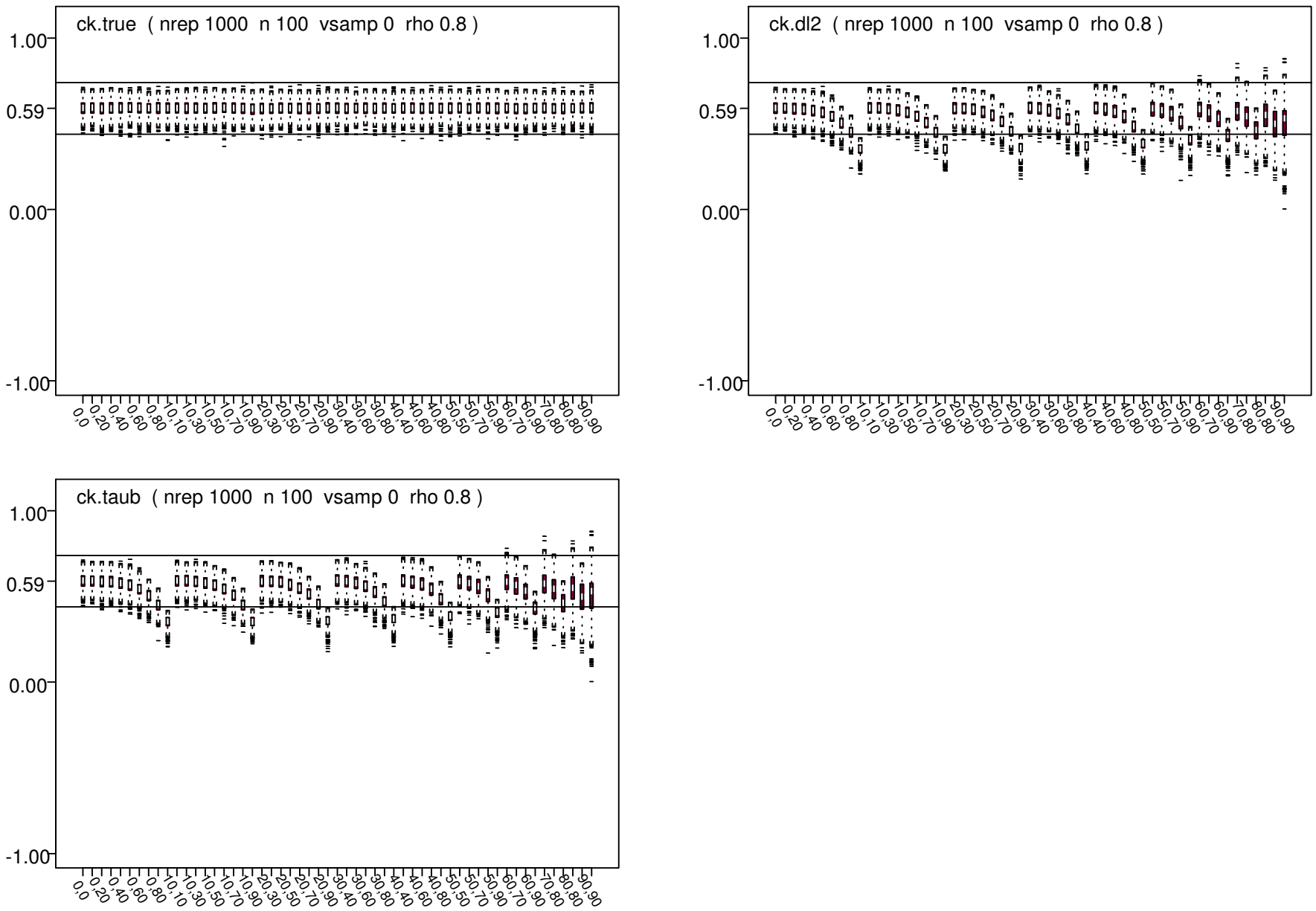

Figure S15. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

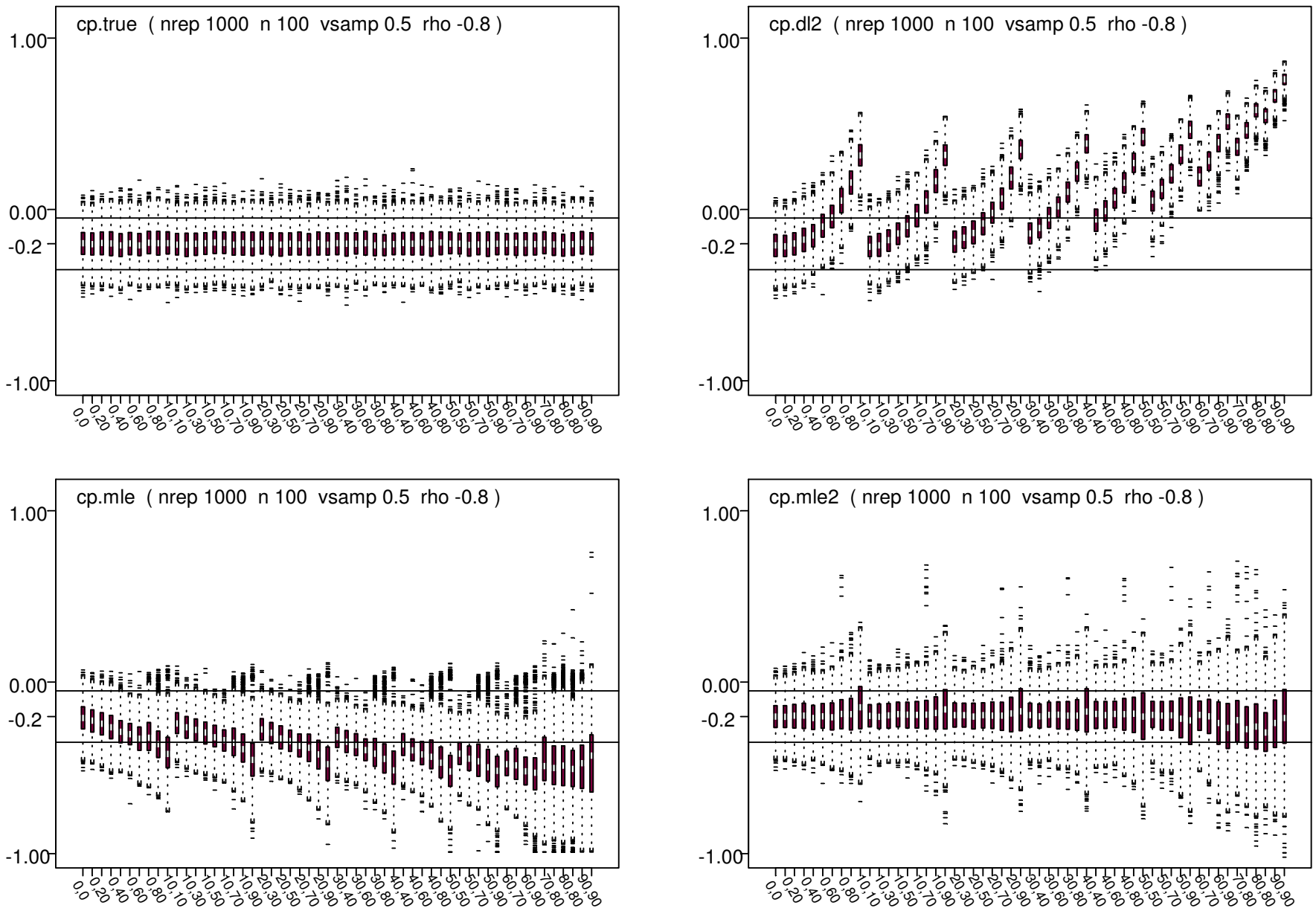

Figure S16. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

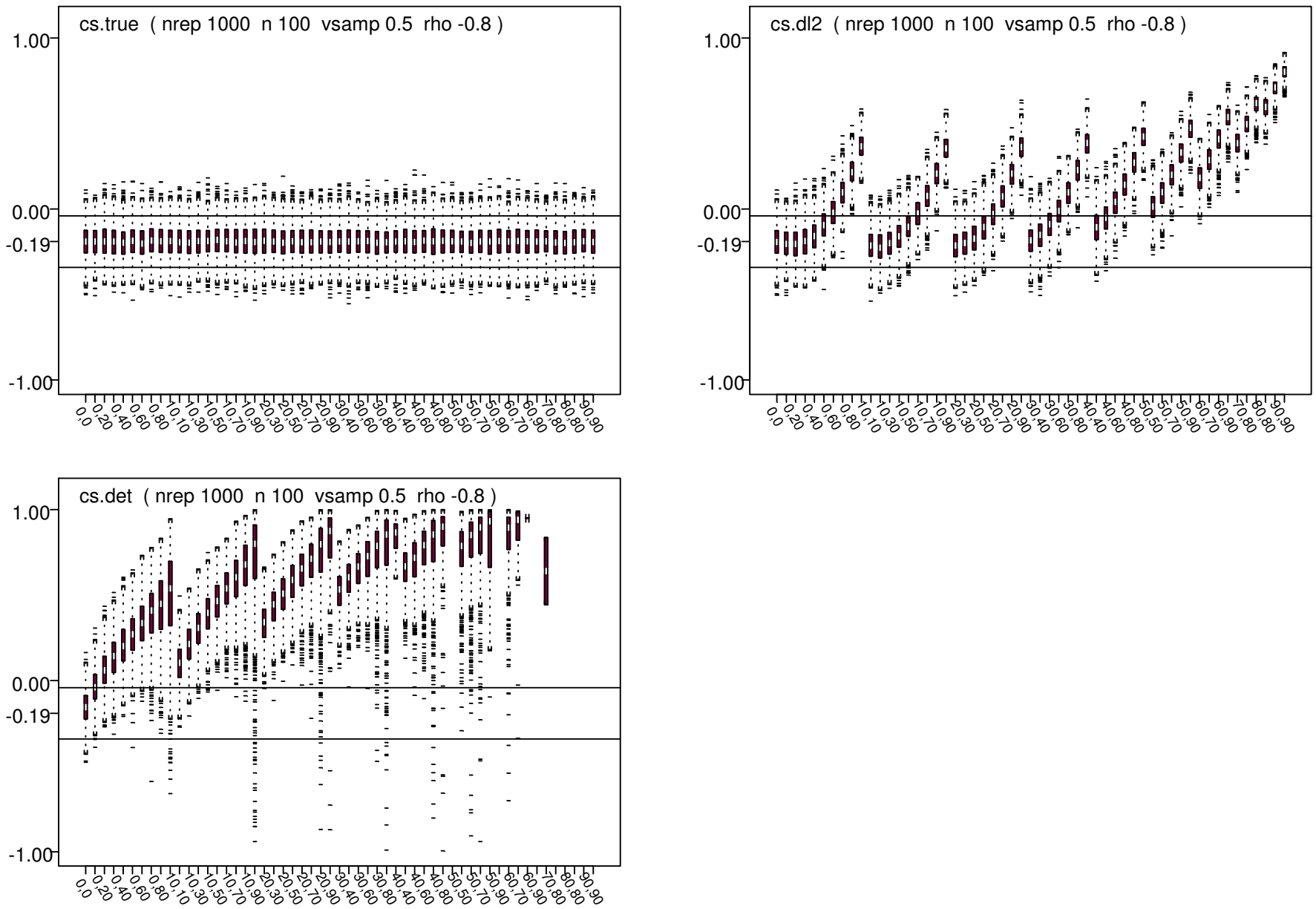

Figure S17. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

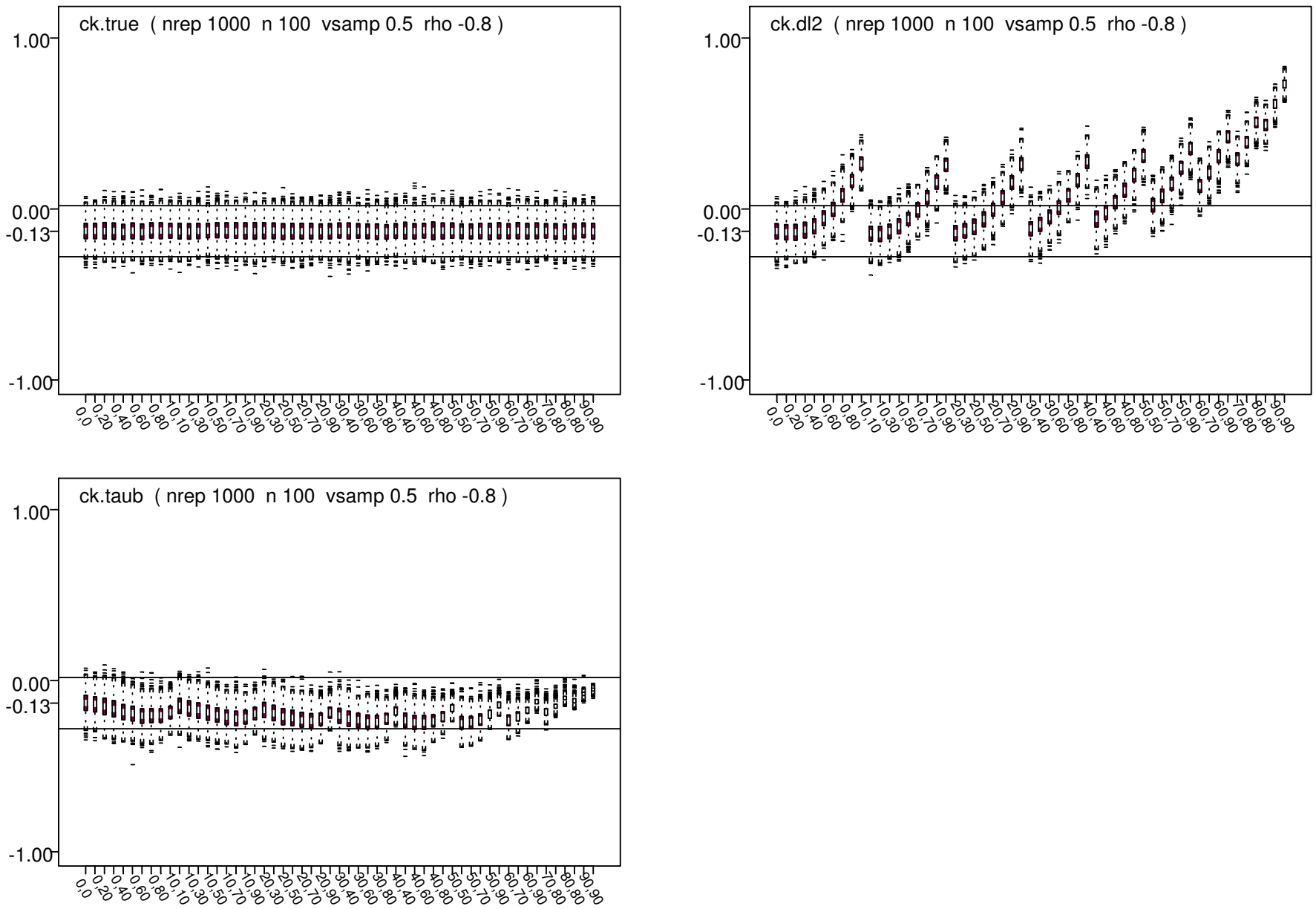

Figure S18. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

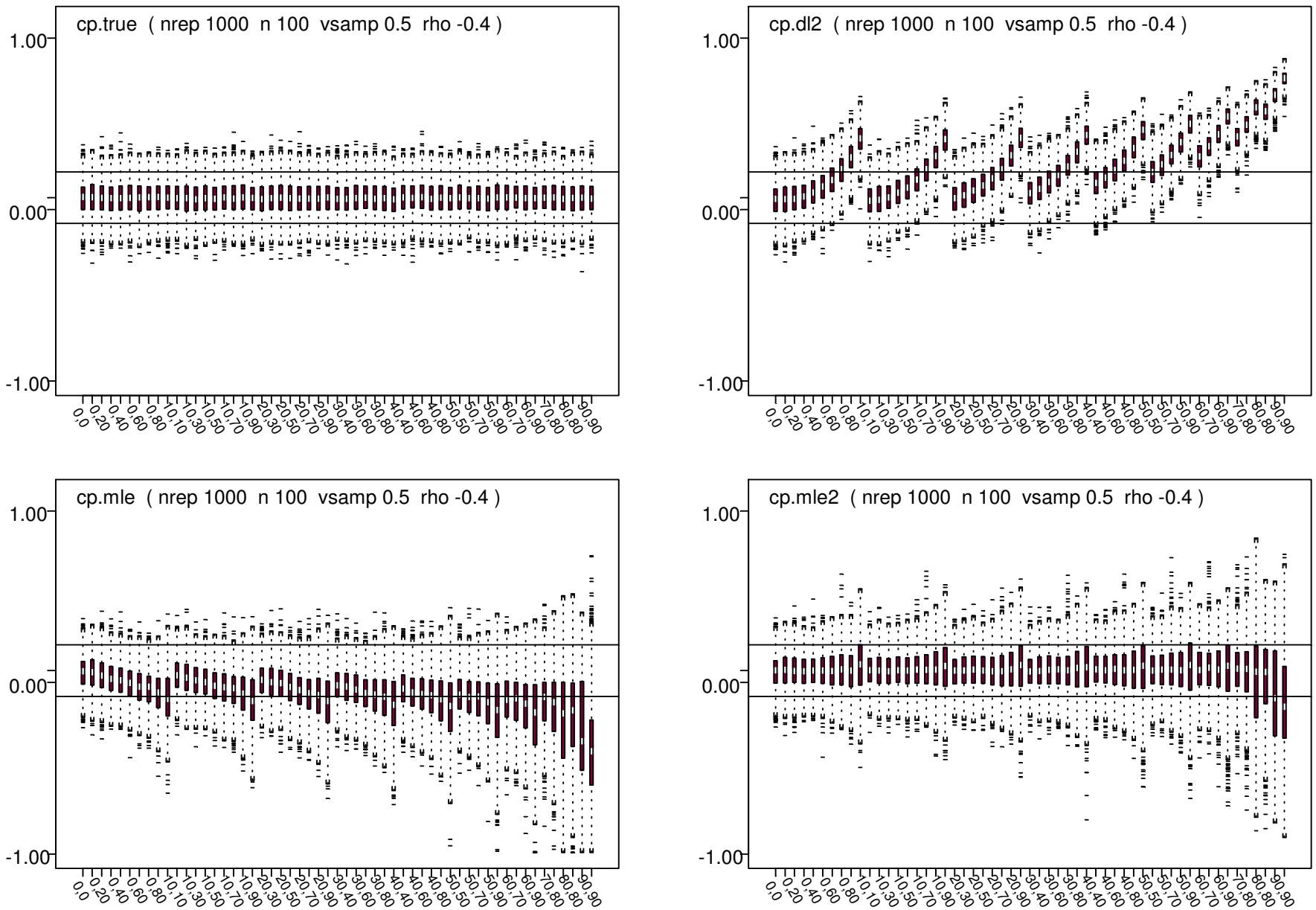

Figure S19. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

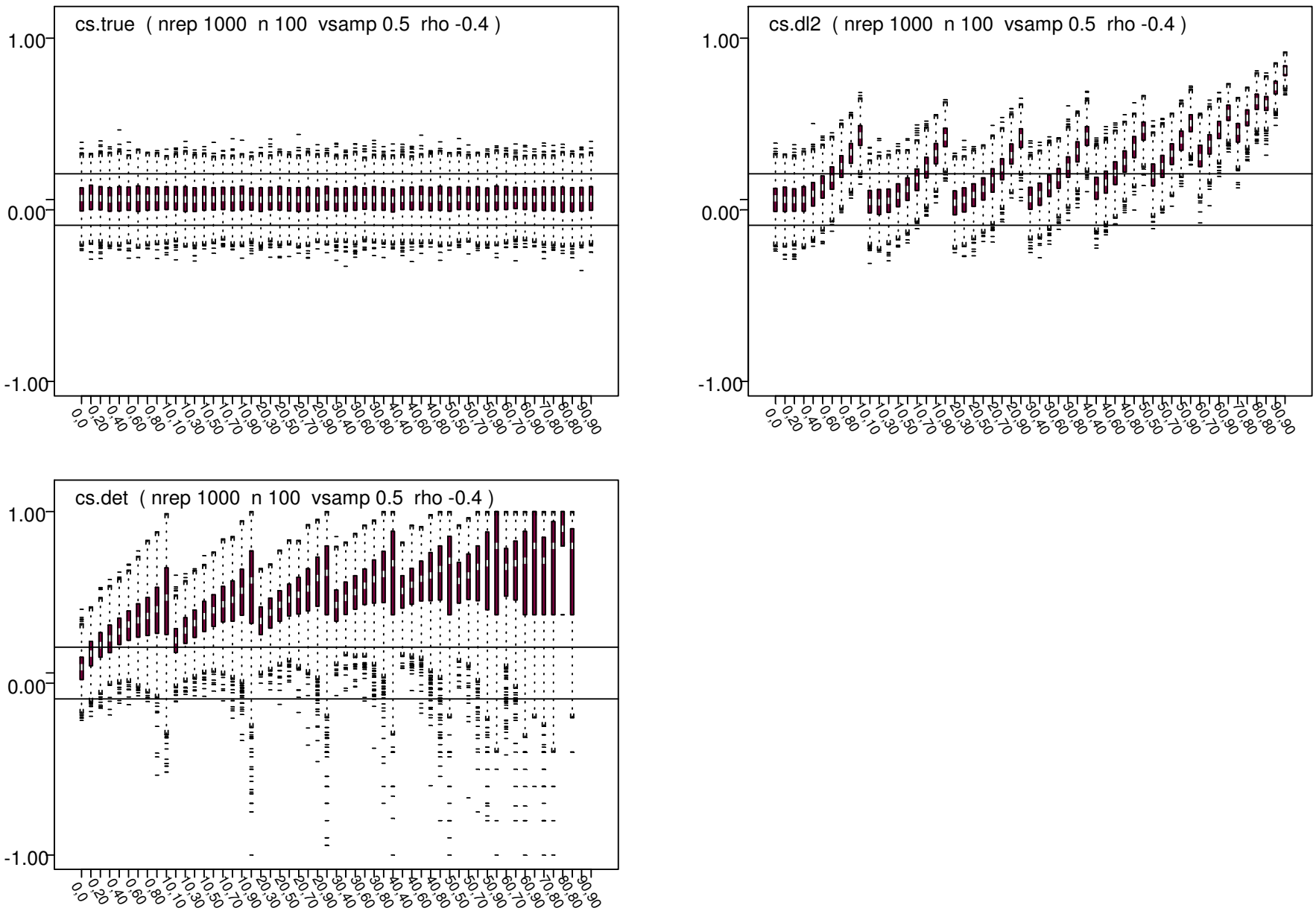

Figure S20. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

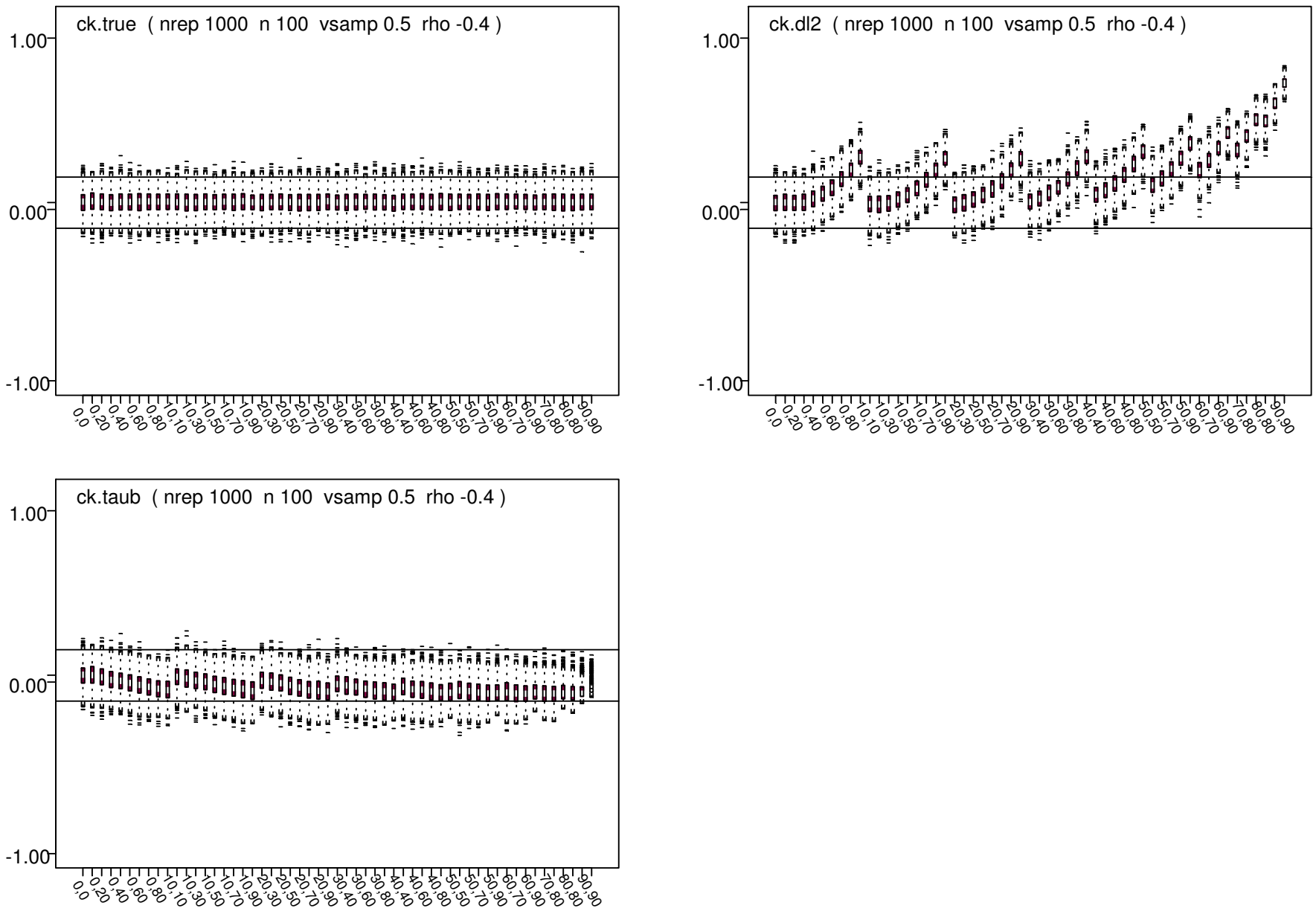

Figure S21. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

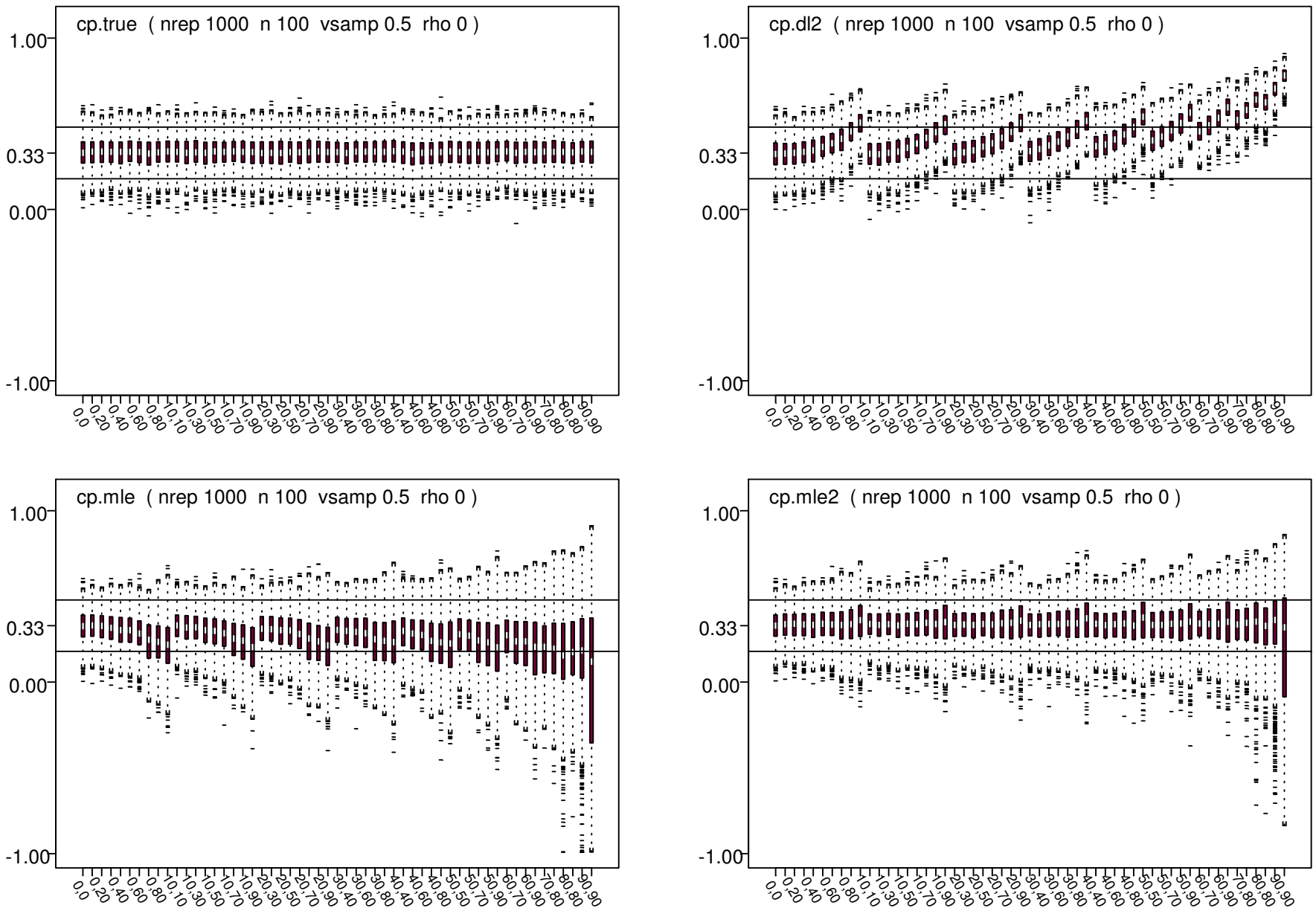

Figure S22. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

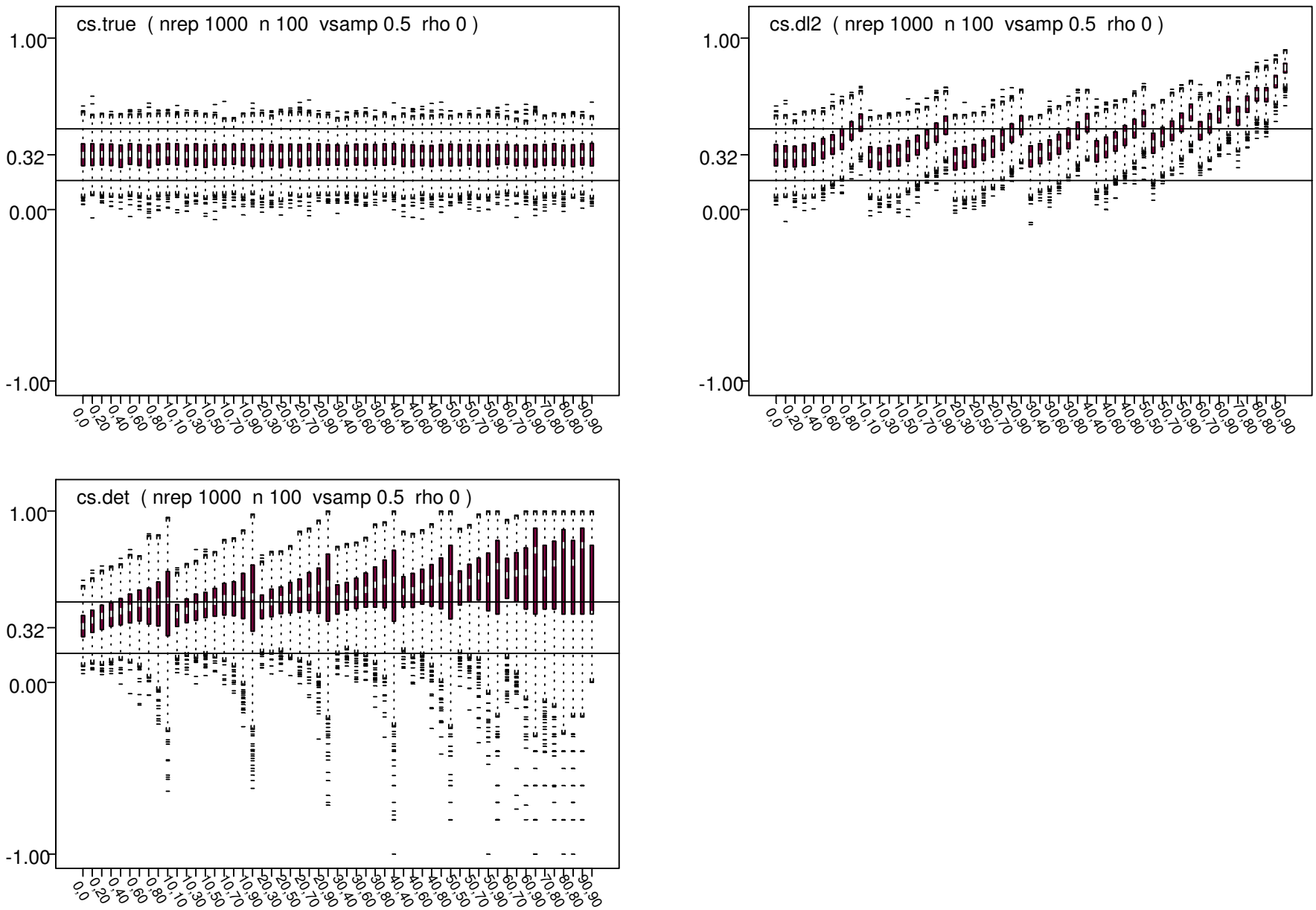

Figure S23. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

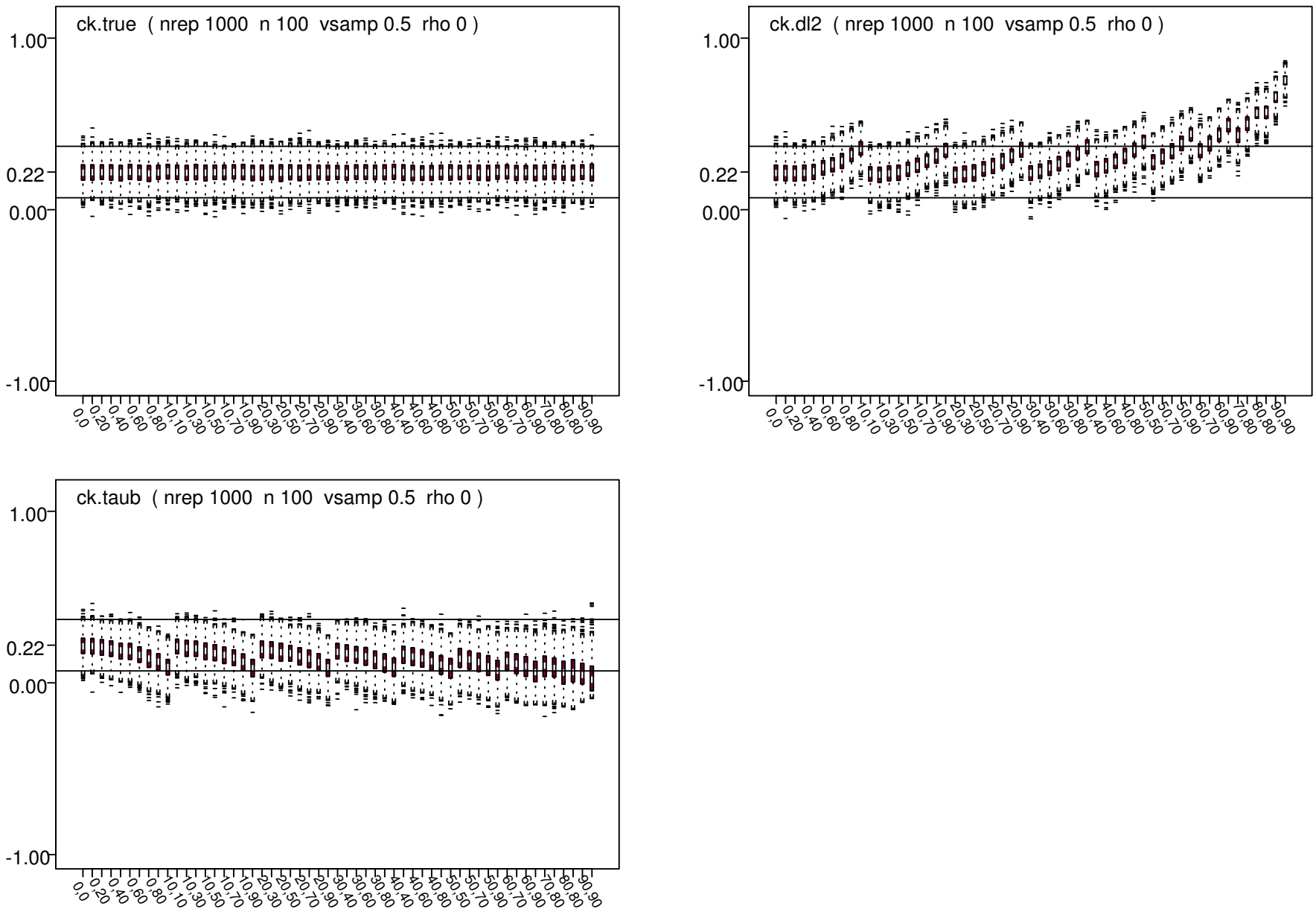

Figure S24. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

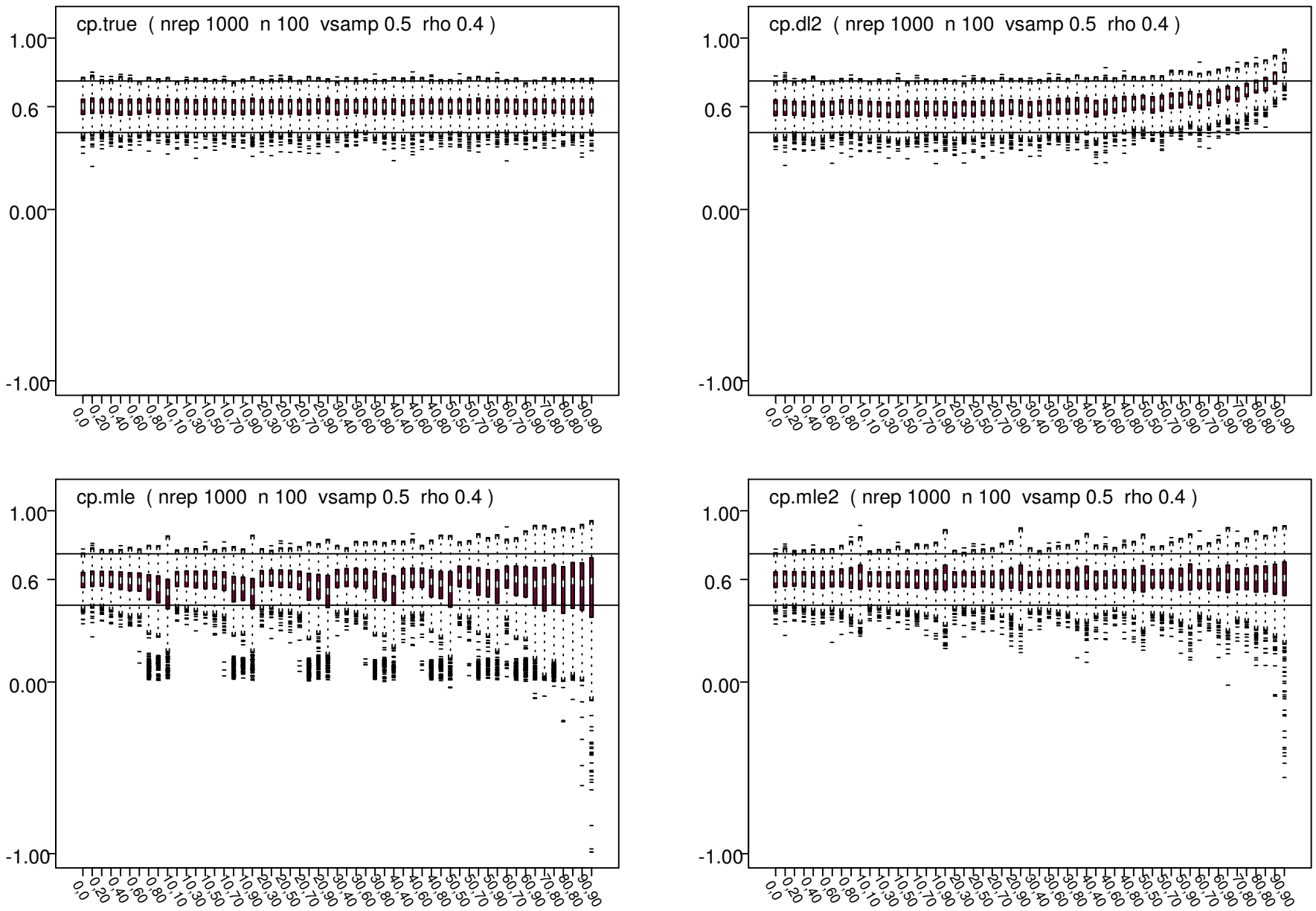

Figure S25. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

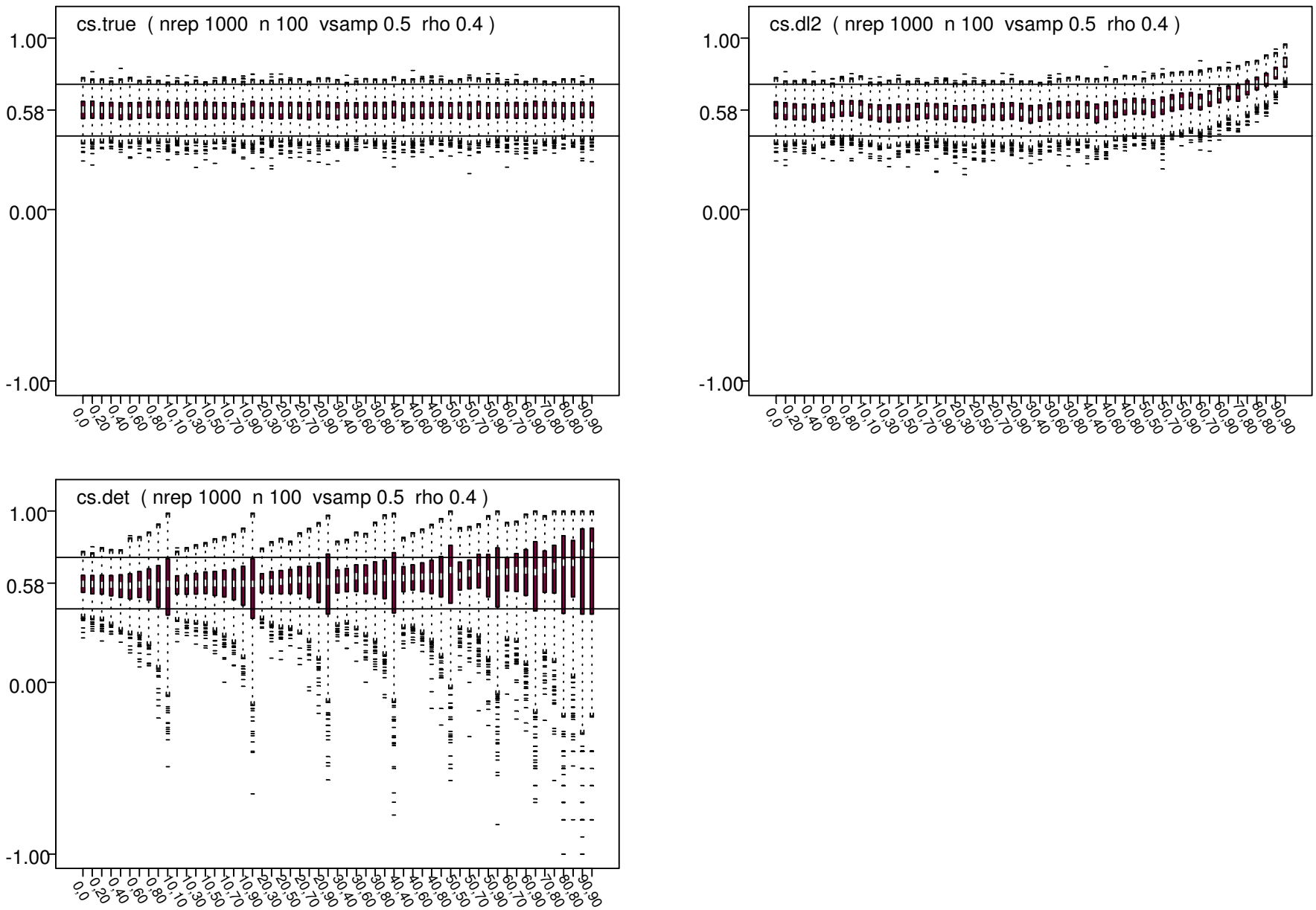

Figure S26. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

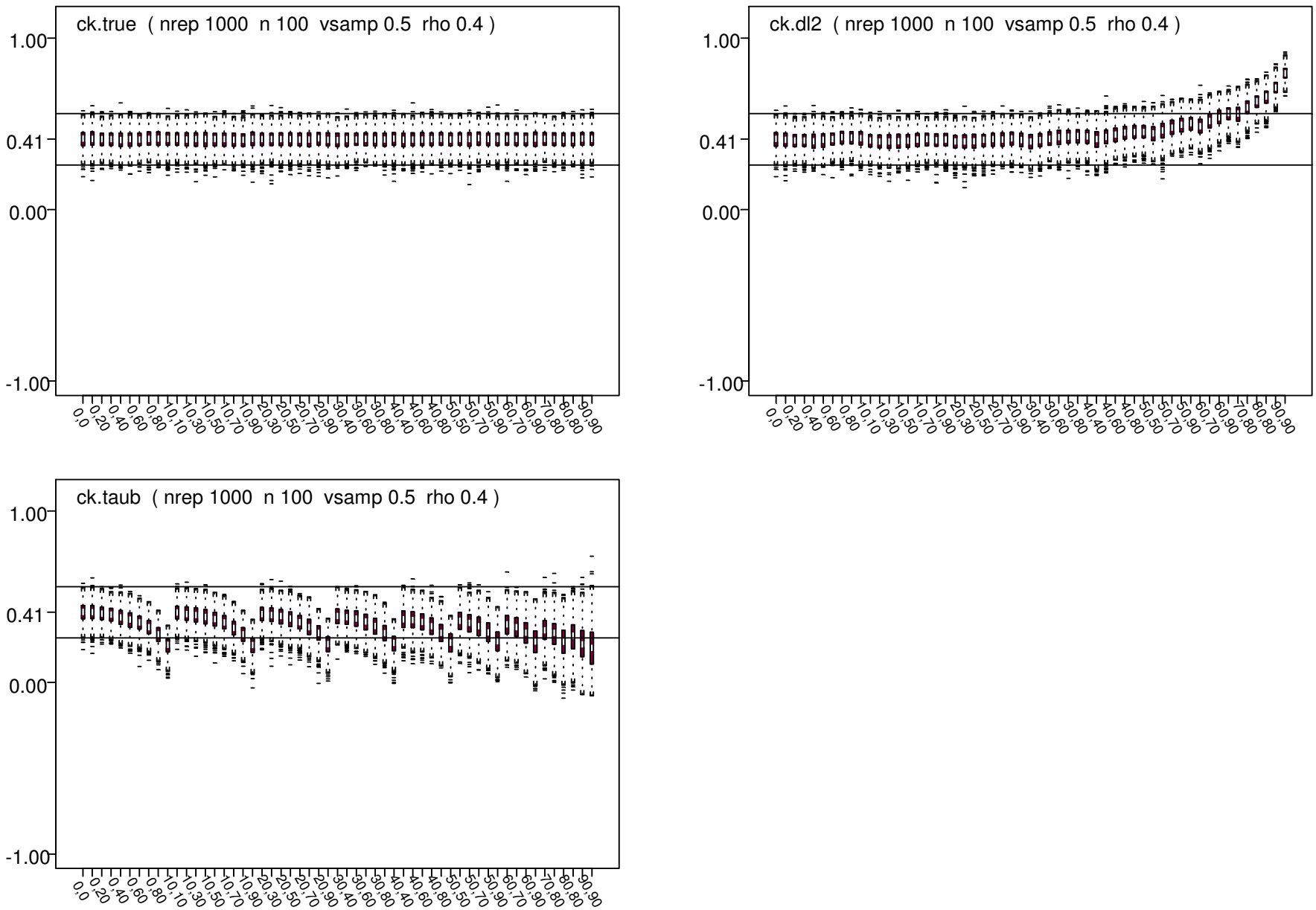

Figure S27. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

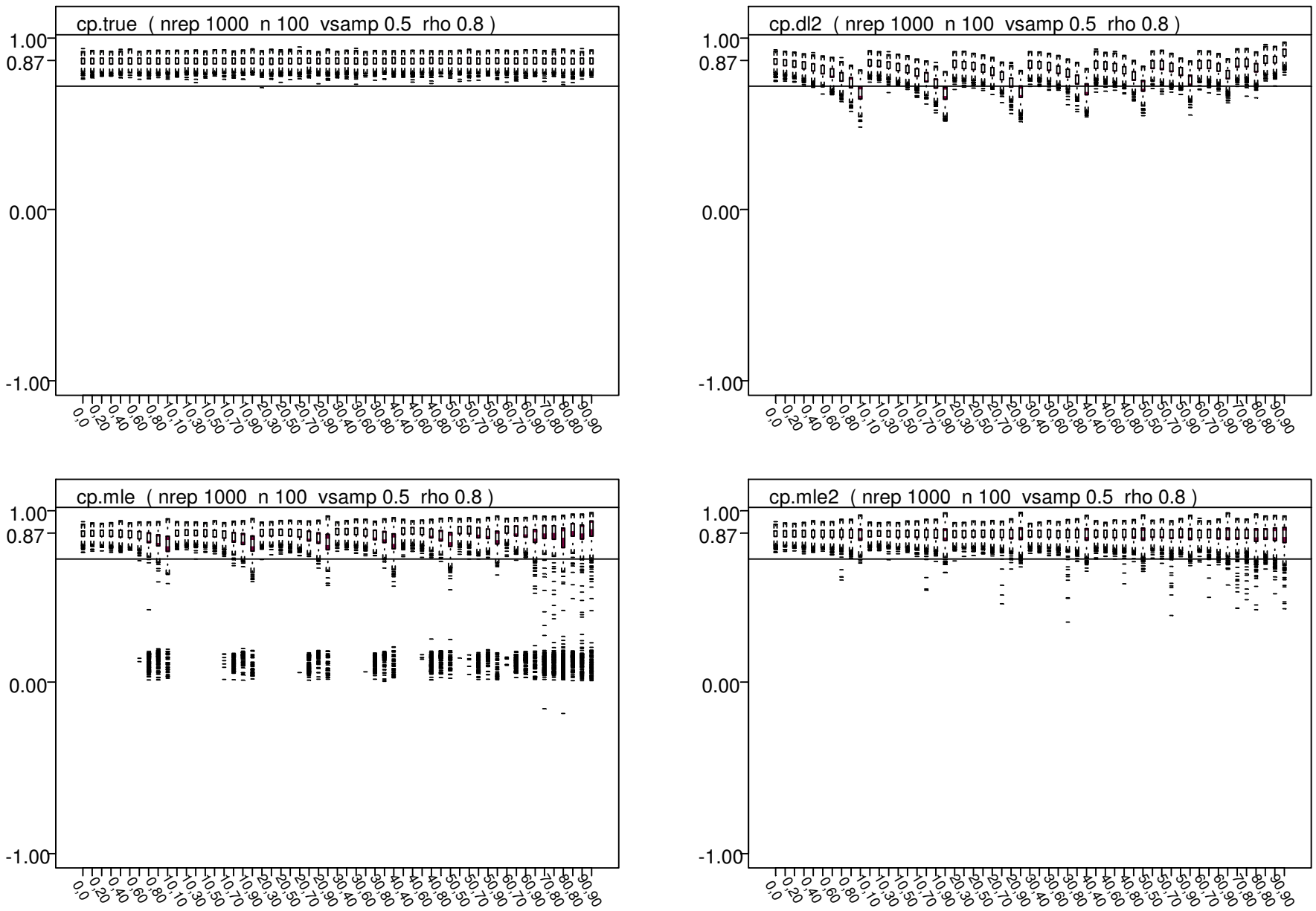

Figure S28. Boxplots of estimates of Pearson correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

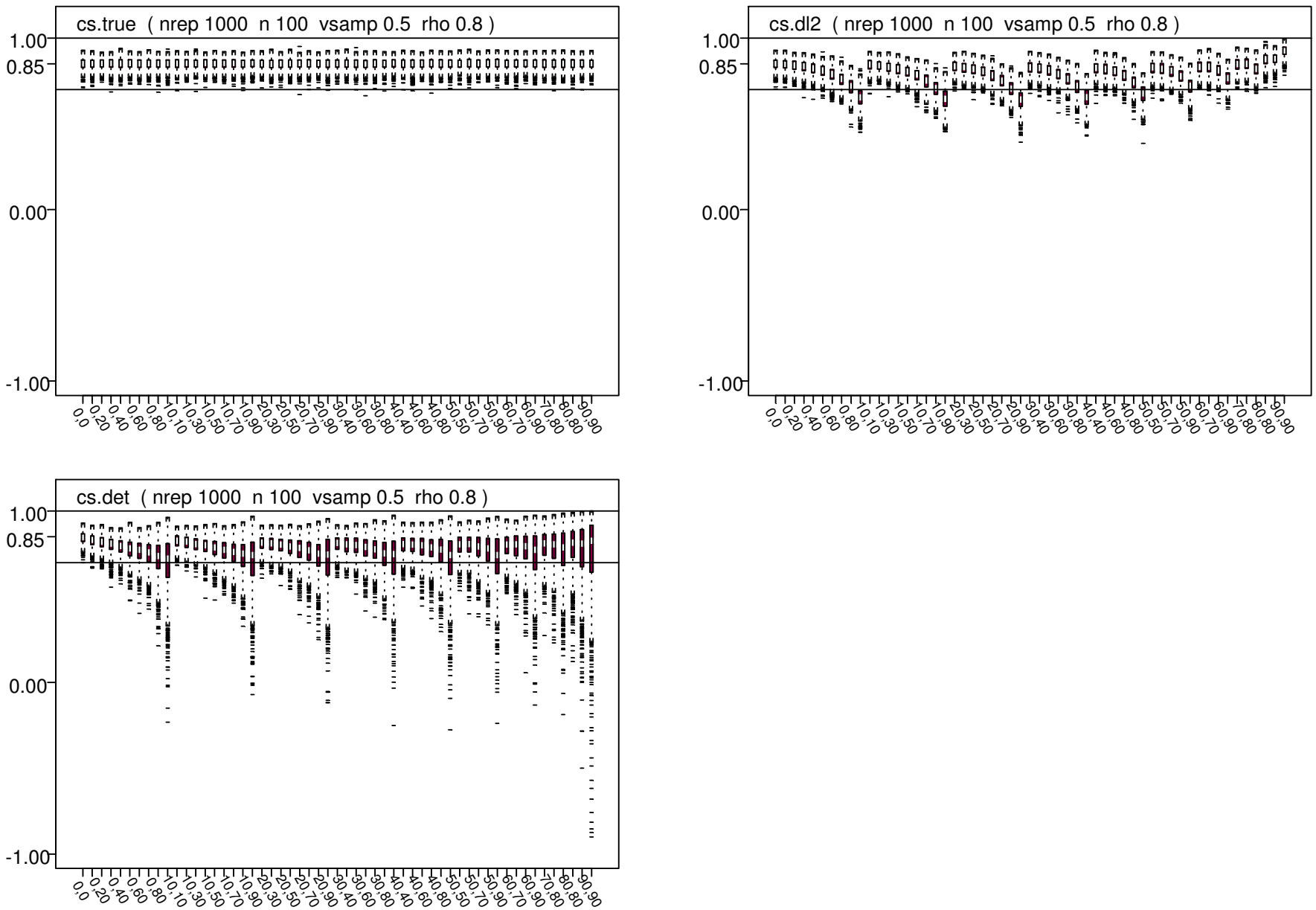

Figure S29. Boxplots of estimates of Spearman correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

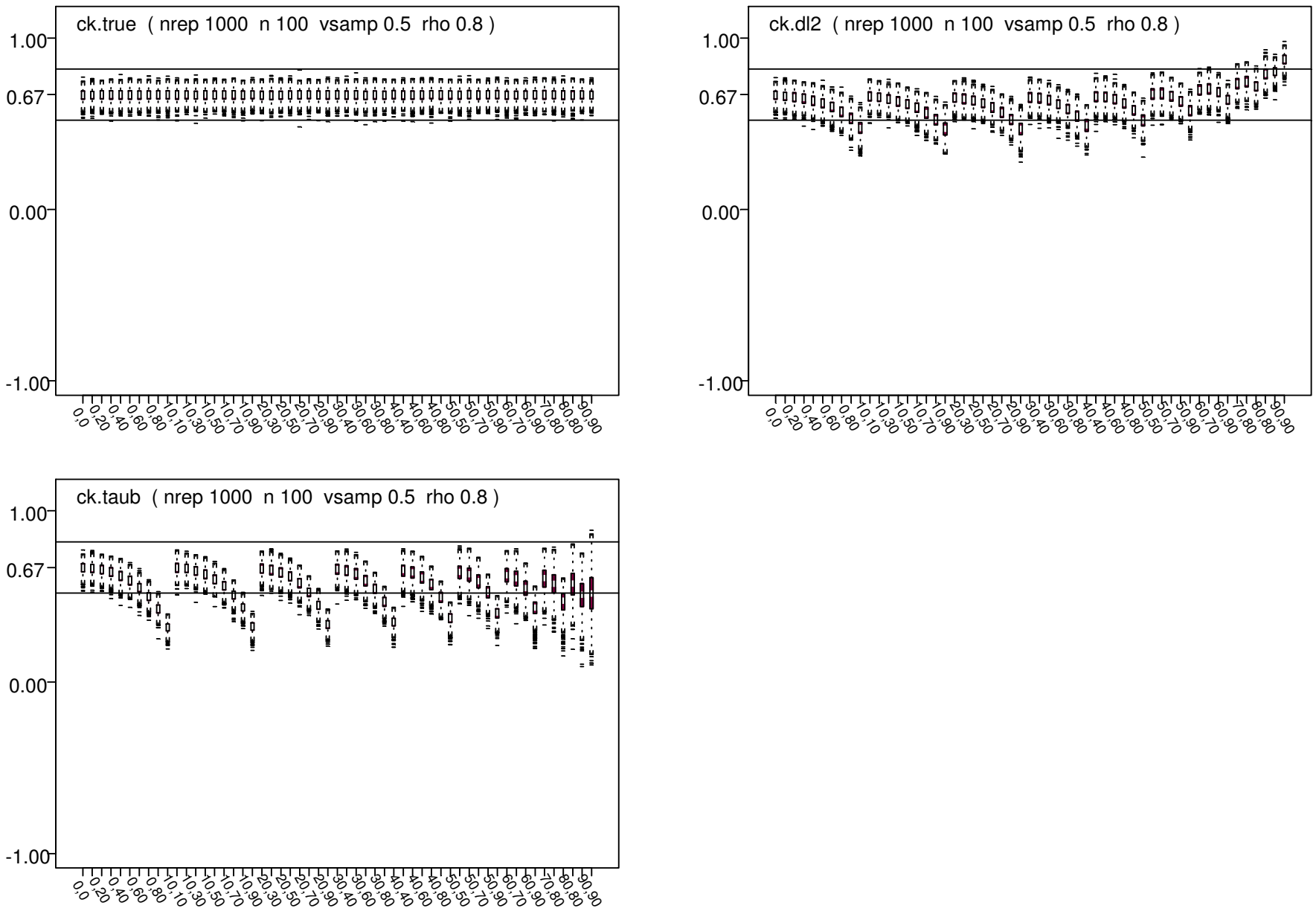

Figure S30. Boxplots of estimates of Kendall correlation. Horizontal axis is percent censored in x and y. Estimators and parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

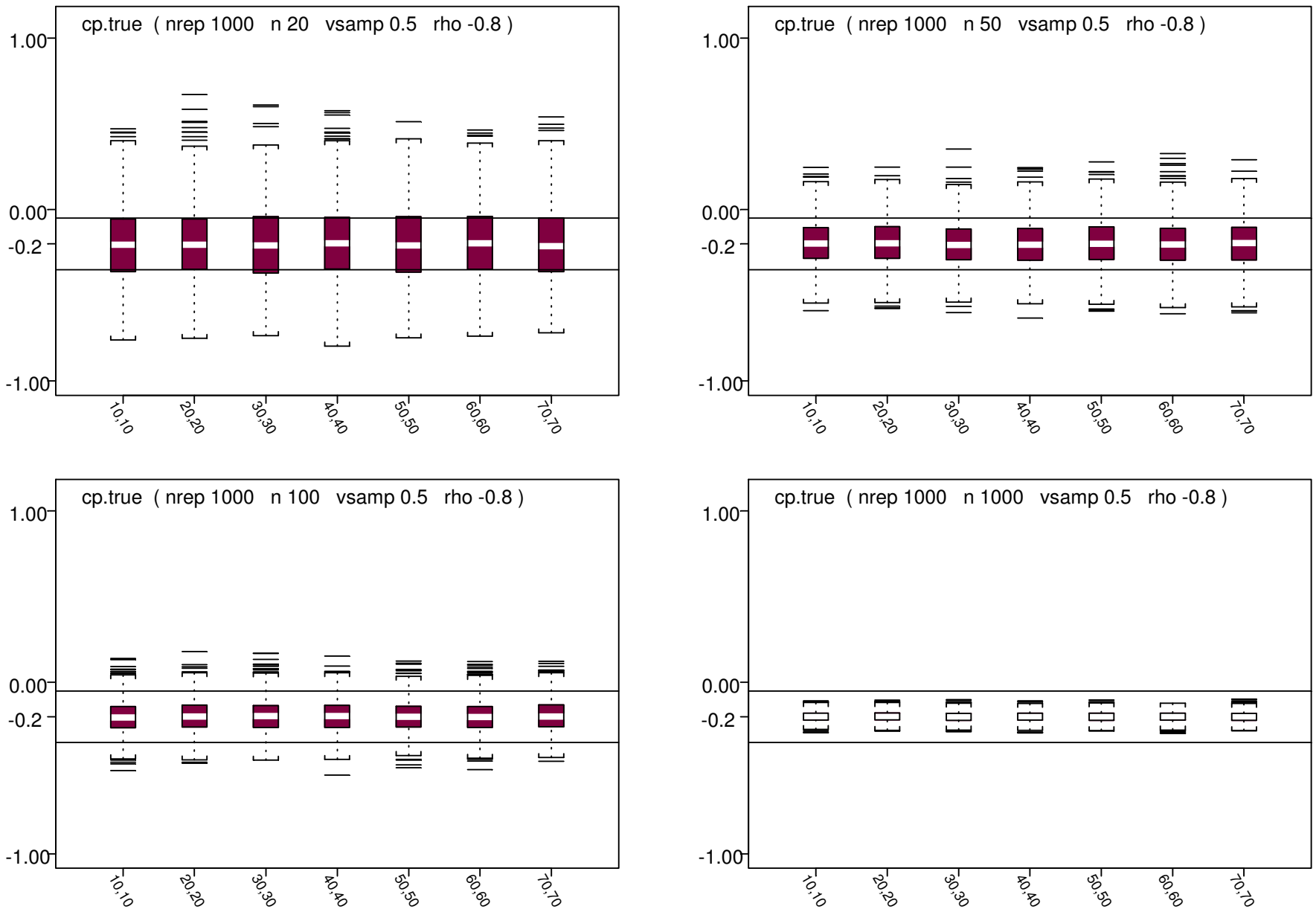

Figure S31. Box plots of estimates of correlation. Estimator = cp true. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

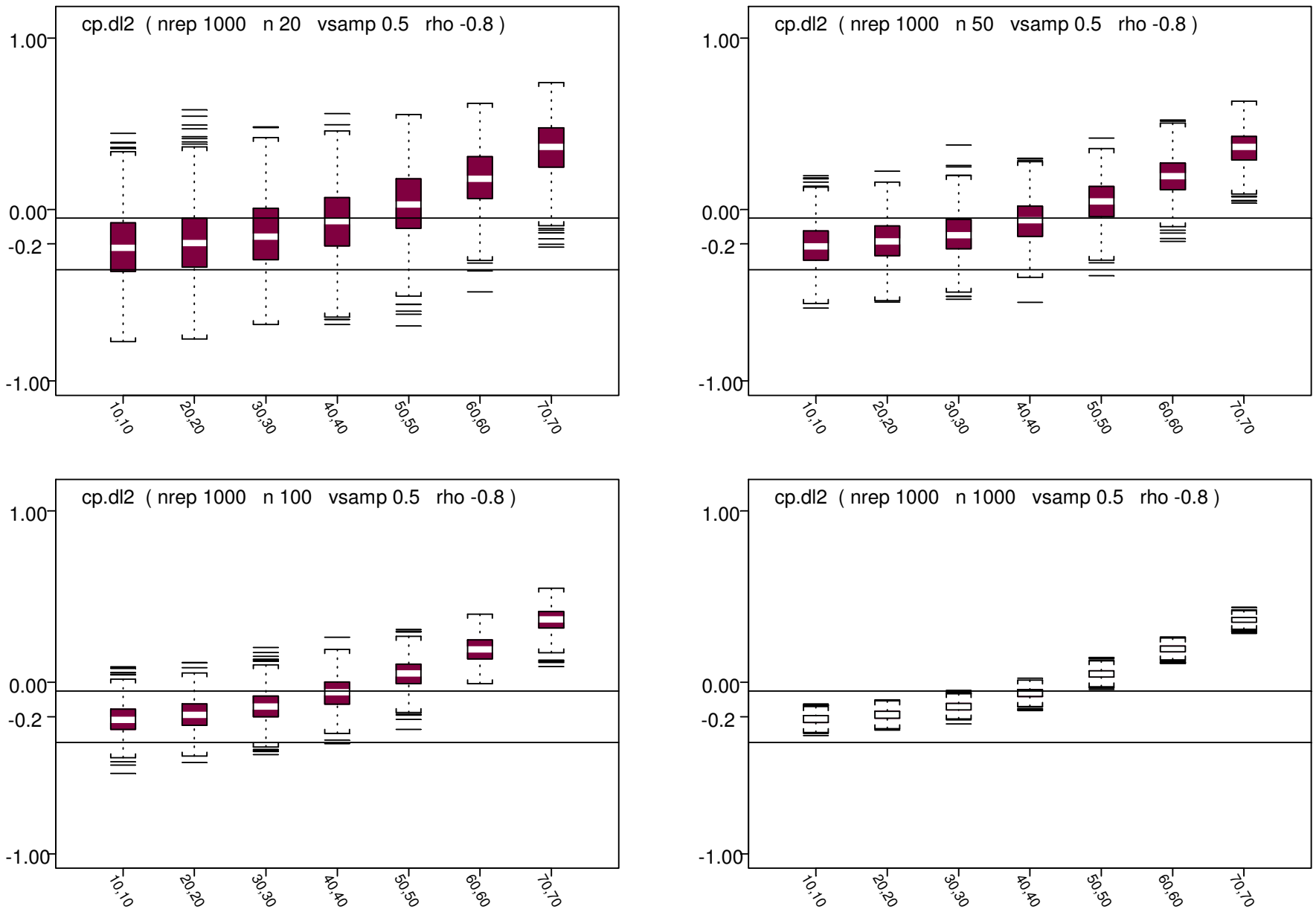

Figure S32. Box plots of estimates of correlation. Estimator = cp dl2. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

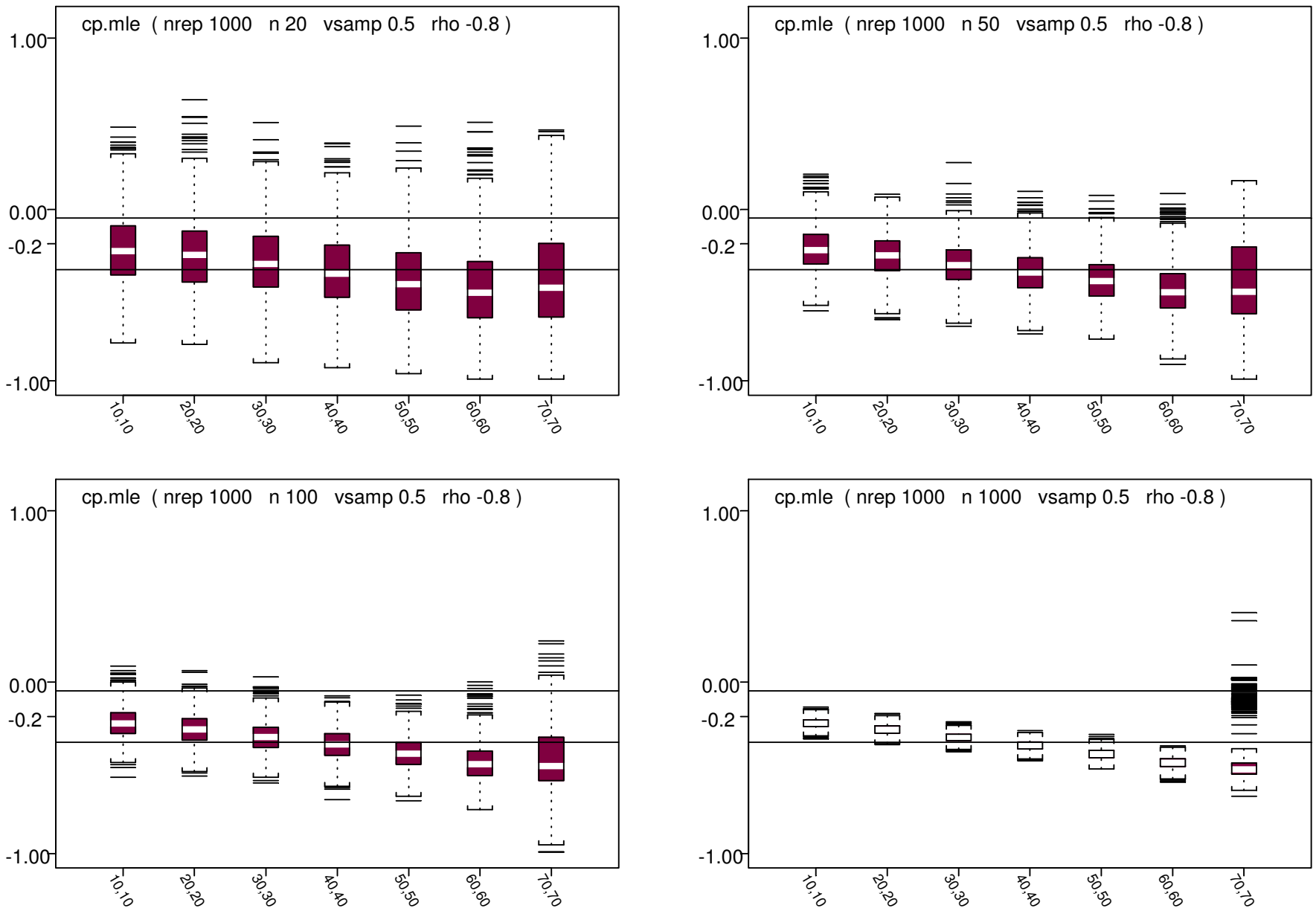

Figure S33. Box plots of estimates of correlation. Estimator = cp mle. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

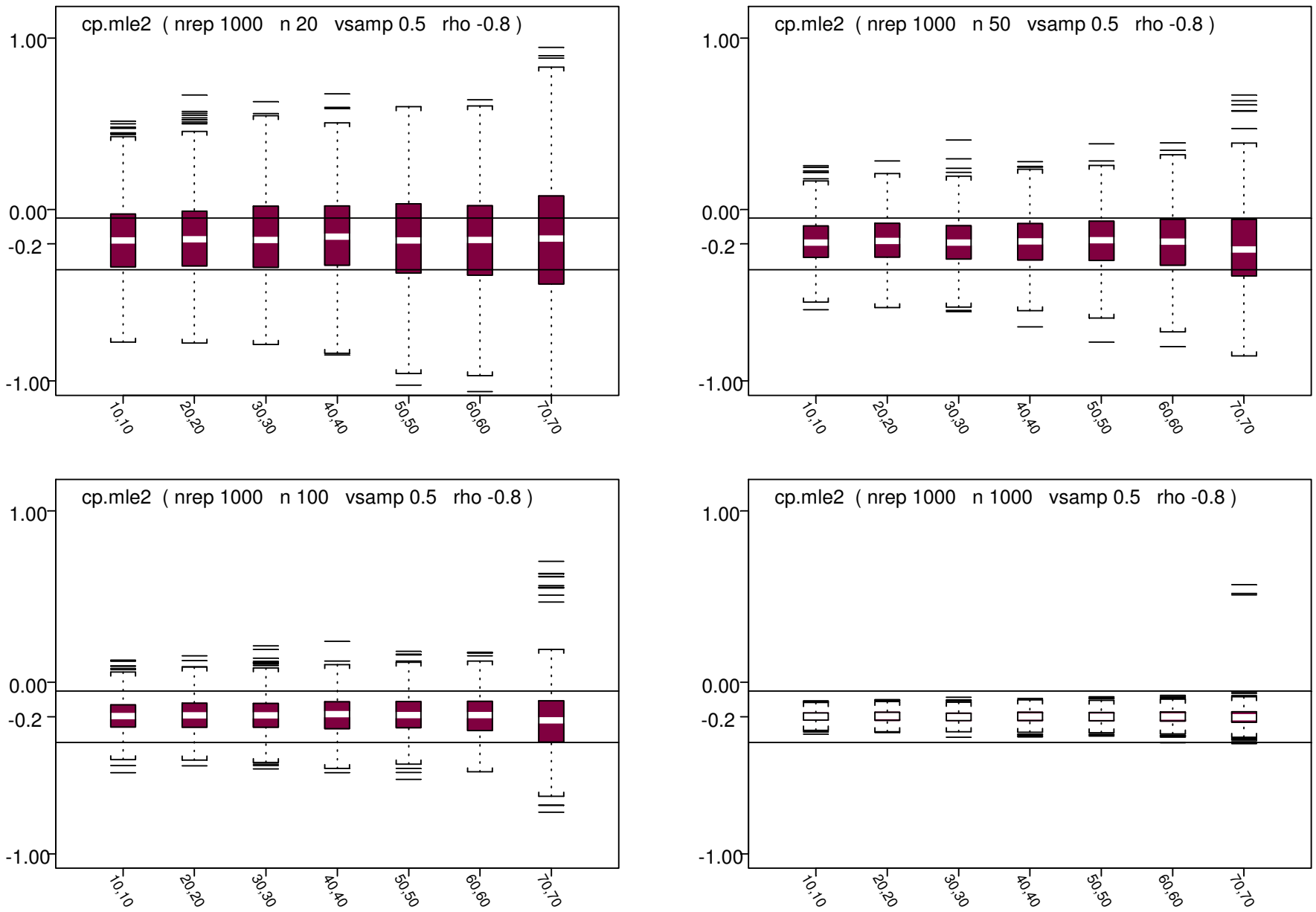

Figure S34. Box plots of estimates of correlation. Estimator = cp mle2. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

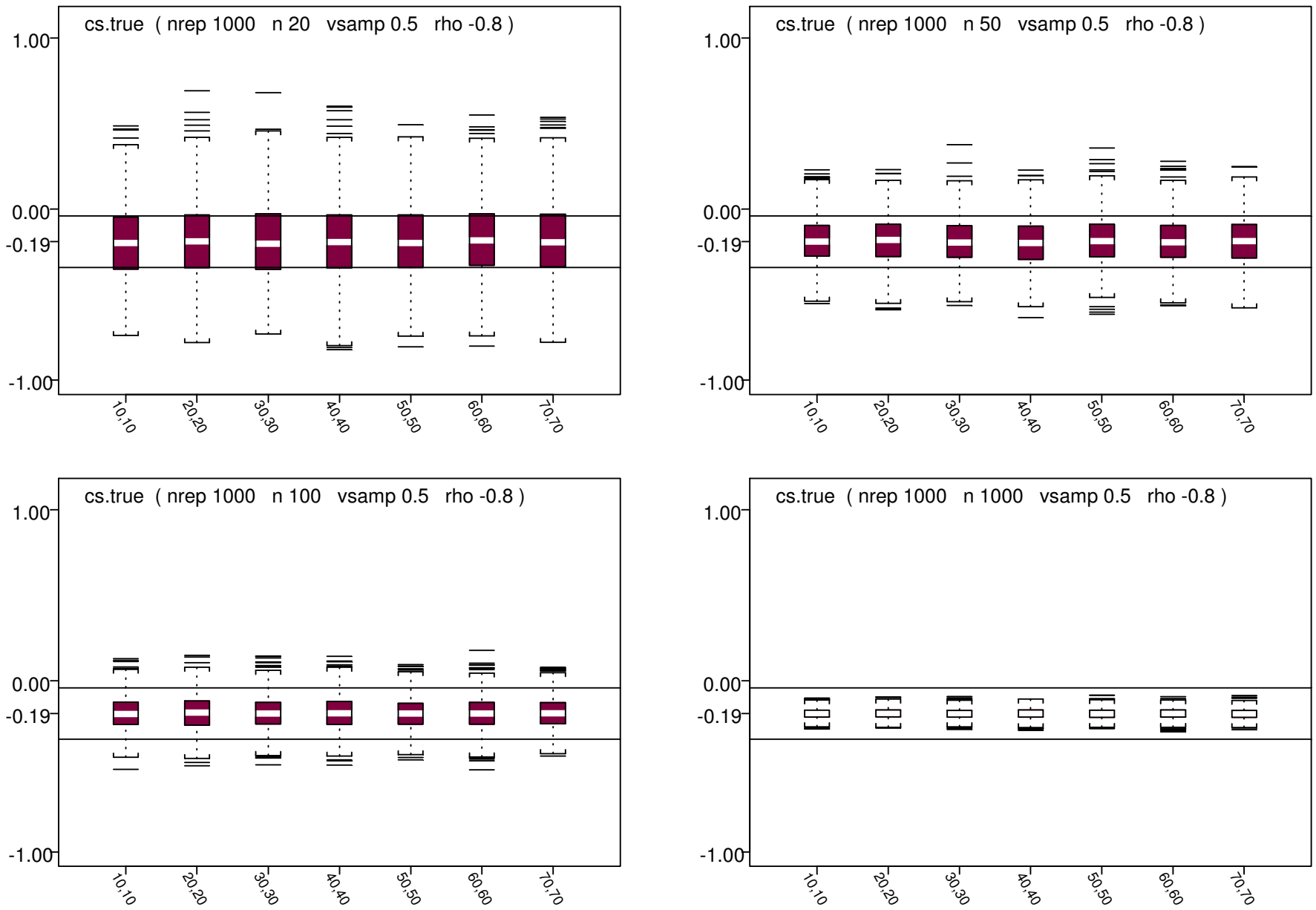

Figure S35. Box plots of estimates of correlation. Estimator = cs true. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

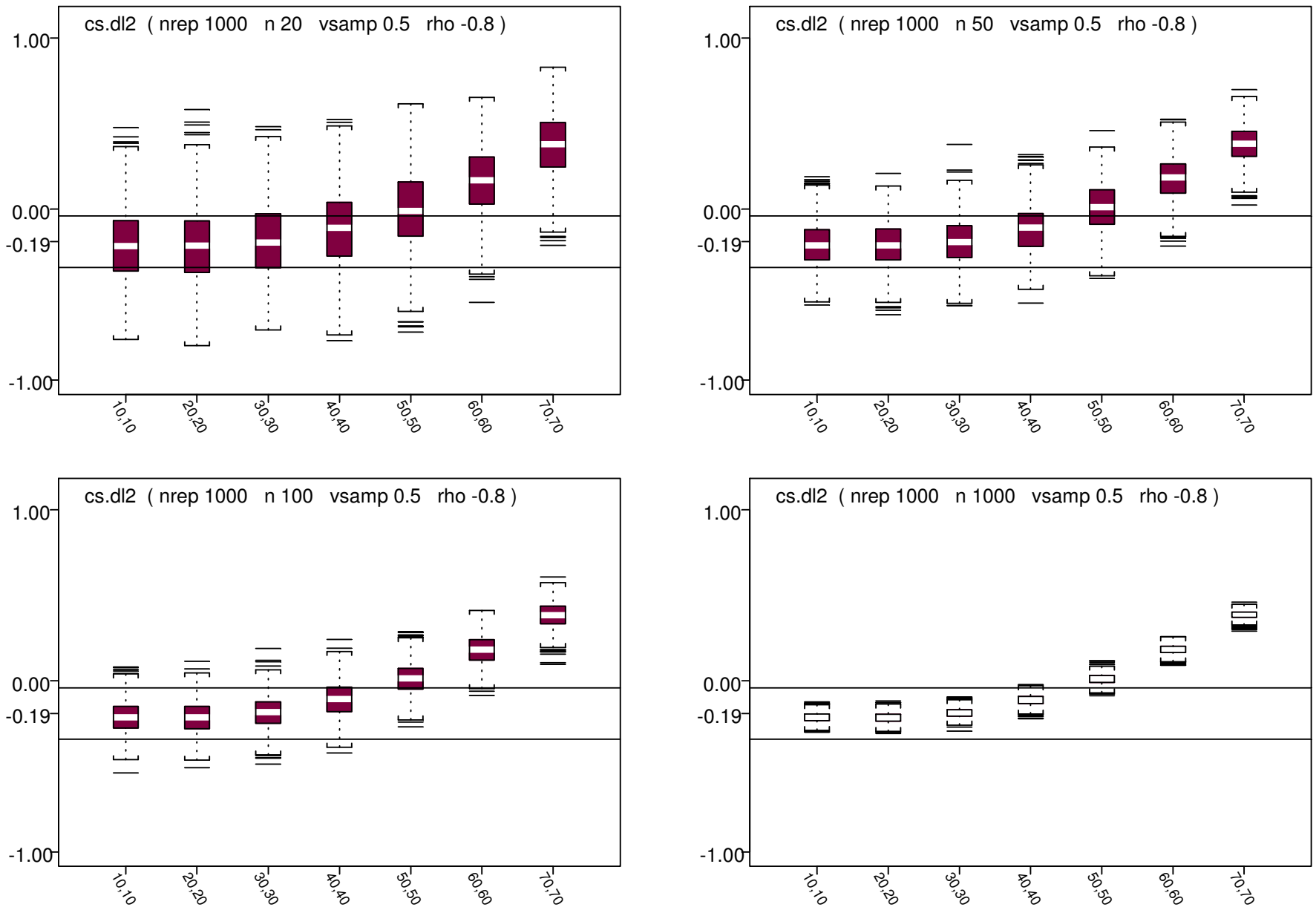

Figure S36. Box plots of estimates of correlation. Estimator = cs dl2. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

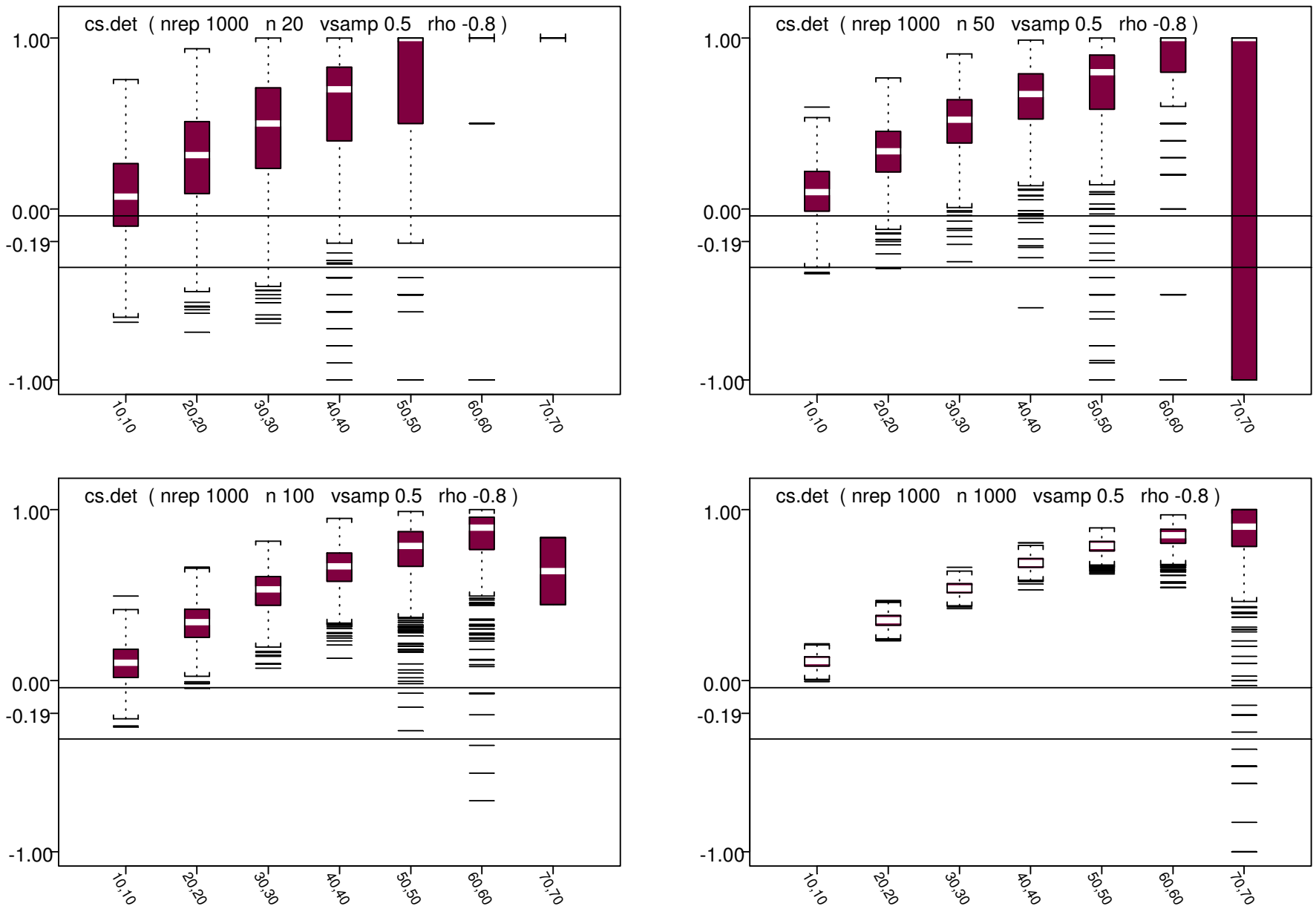

Figure S37. Box plots of estimates of correlation. Estimator = cs det. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

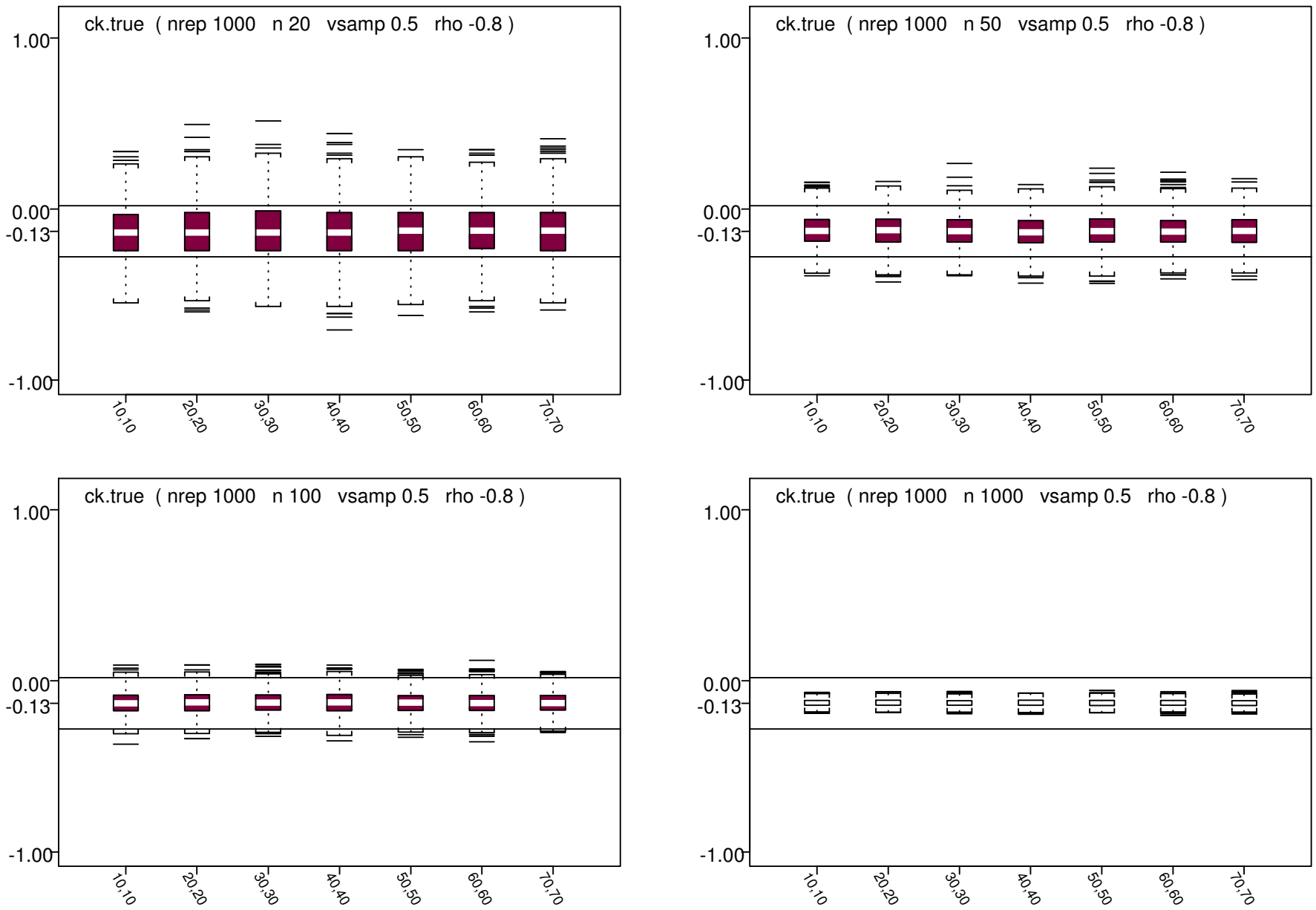

Figure S38. Box plots of estimates of correlation. Estimator = ck true. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

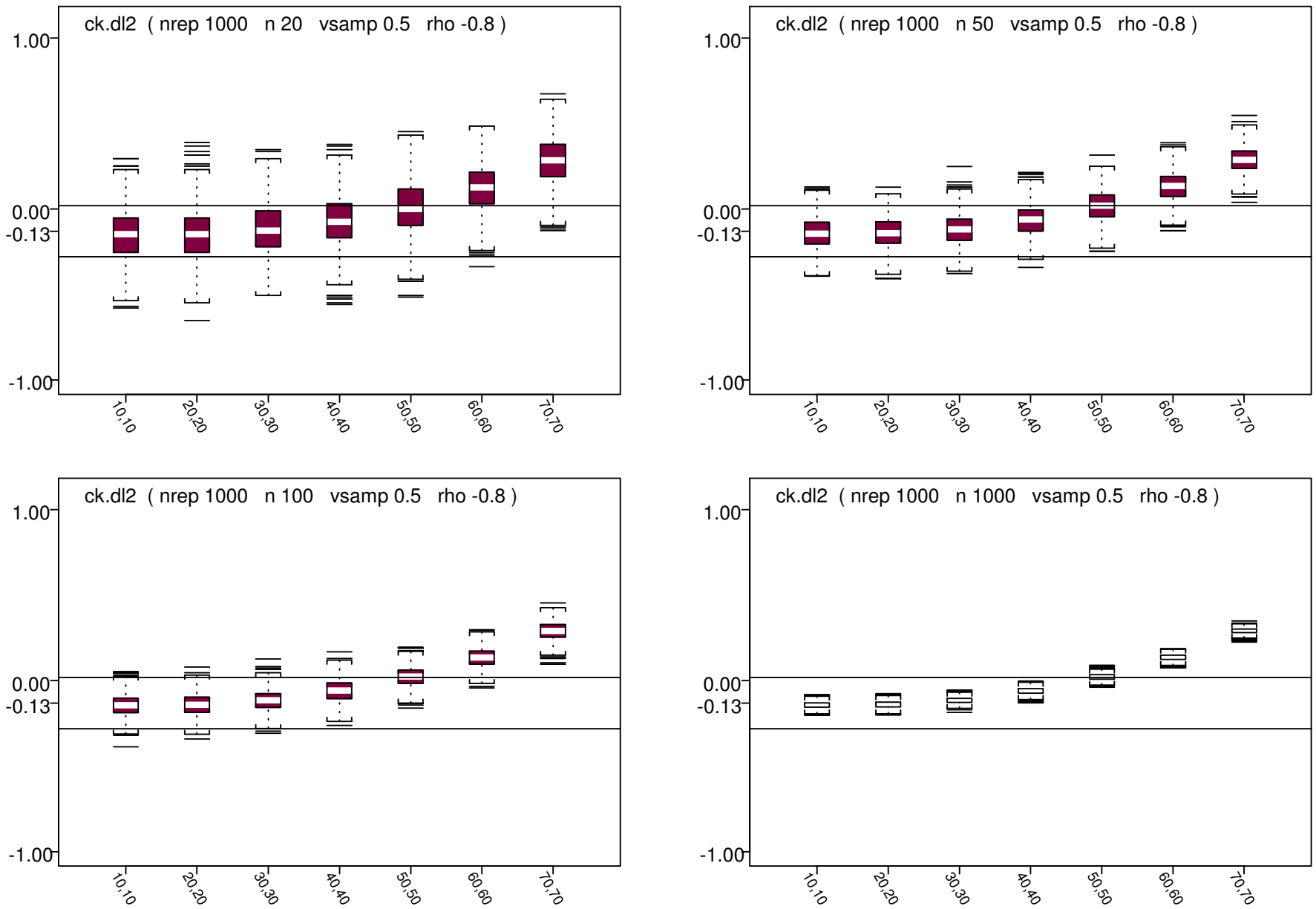

Figure S39. Box plots of estimates of correlation. Estimator = ck dl2. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

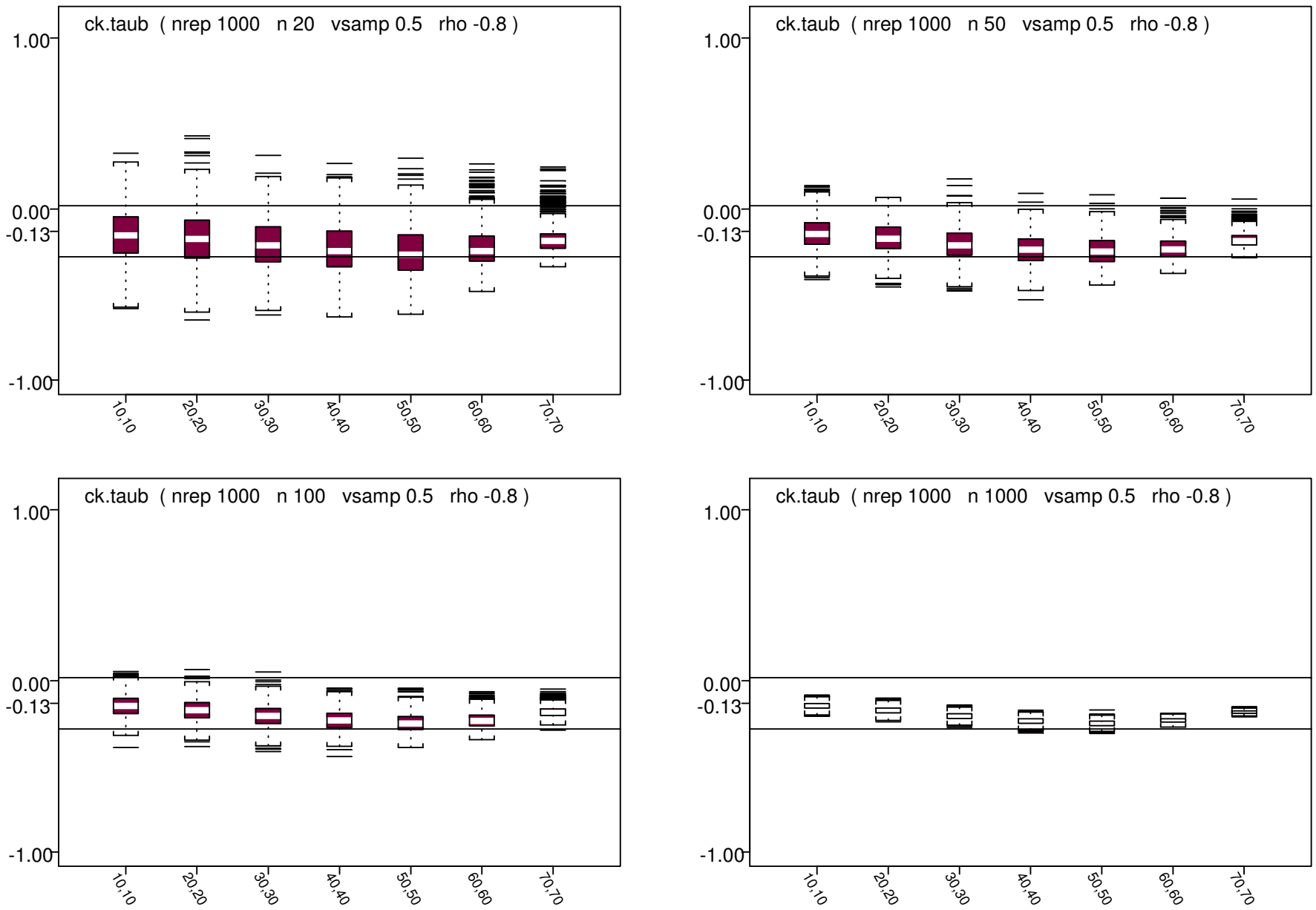

Figure S40. Box plots of estimates of correlation. Estimator = ck taub. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)

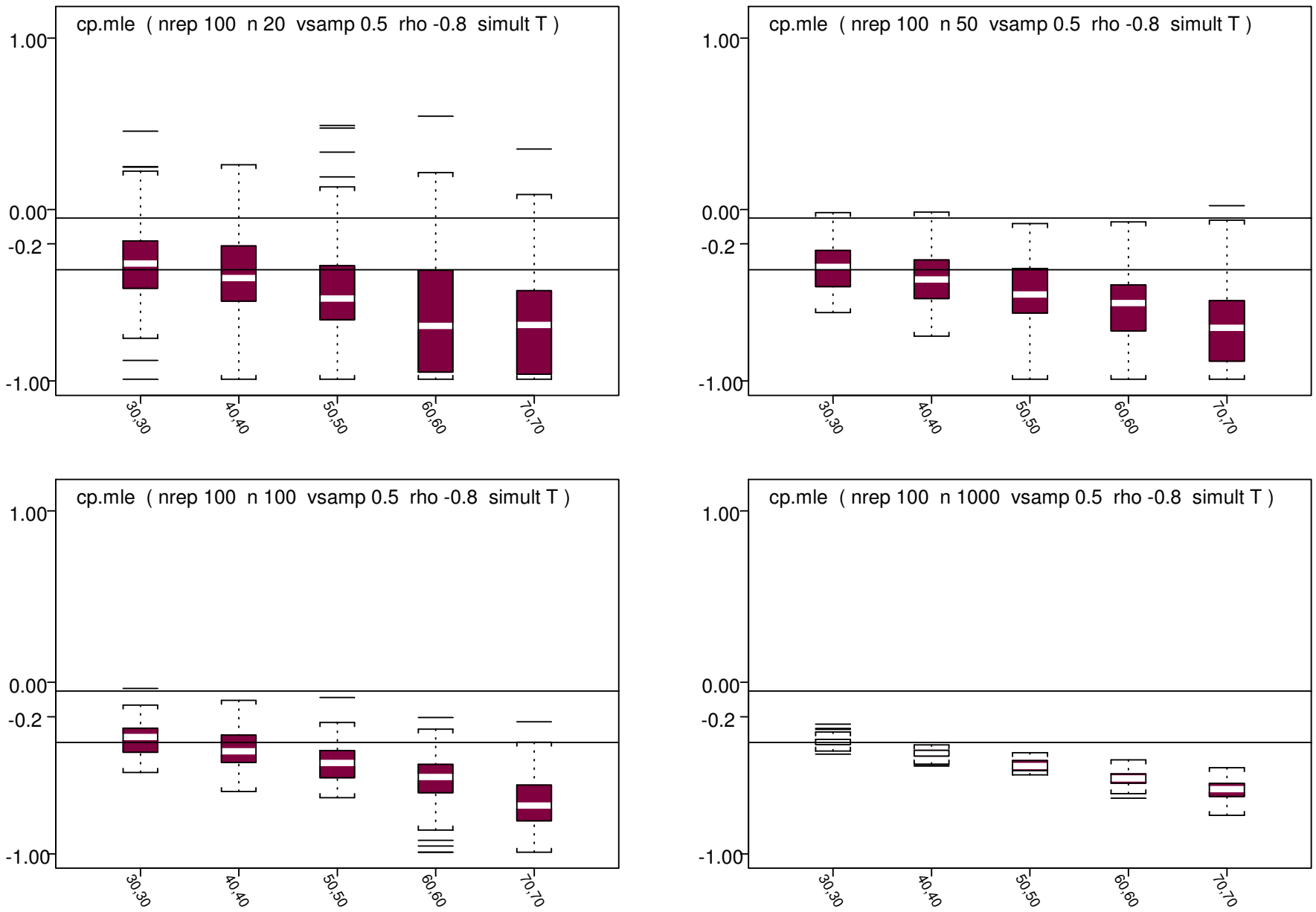

Figure S41. Box plots of estimates of correlation. Estimator = cp.mle. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)
(simult=T indicates correlation estimated simultaneously with means and standard deviations)

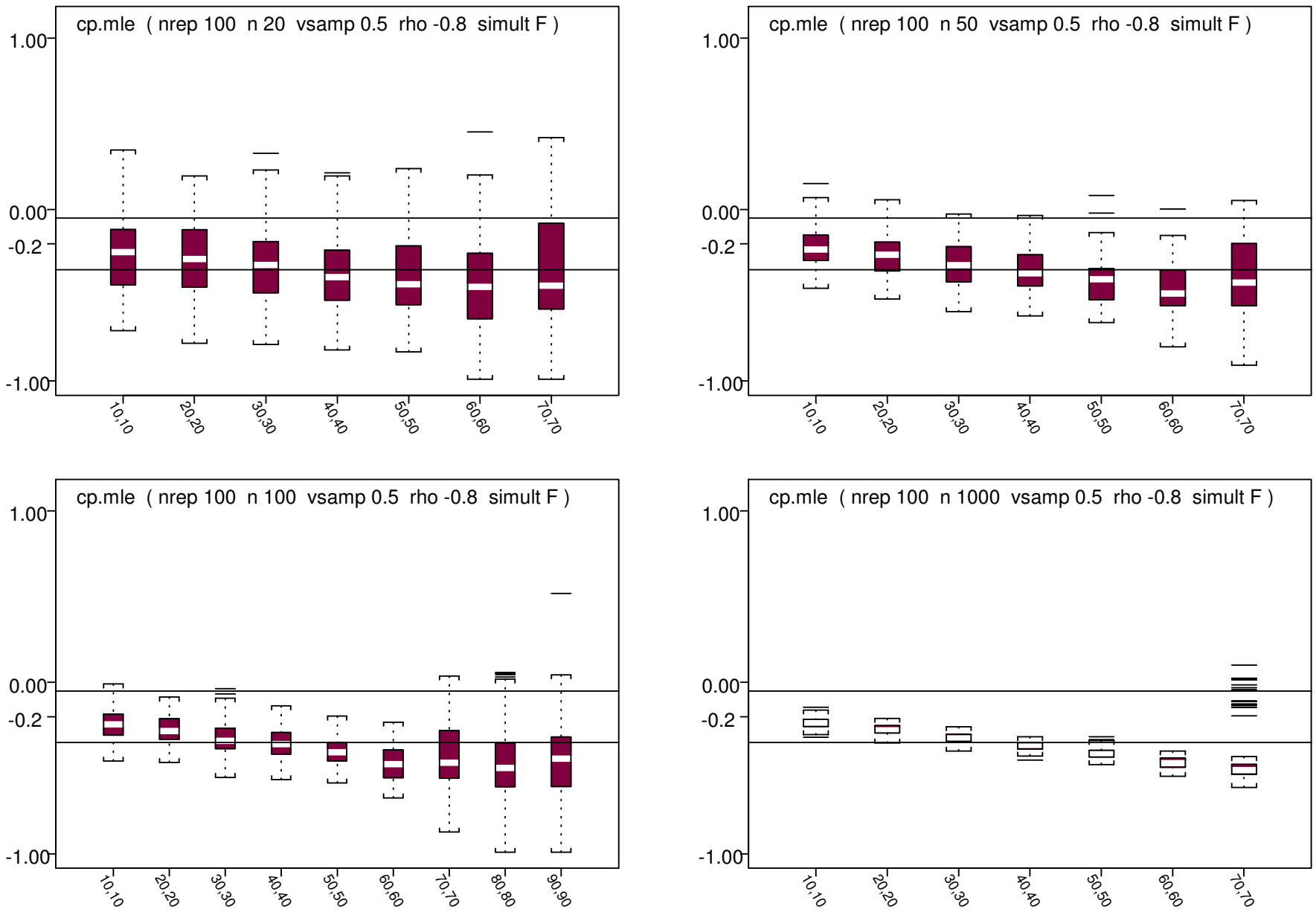

Figure S42. Box plots of estimates of correlation. Estimator = cp.mle. Horizontal axis is percent censored in x and y. Parameters as indicated.
(nrep is number of replications, n is number of observations, vsamp is variance of physical sample size, rho is correlation of unadjusted data)
(simult=F indicates correlations estimated after means and standard deviations)

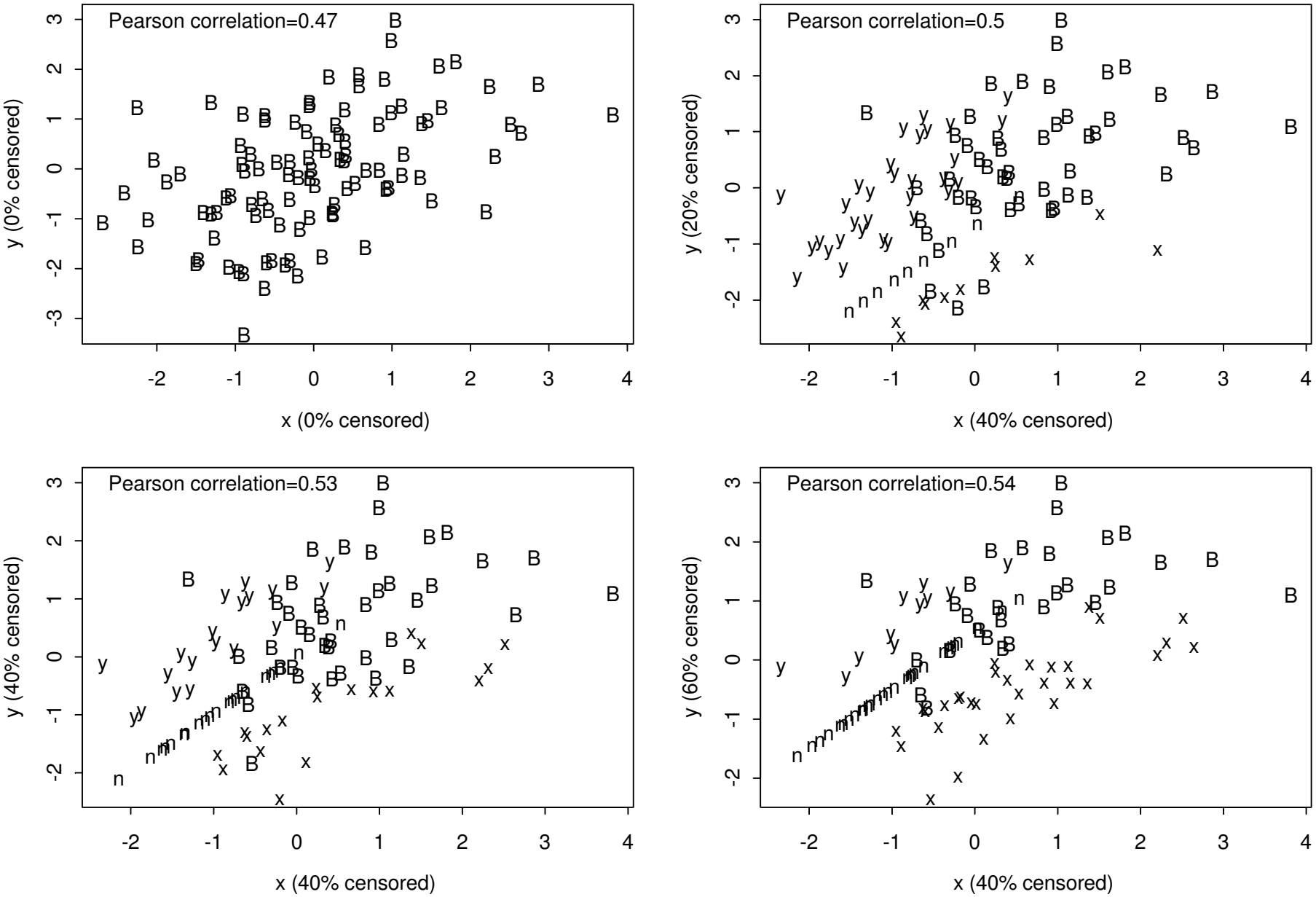

Figure S43. Scatter plots of simulated multiply censored data.
(B=both detected; x=x detected, y at DL/2; y=y detected, x at DL/2; n=both at DL/2)
(number of observations=100, variance of physical sample size=0.5, correlation of unadjusted data=0.2)

Example of computation of Kendall's tau and Kendall's tau-b.

If x = (5, 4, 3, 2, 6) and y = (1, 2, 3, 2, 4), there are N=5*4/2 = 10 comparisons in x and in y:

x_i	x_j	Sign(x_i-x_j)	y_i	y_j	Sign(y_i-y_j)	Sign(x_i-x_j) * Sign(y_i-y_j)
5	4	1	1	2	-1	-1
5	3	1	1	3	-1	-1
5	2	1	1	2	-1	-1
5	6	-1	1	4	-1	1
4	3	1	2	3	-1	-1
4	2	1	2	2	0	0
4	6	-1	2	4	-1	1
3	2	1	3	2	1	1
3	6	-1	3	4	-1	1
2	6	-1	2	4	-1	1

Here, the Kendall correlation = (Ncon–Ndis)/sqrt(N-ntx)(N-nty)) = 0.105, where Ncon = number concordant=5, Ndis = number discordant=4, ntx = number of ties in x=0, nty = number of ties in y=1. By comparison, the Pearson correlation = 0.277 and the Spearman correlation = 0.205.

If the data are censored, a common practice is to impute censored values as the detection limit divided by a constant. For example, if x = (5, <4, <3, 2, 6) and y = (1, <2, 3, 2, <4), and the non-detects are set to the detection limits, the correlations are estimated as above. However, if the non-detects are set to the detection limit divided by two, the estimated correlations are:
rp= -0.321, rs= -0.433, and rk= -0.354.

In estimating Kendall's tau-b, a non-detect compared with a non-detect is considered a tie regardless of the values of the two DL's. A non-detect compared with a detect is considered a tie if the detect value is lower than the DL for the non-detect. This method can result in a large number of ties and difficulties in estimating the variance of the statistic *(4)*.

For instance, using the above example, the comparisons are:

x_i	x_j	Sign(x_i-x_j)	y_i	y_j	Sign(y_i-y_j)	Sign(x_i-x_j) * Sign(y_i-y_j)
5	<4	1	1	<2	0	0
5	<3	1	1	3	-1	-1
5	2	1	1	2	-1	-1
5	6	-1	1	<4	0	0
<4	<3	0	<2	3	-1	0
<4	2	0	<2	2	-1	0
<4	6	-1	<2	<4	0	0
<3	2	0	3	2	1	0
<3	6	-1	3	<4	0	0
2	6	-1	2	<4	0	0

Here Ncon = 0, Ndis = 2, ntx = 3 and nty = 5, so that tau= -0.338.