Supplement to "Simulation-based comparison of multivariate ensemble post-processing methods"

Sebastian Lerch*[1], Sándor Baran[2], Annette Möller[3], Jürgen Groß[4], Roman Schefzik[5], Stephan Hemri[6] and Maximiliane Graeter[1]

[1]Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]University of Debrecen, Debrecen, Hungary
[3]Technical University of Clausthal, Clausthal, Germany
[4]University of Hildesheim, Hildesheim, Germany
[5]Federal Office of Meteorology and Climatology MeteoSwiss Zurich-Airport, Switzerland
[6]German Cancer Research Center (DKFZ), Heidelberg, Germany

December 30, 2019

The supplemental material provides additional figures with results for all parameter combination in all simulation settings.

1 Additional results for Setting 1

An overview the additional results for Setting 1 is provided in Table 1.

*Corresponding author, Sebastian.Lerch@kit.edu

Table 1: Overview of parameter combinations for Setting 1 with references to corresponding figures in this supplement.

ϵ	σ	Figure
0	0.5	1
0	1	2
0	$\sqrt{2}$	3
0	$\sqrt{5}$	4
1	0.5	5
1	1	6
1	$\sqrt{2}$	7
1	$\sqrt{5}$	8
3	0.5	9
3	1	10
3	$\sqrt{2}$	11
3	$\sqrt{5}$	12

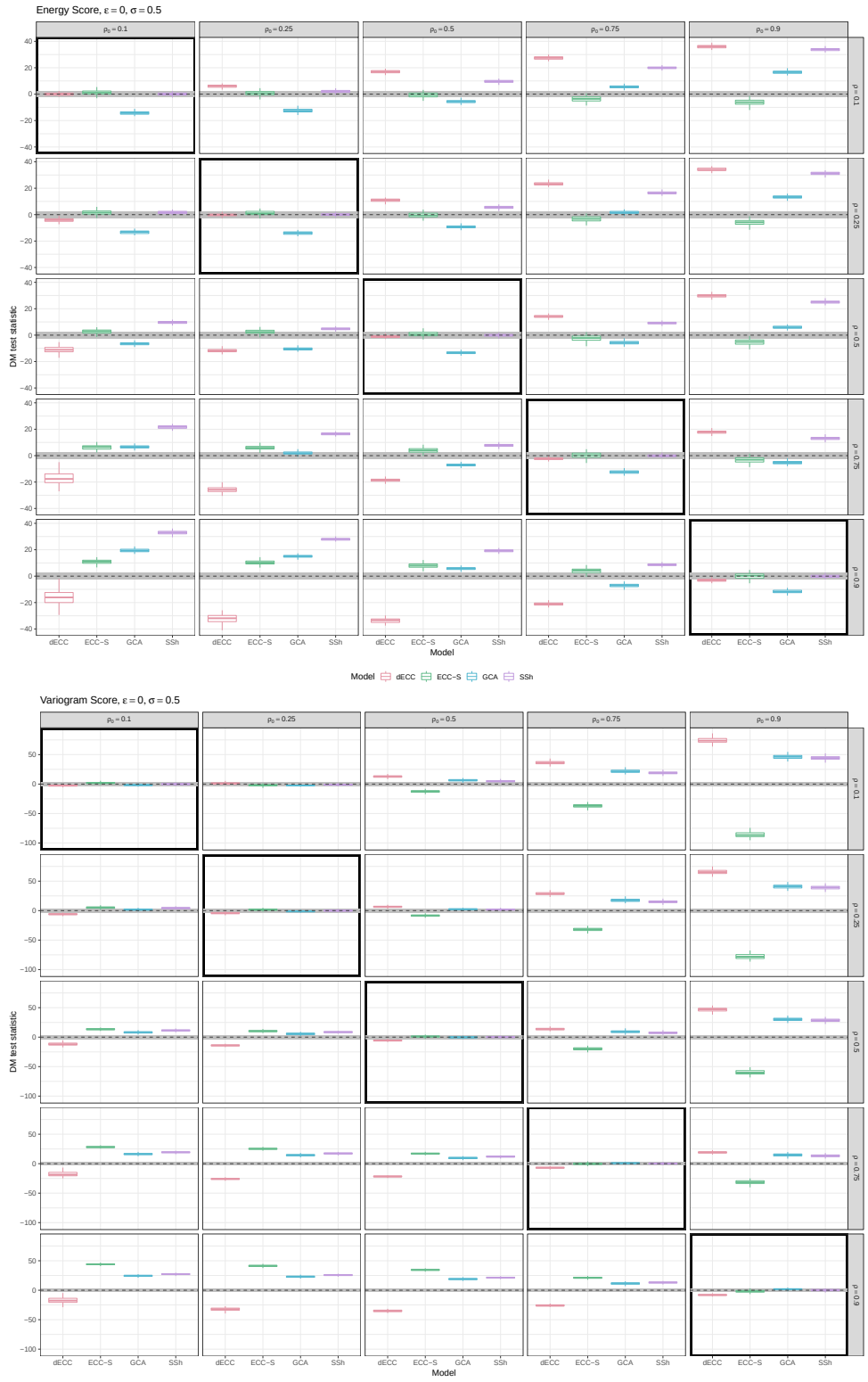
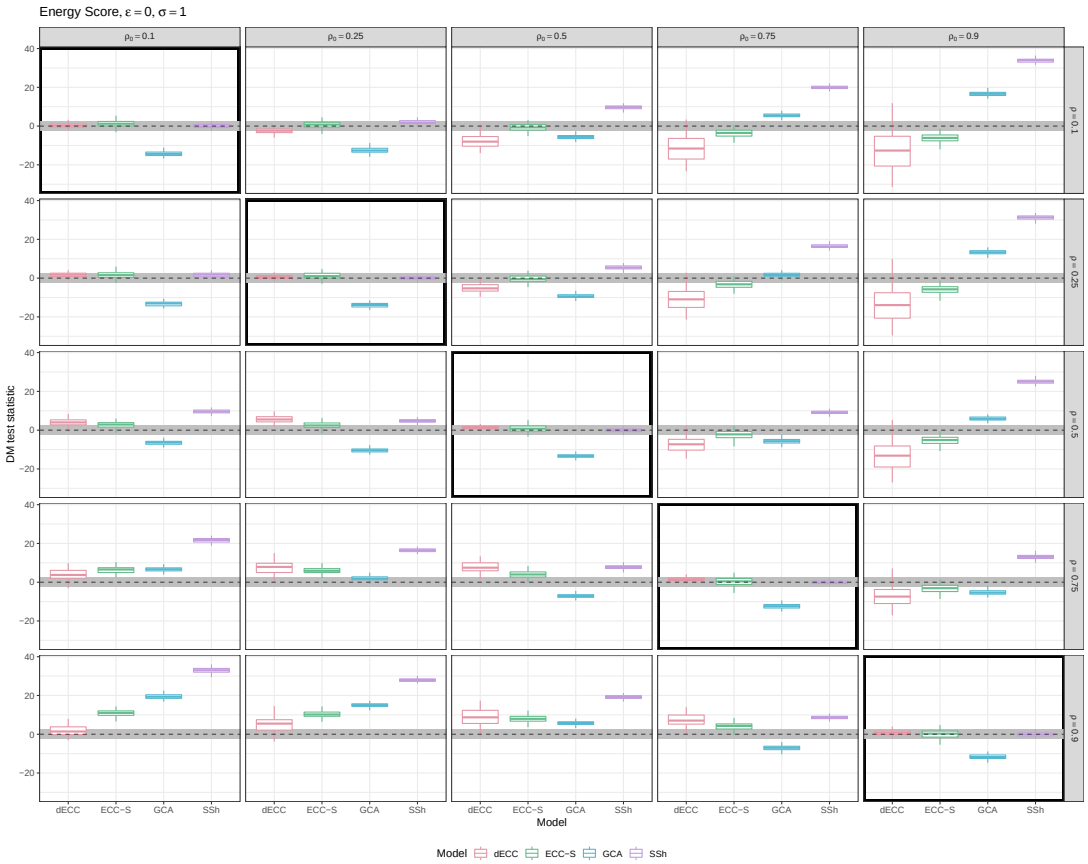


Figure 1: Summaries of DM test statistic values based on ES (top) and VS (bottom) for Setting 1 with $\epsilon = 0$, and $\sigma = 0.5$. ECC-Q forecasts are used as reference model such that positive values of the test statistic indicate improvements over ECC-Q and negative value indicated deterioration of forecast skill. Boxplots summarize results of the 100 Monte Carlo repetitions of each individual experiment. The horizontal gray stripe indicates the acceptance region of the two-sided DM test under the null hypothesis of equal predictive performance at a level of 0.05. Simulation parameter choices where the correlation structure of the raw ensemble is correctly specified ($\rho = \rho_0$) are surrounded by black boxes.



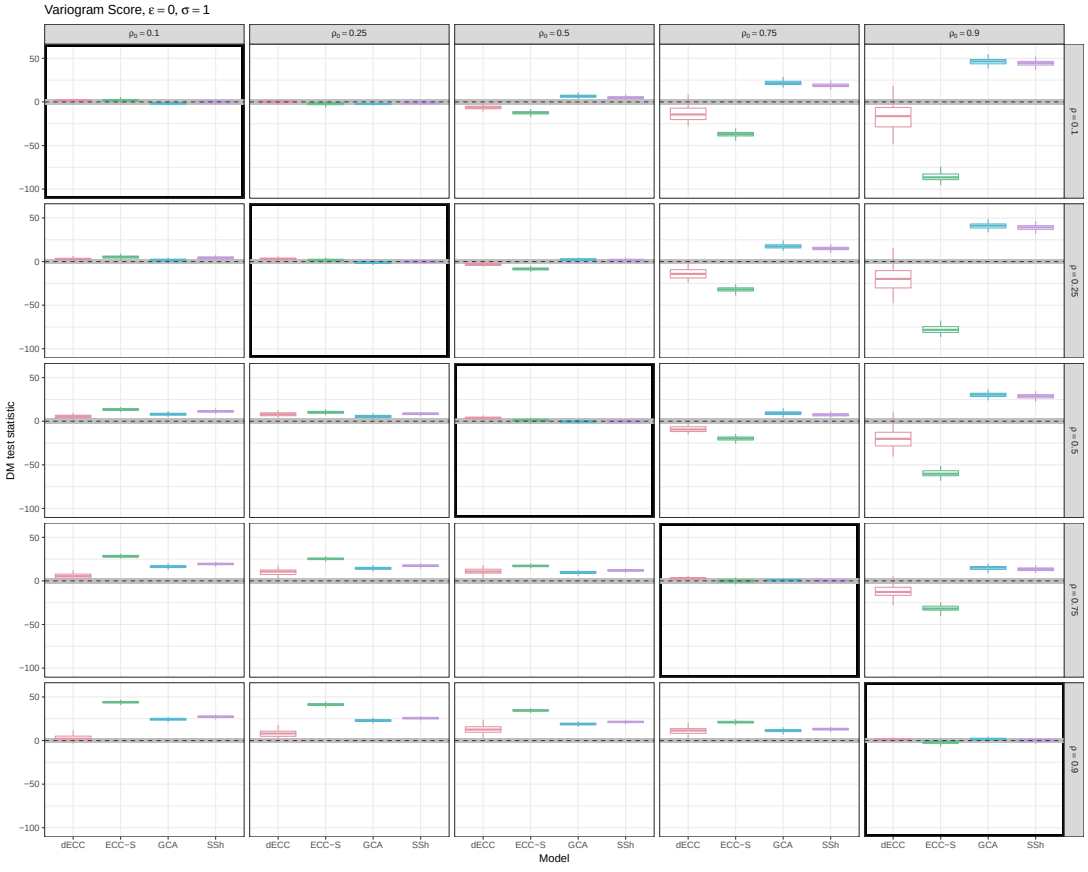
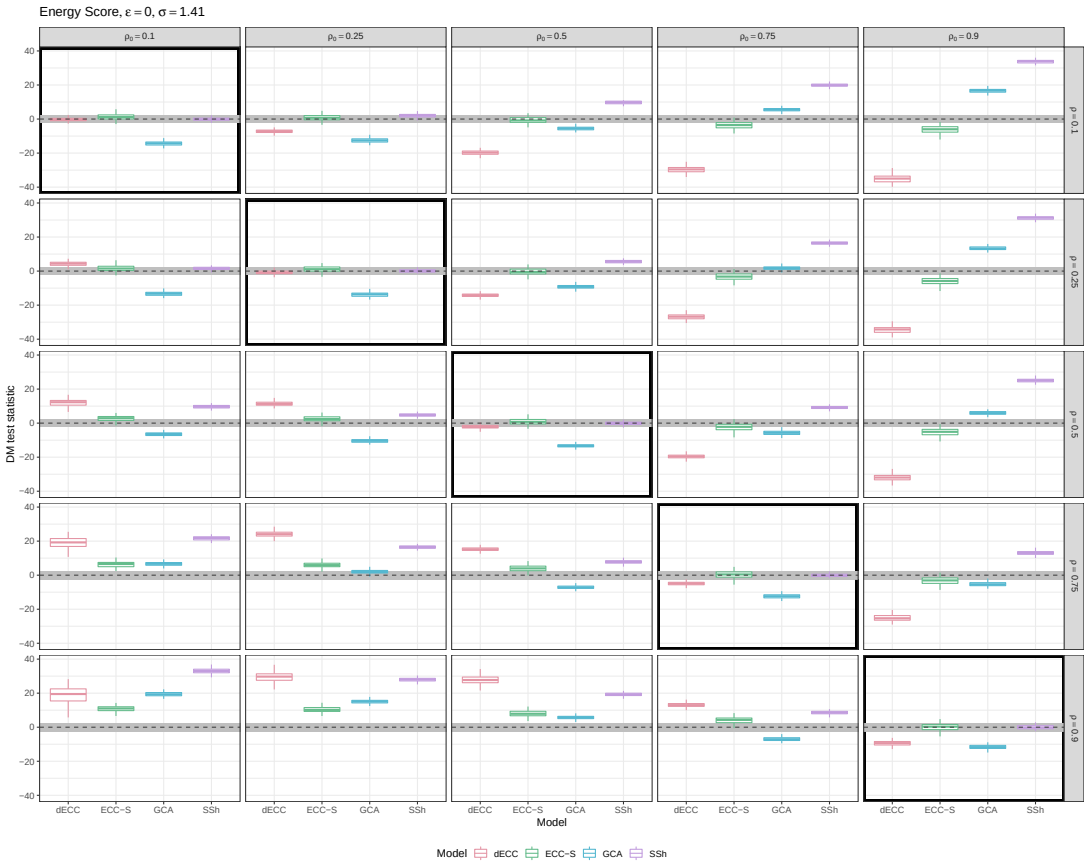


Figure 2: As Figure 1, but for $\epsilon = 0$, and $\sigma = 1$.



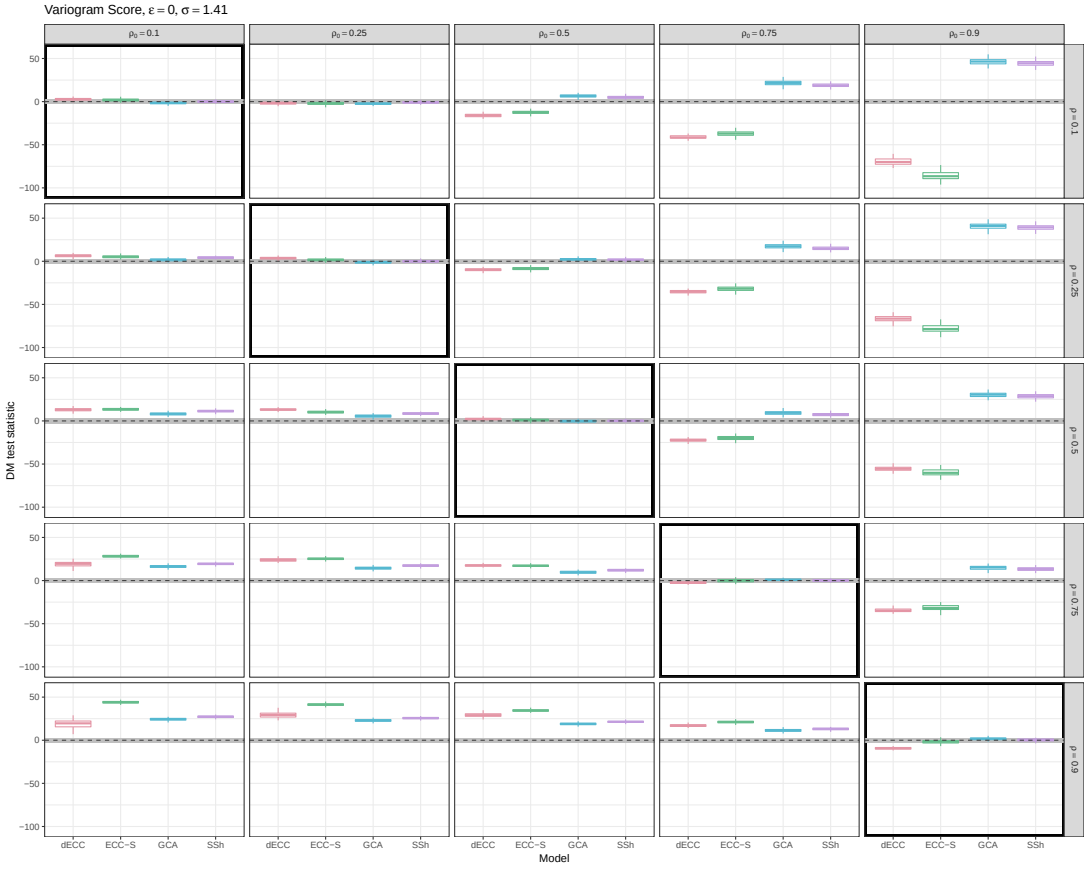
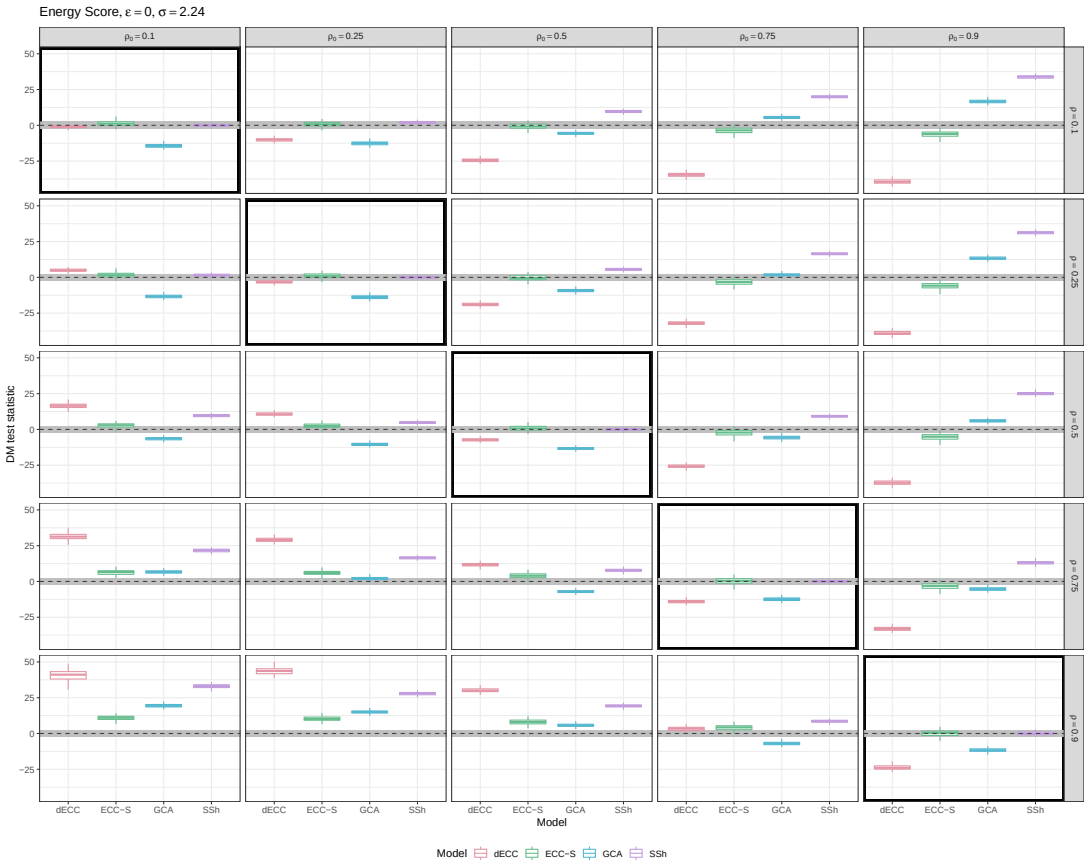


Figure 3: As Figure 1, but for $\epsilon = 0$, and $\sigma = \sqrt{2}$.



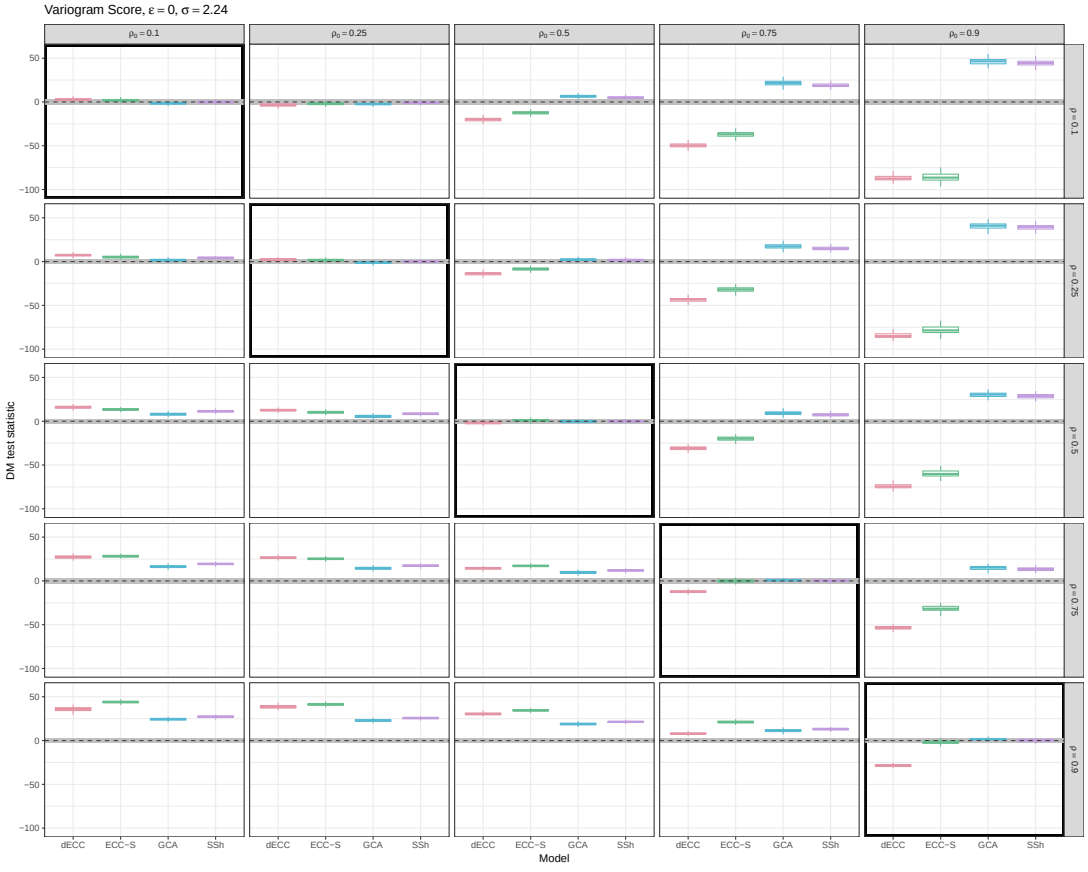


Figure 4: As Figure 1, but for $\epsilon = 0$, and $\sigma = \sqrt{5}$.

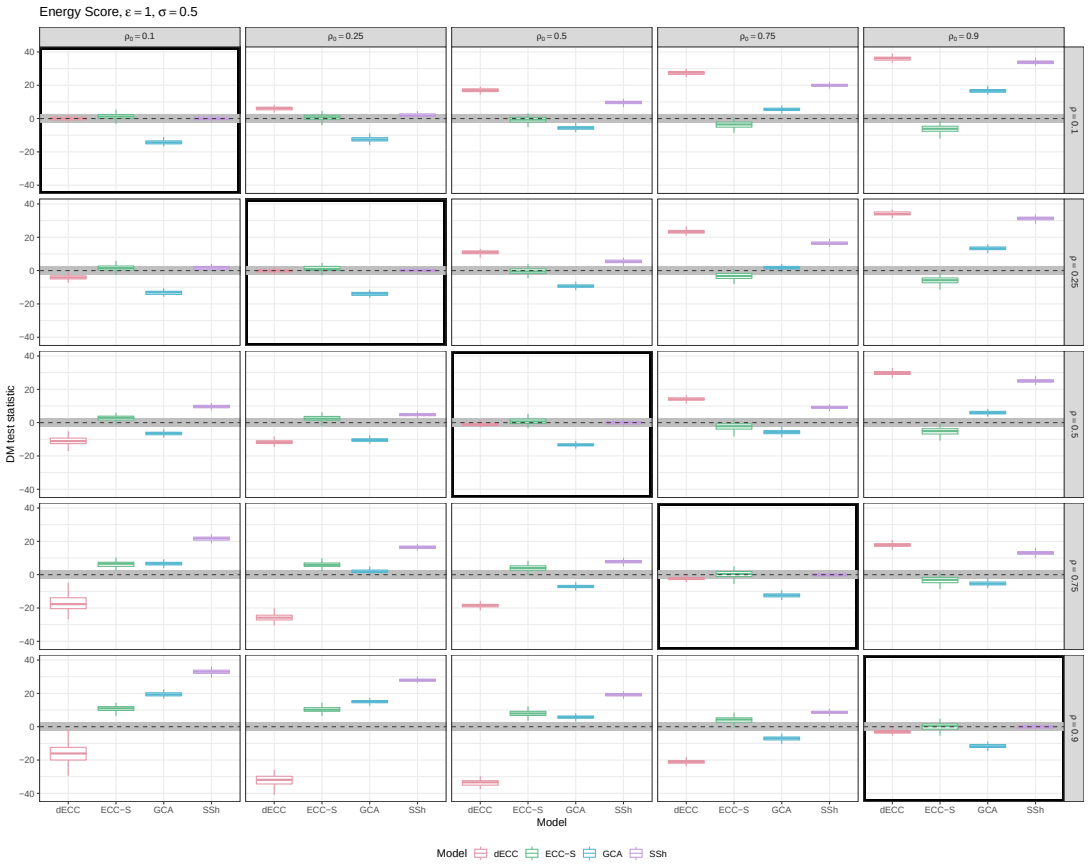
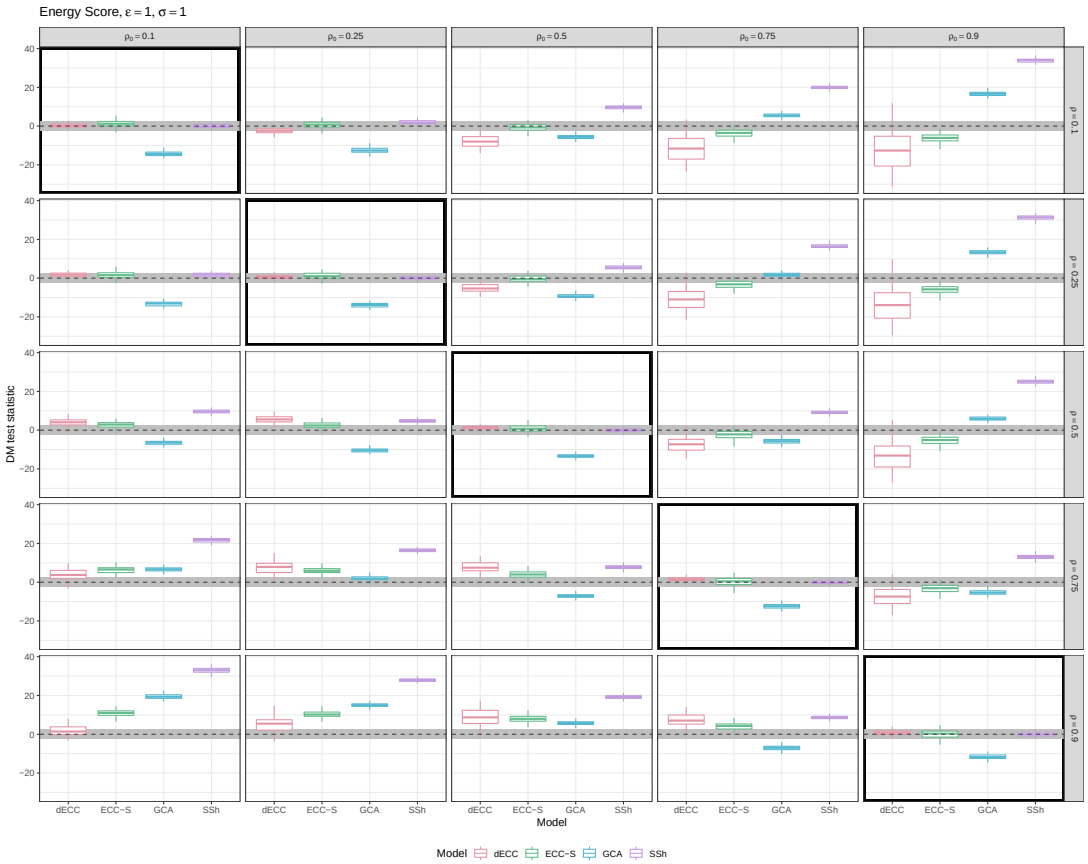


Figure 5: As Figure 1, but for $\epsilon = 1$, and $\sigma = 0.5$.



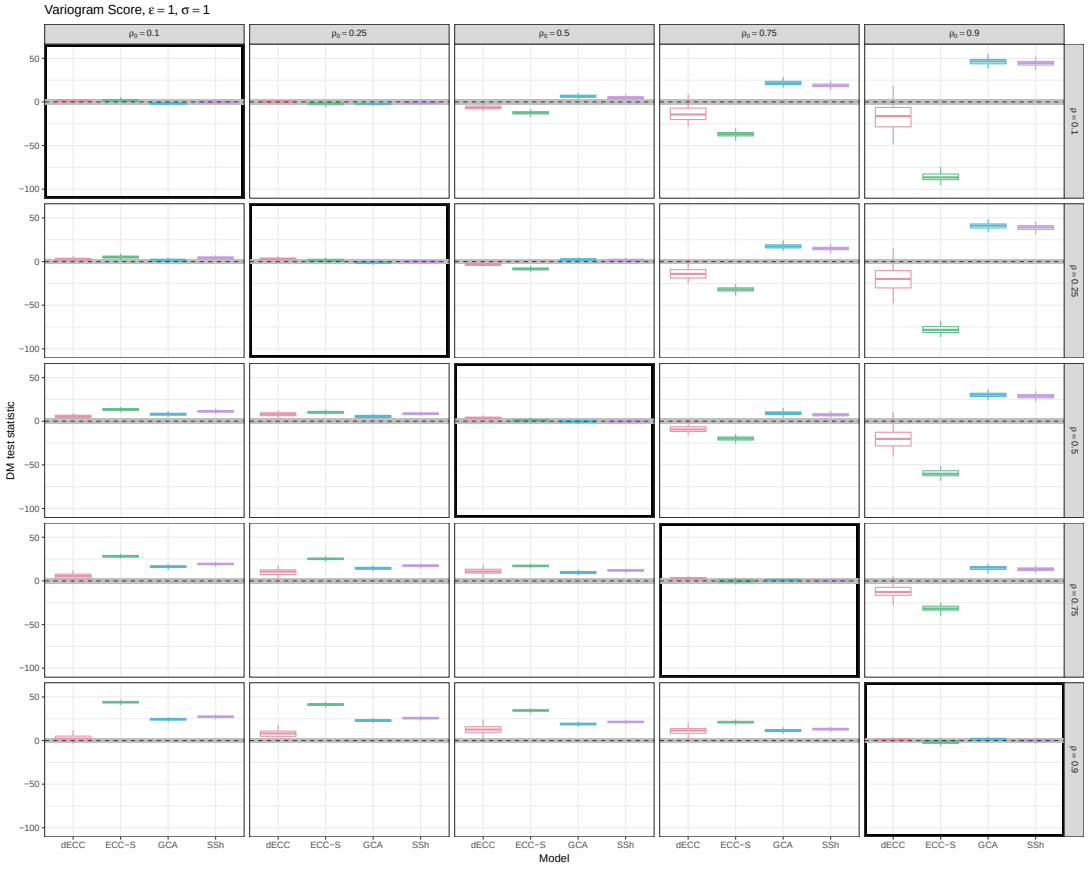
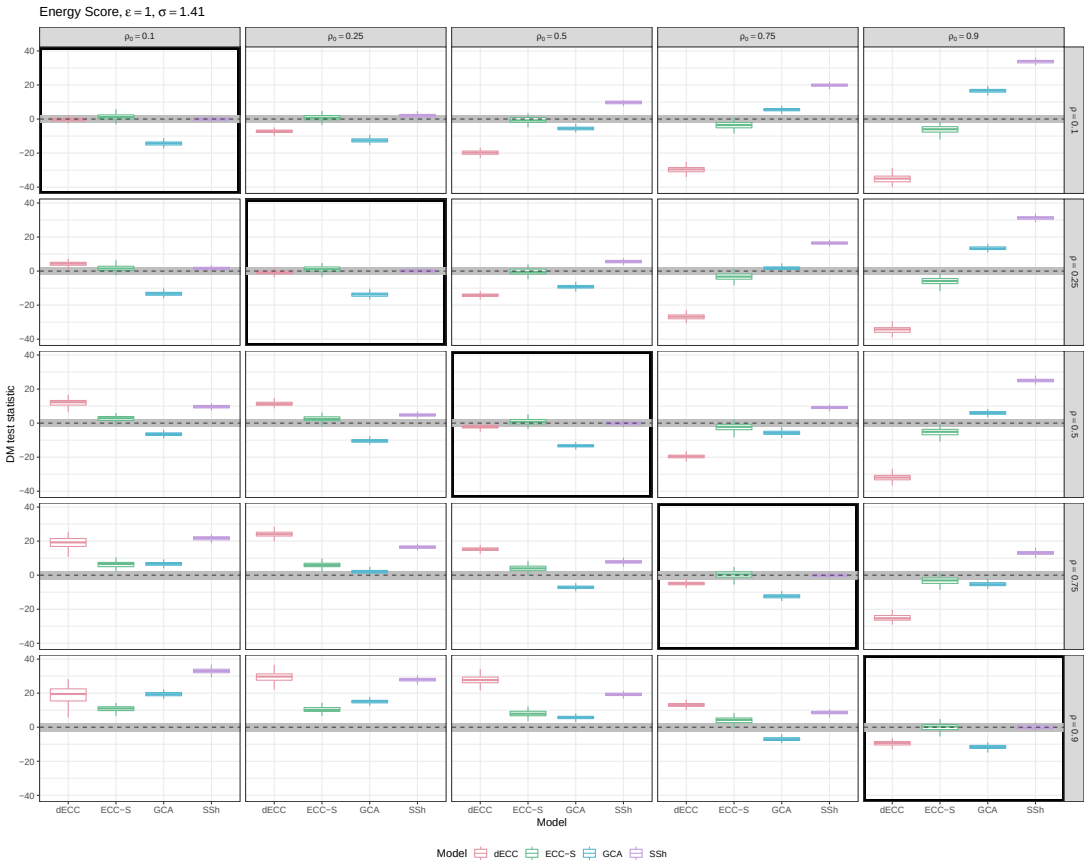


Figure 6: As Figure 1, but for $\epsilon = 1$, and $\sigma = 1$.



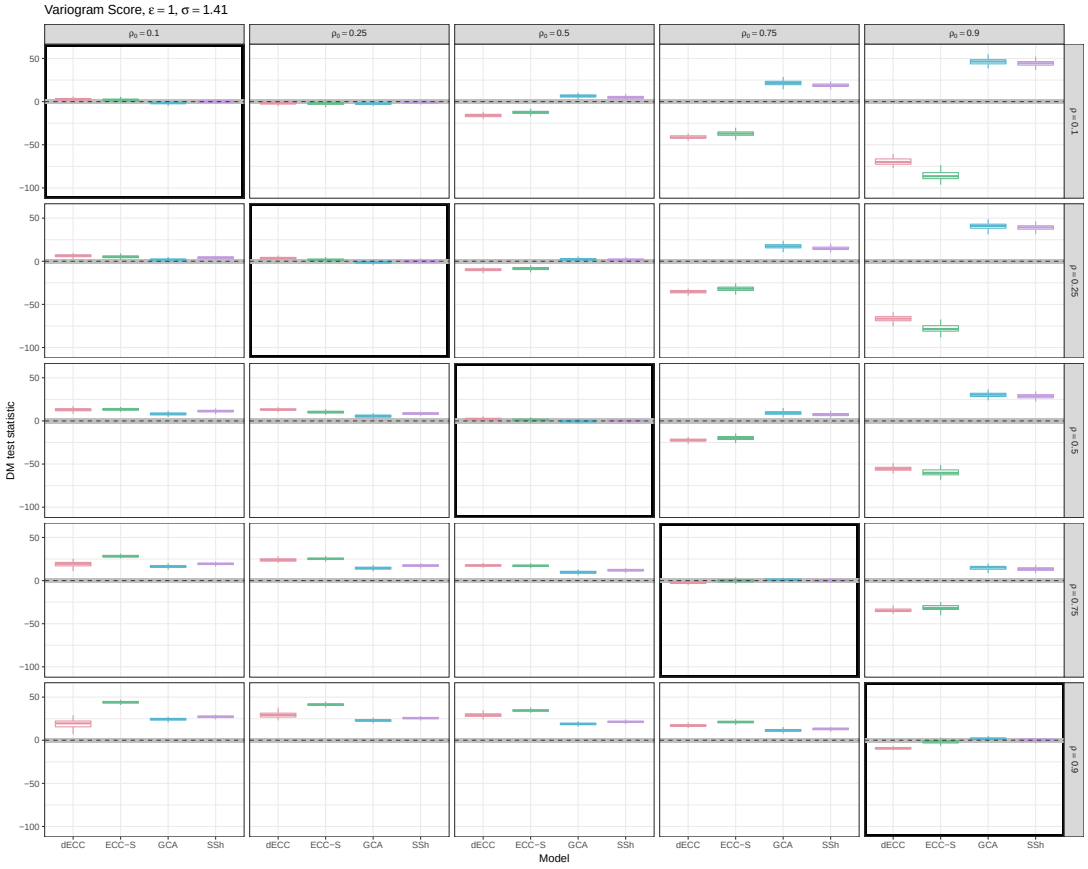
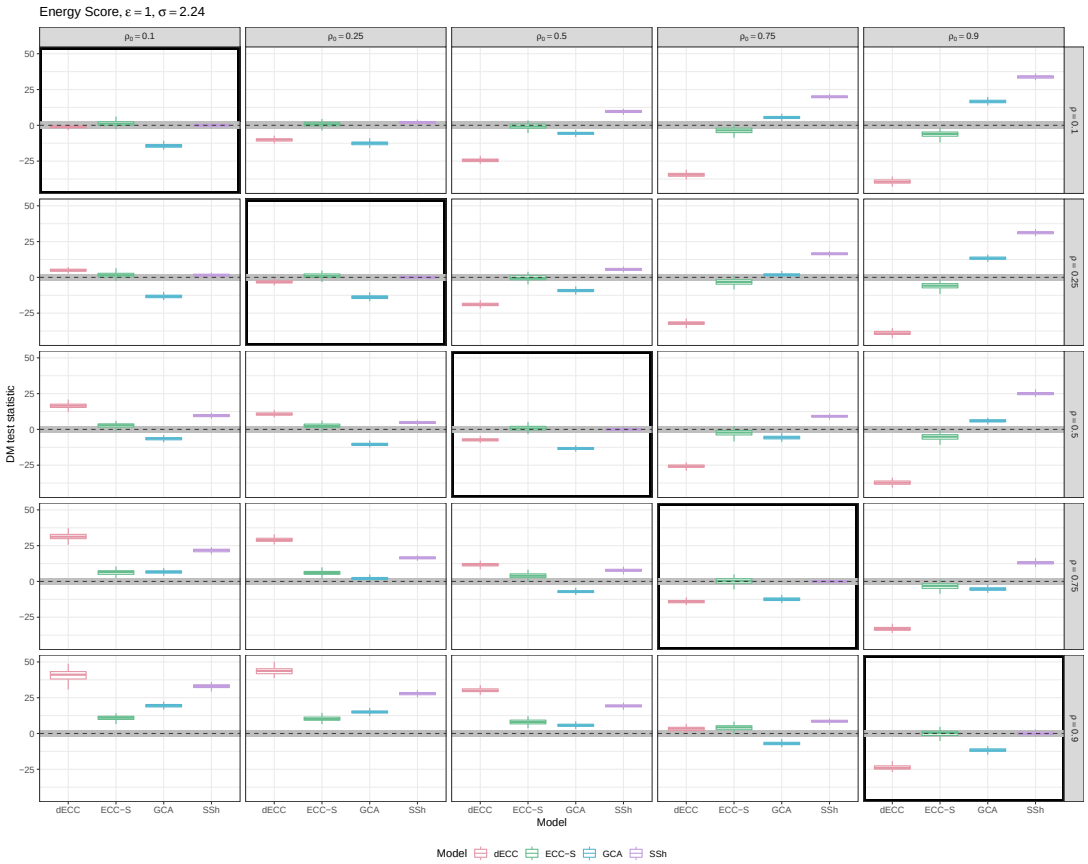


Figure 7: As Figure 1, but for $\epsilon = 1$, and $\sigma = \sqrt{2}$.



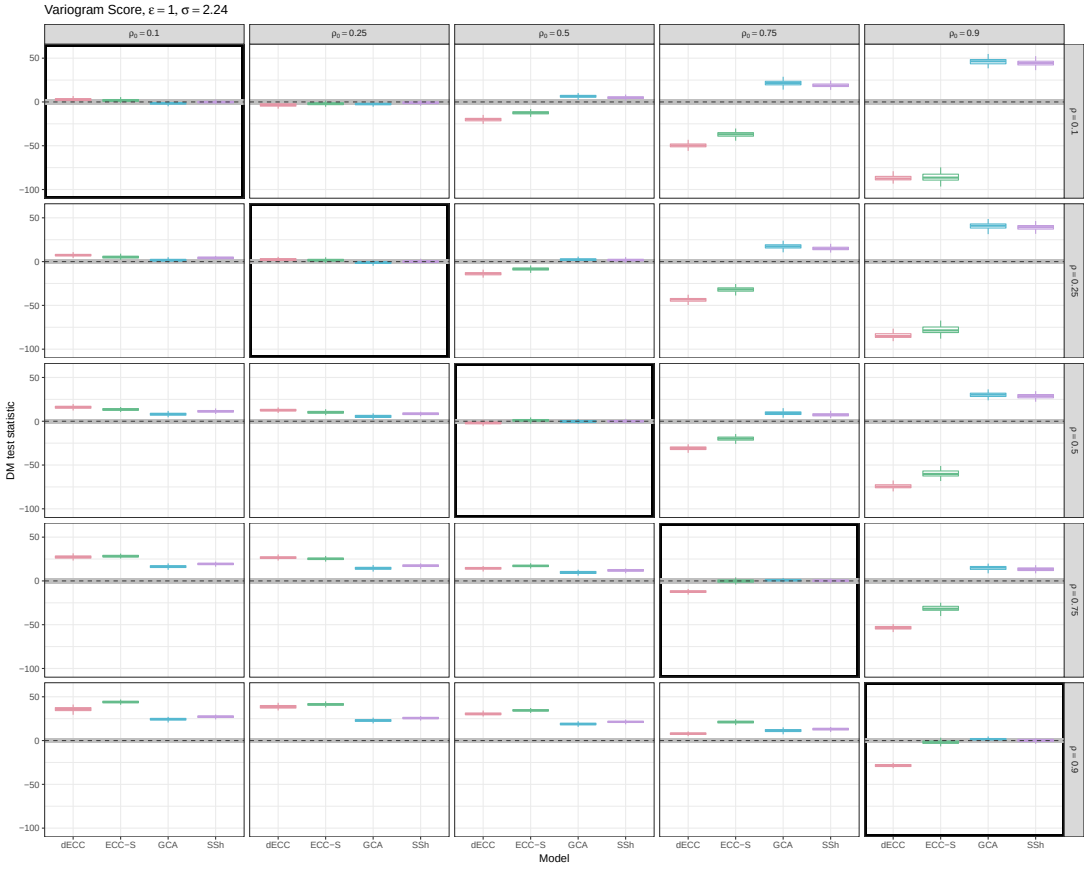


Figure 8: As Figure 1, but for $\epsilon = 1$, and $\sigma = \sqrt{5}$.

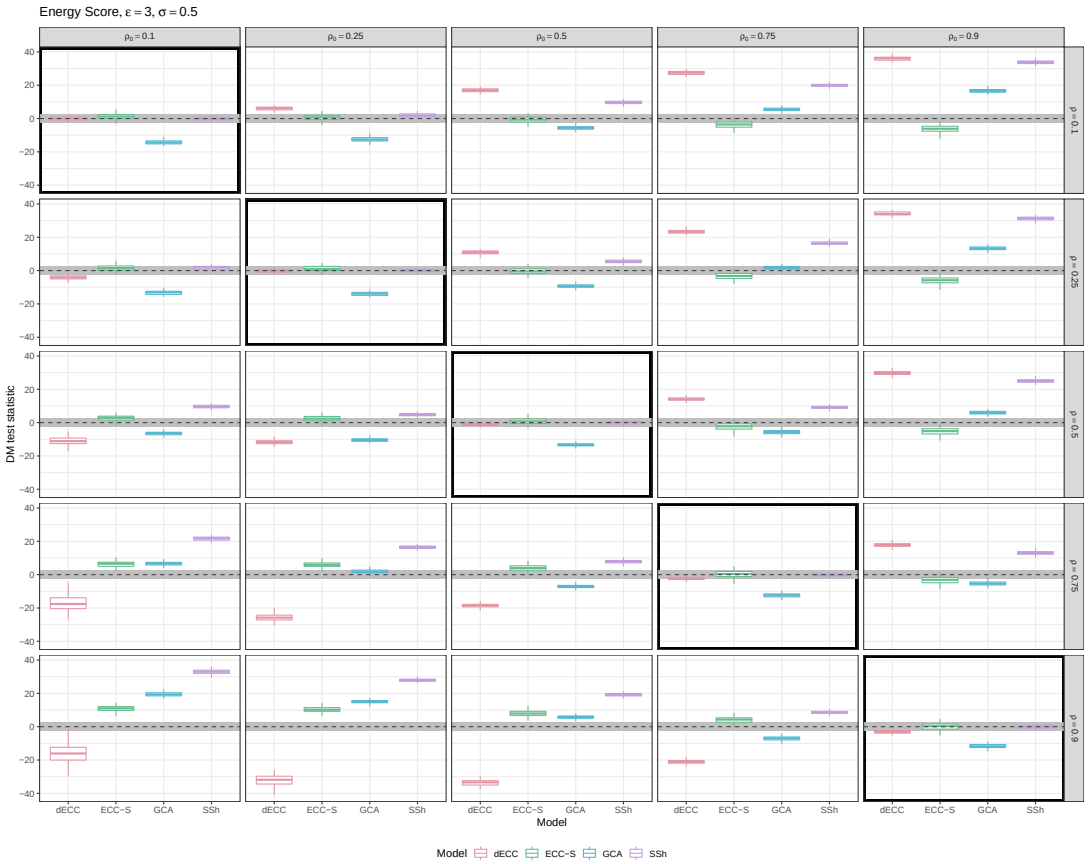
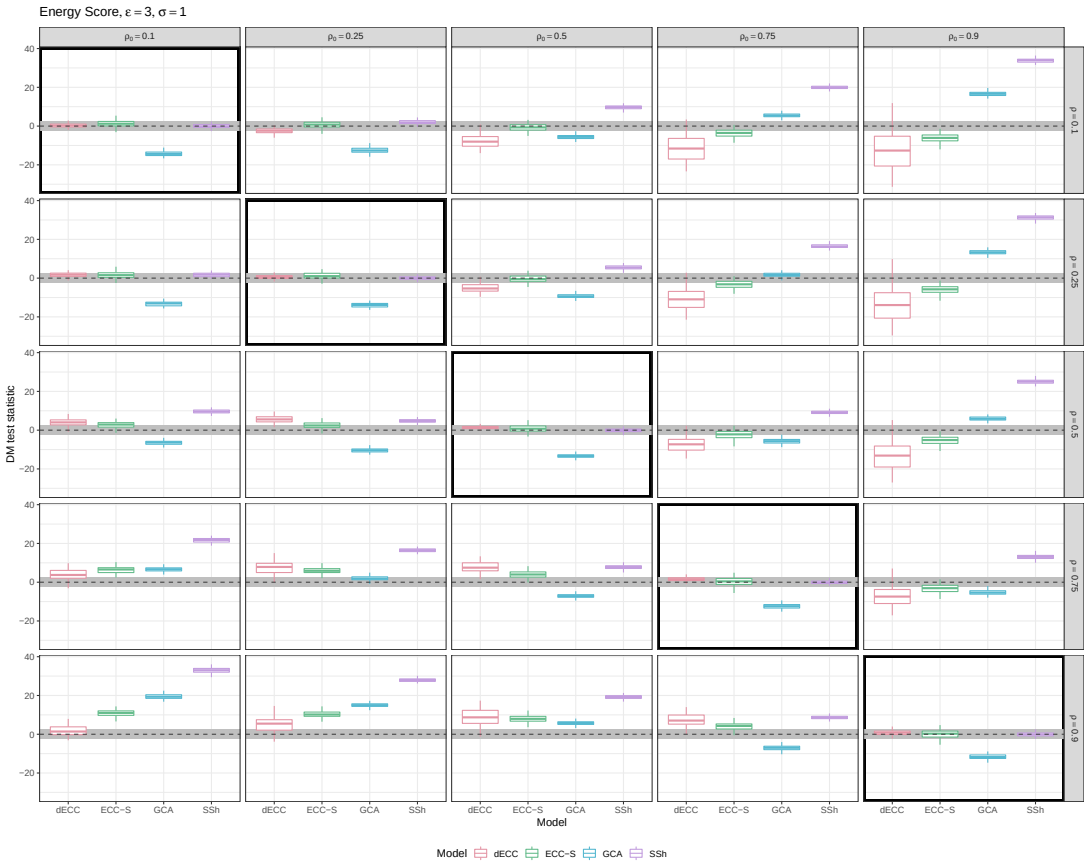


Figure 9: As Figure 1, but for $\epsilon = 3$, and $\sigma = 0.5$.



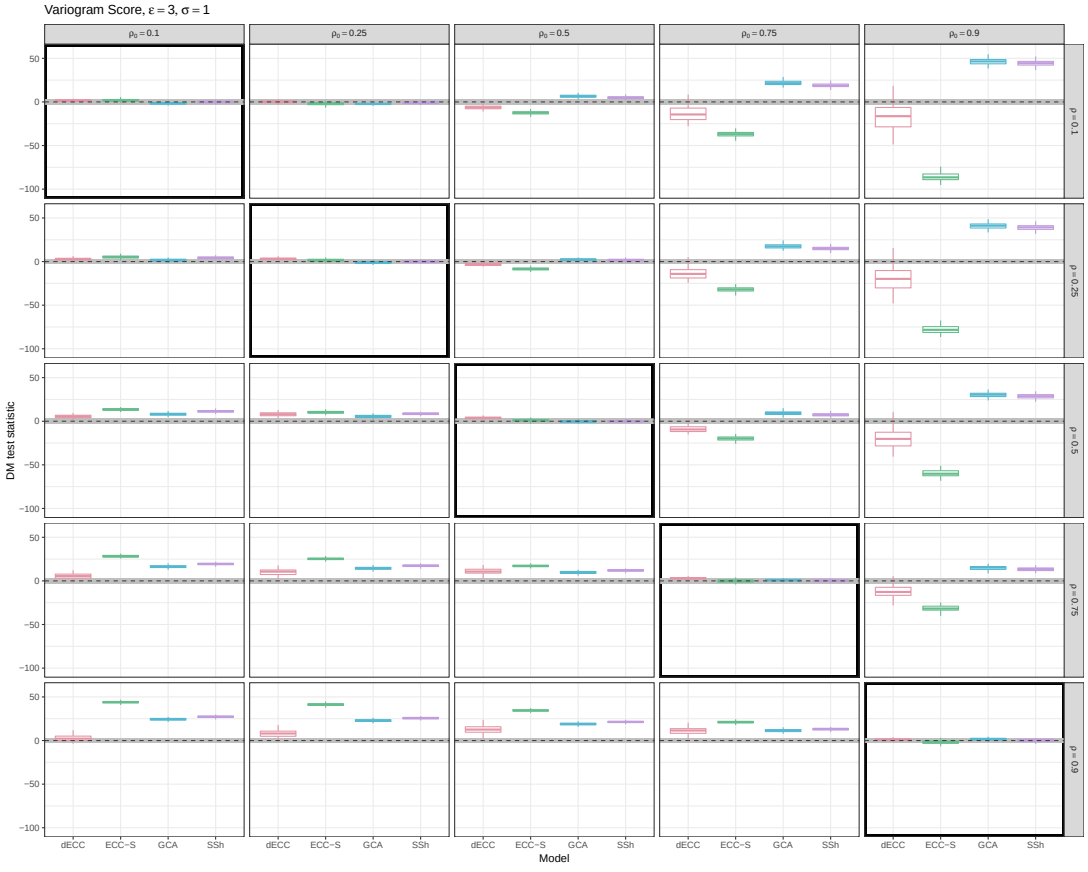
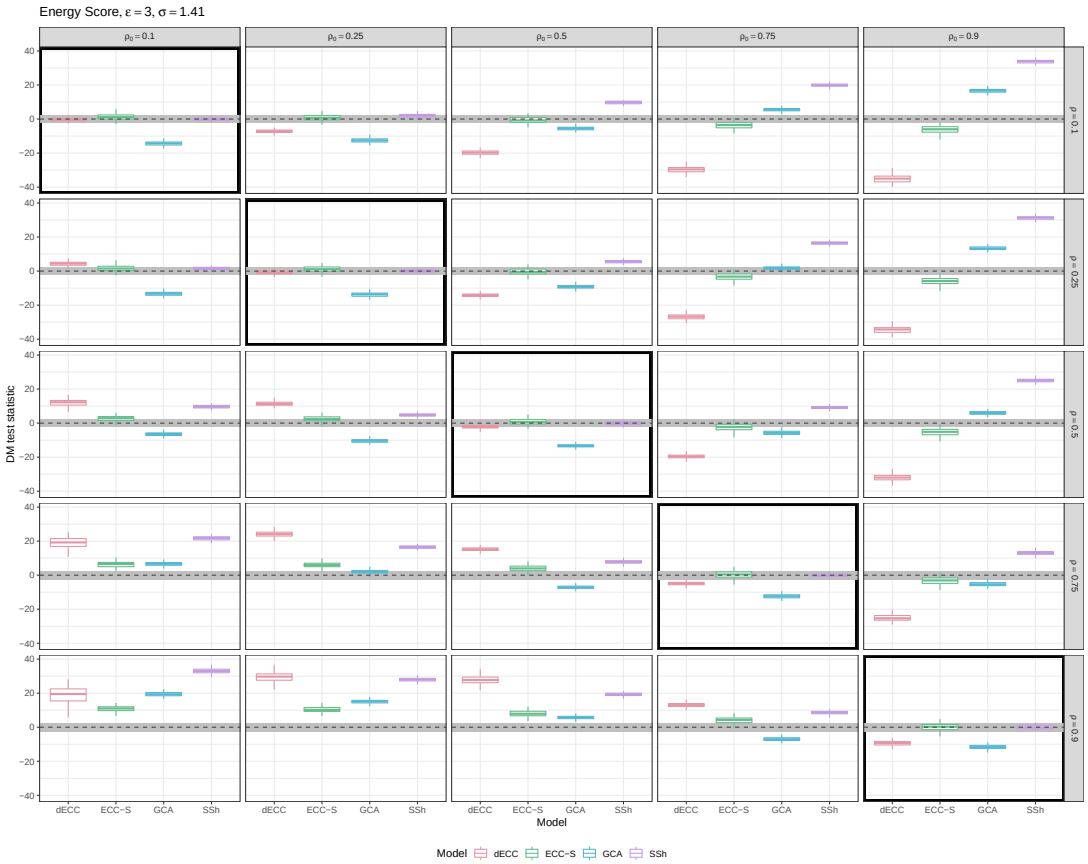


Figure 10: As Figure 1, but for $\epsilon = 3$, and $\sigma = 1$.



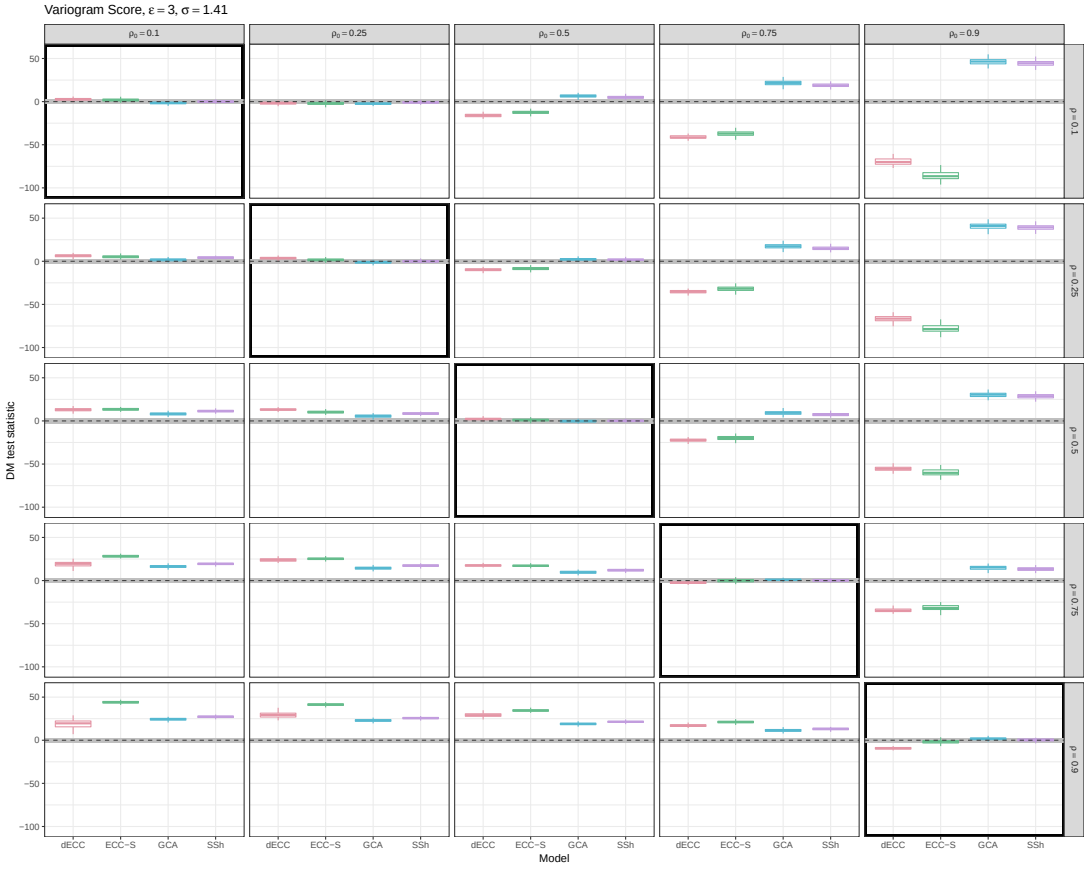
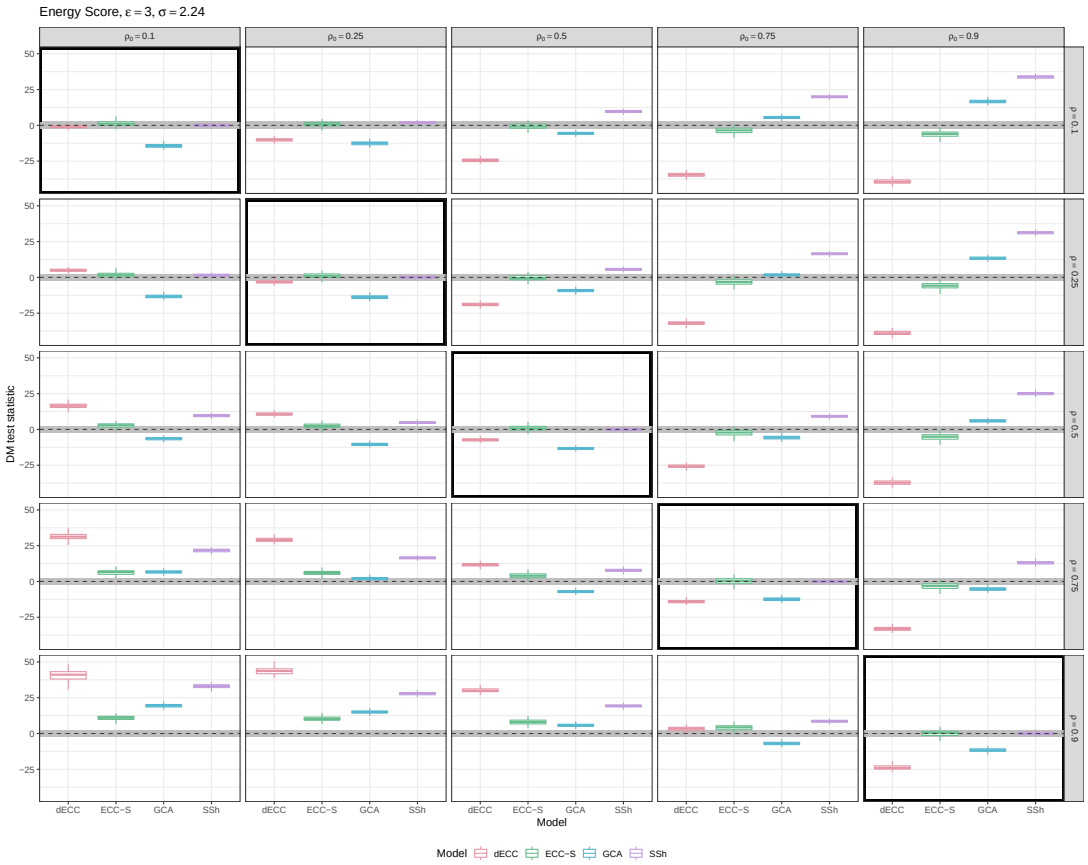


Figure 11: As Figure 1, but for $\epsilon = 3$, and $\sigma = \sqrt{2}$.



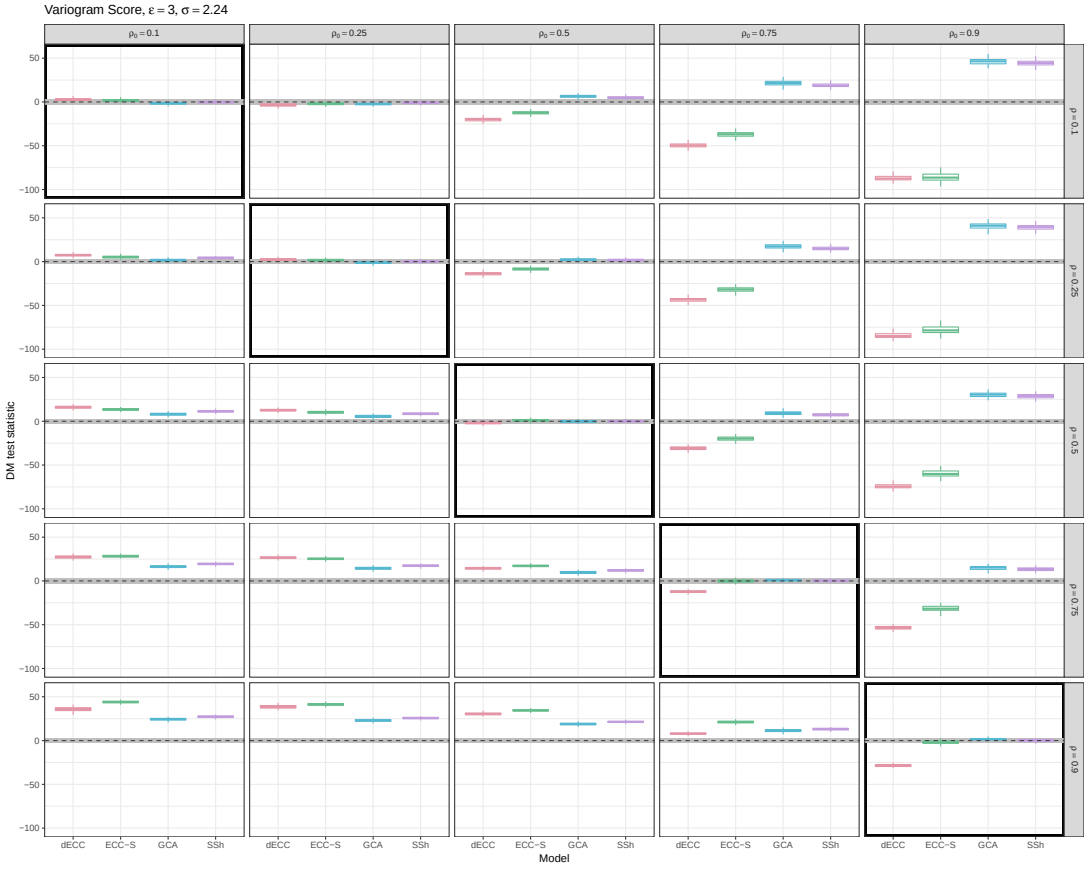


Figure 12: As Figure 1, but for $\epsilon = 3$, and $\sigma = \sqrt{5}$.

2 Additional results for Setting 2

An overview the additional results for Setting 2 is provided in Table 2.

Table 2: Overview of parameter combinations for Setting 2 with references to corresponding figures in this supplement.

μ_0	ϵ	σ	Figure
2	2	0.25	13
2	3	0.5	14
2	3	1	15
2	3	3	16
2	3	5	17
2	3	0.25	18
2	3	0.5	19
2	3	1	20
2	3	3	21
2	3	5	22
2	5	0.25	23
2	5	0.5	24
2	5	1	25
2	5	3	26
2	5	5	27
3	2	0.25	28
3	3	0.5	29
3	3	1	30
3	3	3	31
3	3	5	32
3	3	0.25	33
3	3	0.5	34
3	3	1	35
3	3	3	36
3	3	5	37
3	5	0.25	38
3	5	0.5	39
3	5	1	40
3	5	3	41
3	5	5	42

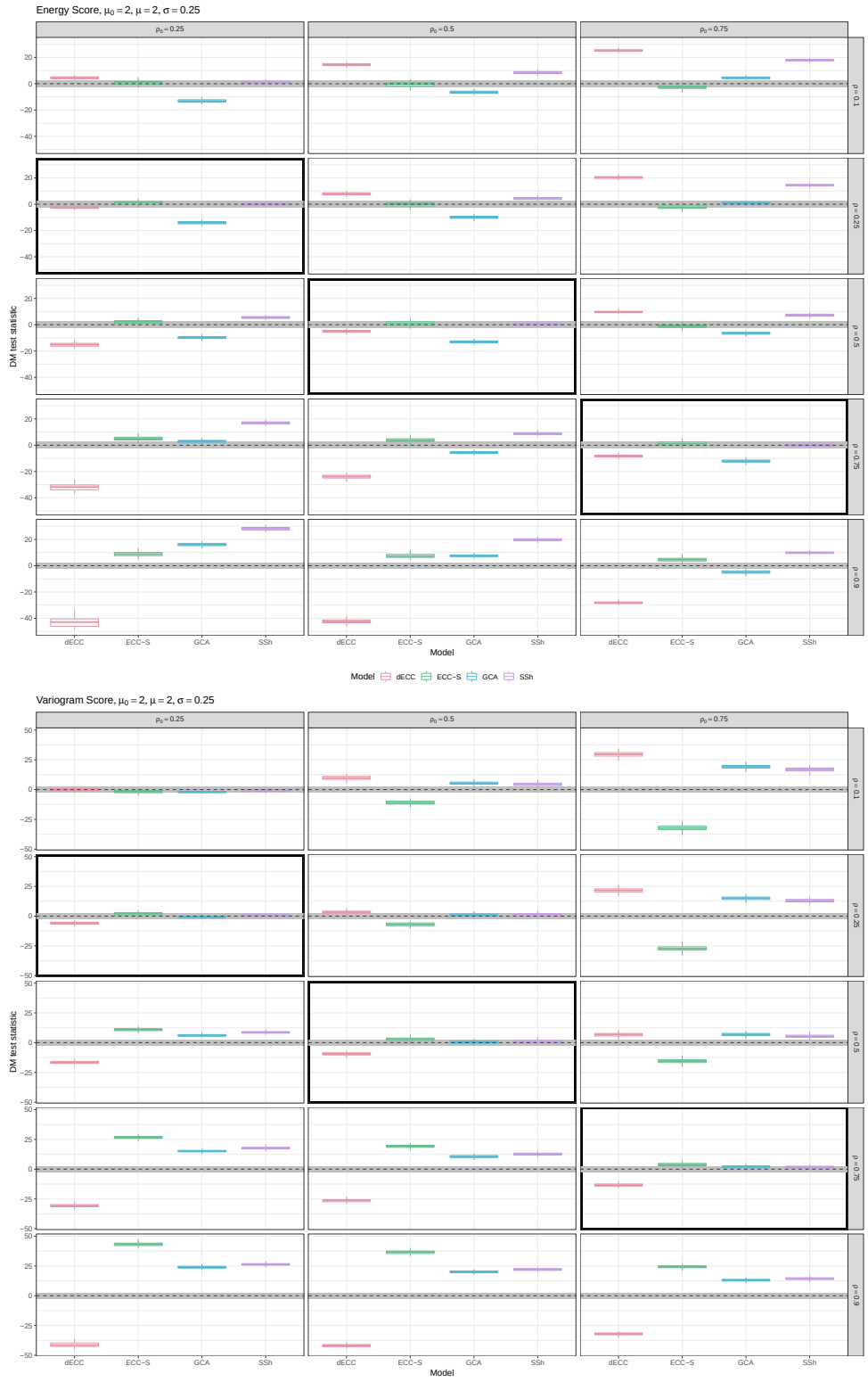


Figure 13: Summaries of DM test statistic values based on ES (top) and VS (bottom) for Setting 2 with $\mu_0 = 2, \epsilon = 2$, and $\sigma = 0.25$. ECC-Q forecasts are used as reference model such that positive values of the test statistic indicate improvements over ECC-Q and negative value indicated deterioration of forecast skill. Boxplots summarize results of the 100 Monte Carlo repetitions of each individual experiment. The horizontal gray stripe indicates the acceptance region of the two-sided DM test under the null hypothesis of equal predictive performance at a level of 0.05. Simulation parameter choices where the correlation structure of the raw ensemble is correctly specified ($\rho = \rho_0$) are surrounded by black boxes.

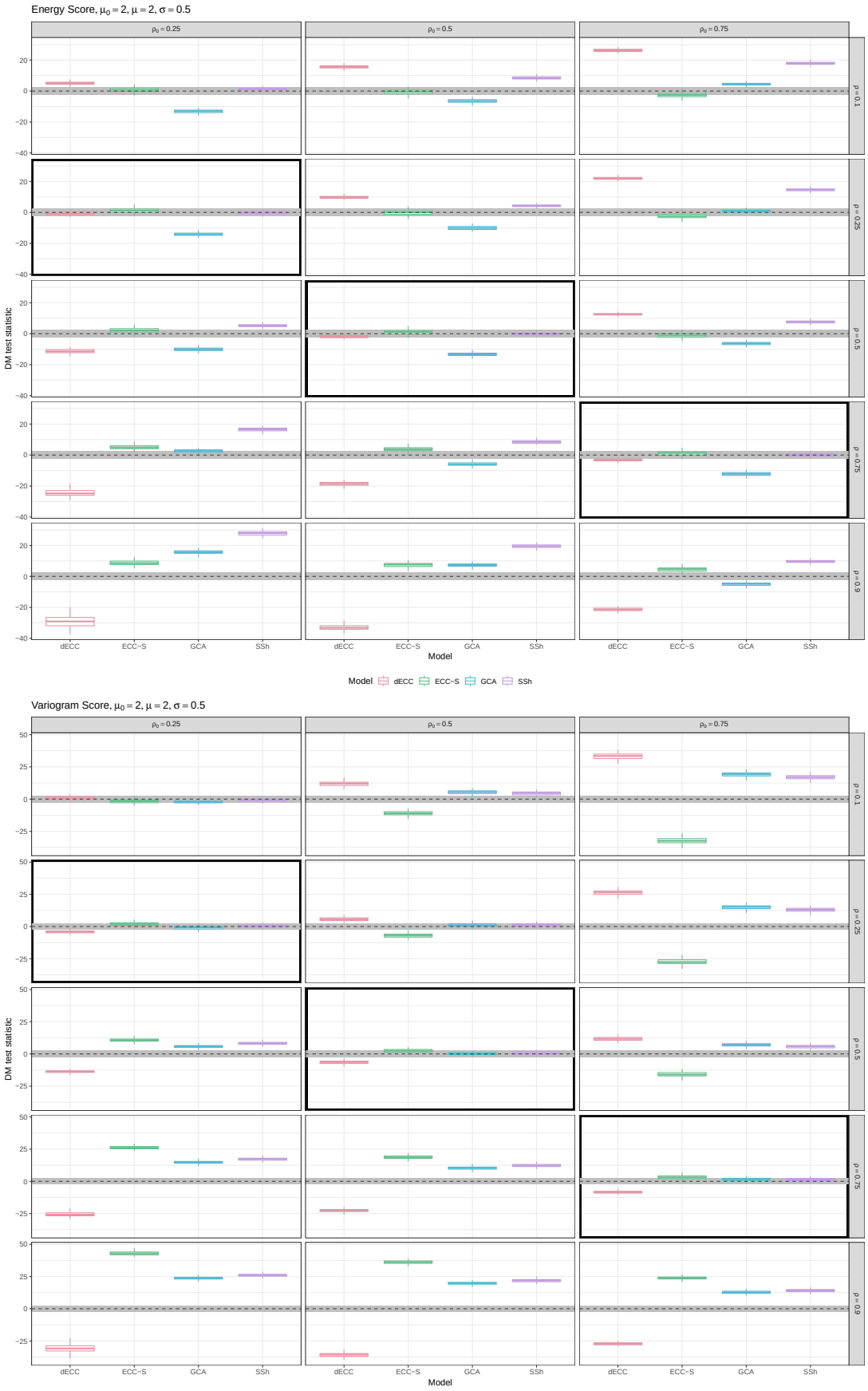


Figure 14: As Figure 13, but for $\mu_0 = 2, \epsilon = 2$, and $\sigma = 0.5$.

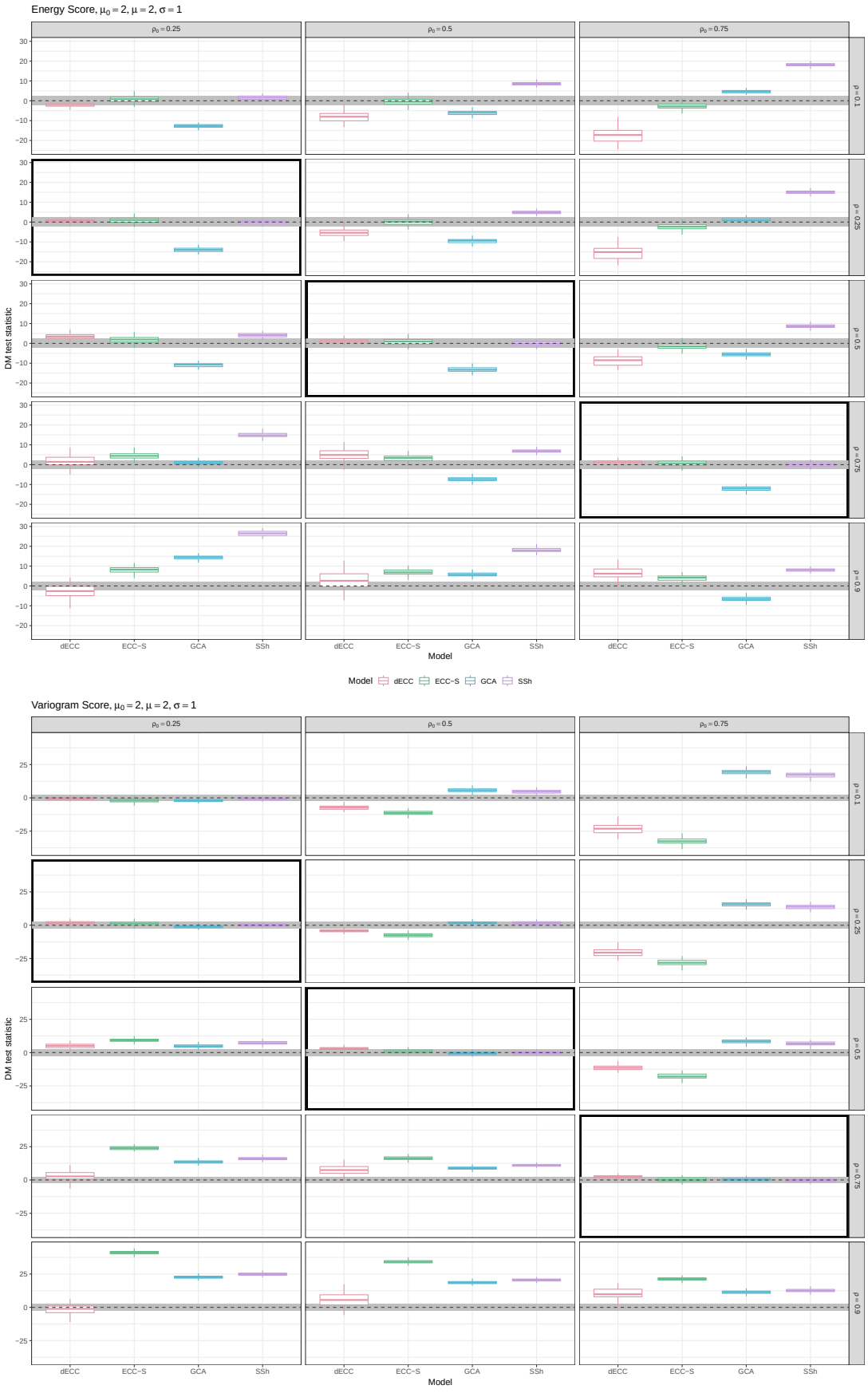


Figure 15: As Figure 13, but for $\mu_0 = 2, \epsilon = 2$, and $\sigma = 1$.

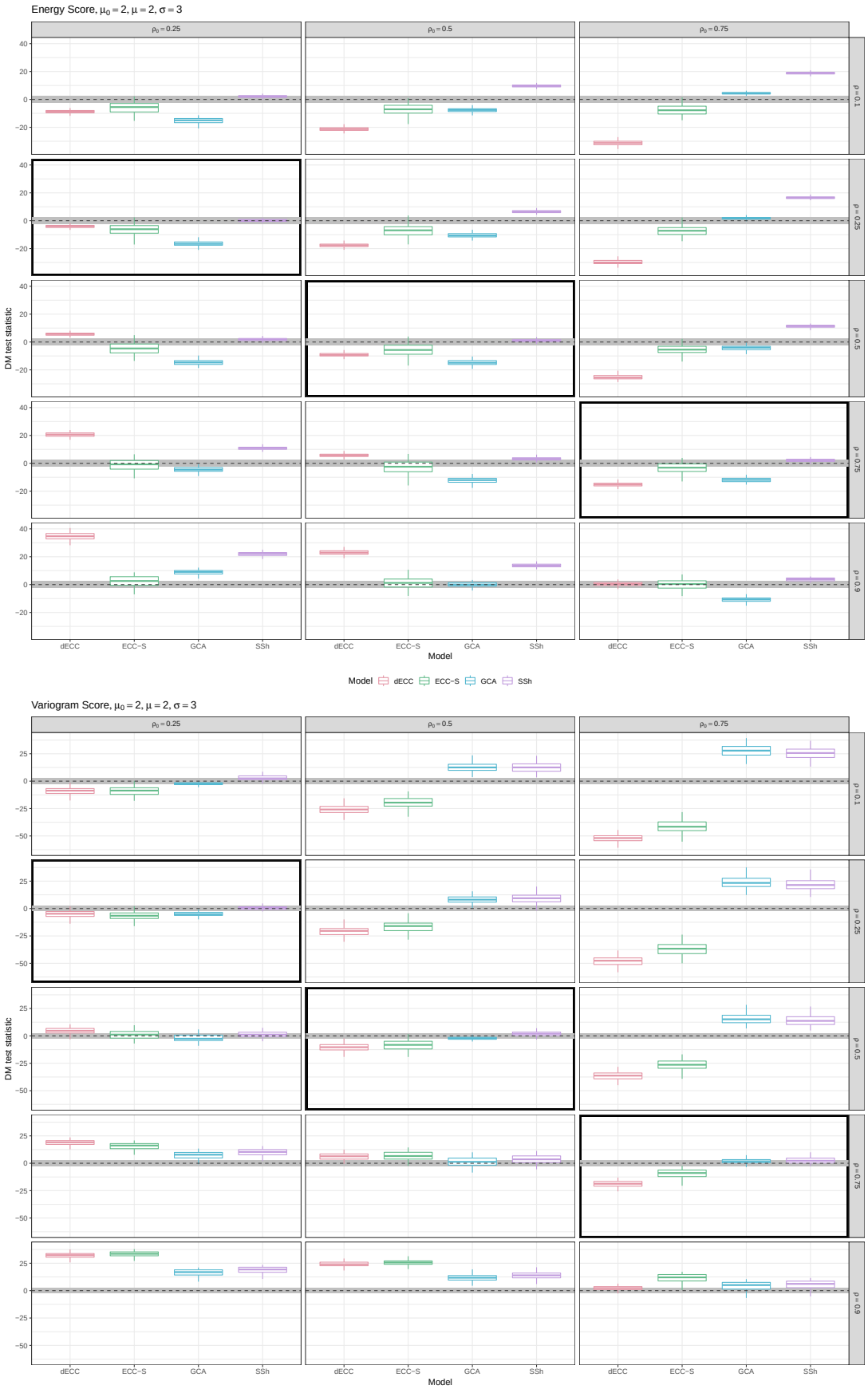


Figure 16: As Figure 13, but for $\mu_0 = 2, \epsilon = 2$, and $\sigma = 3$.

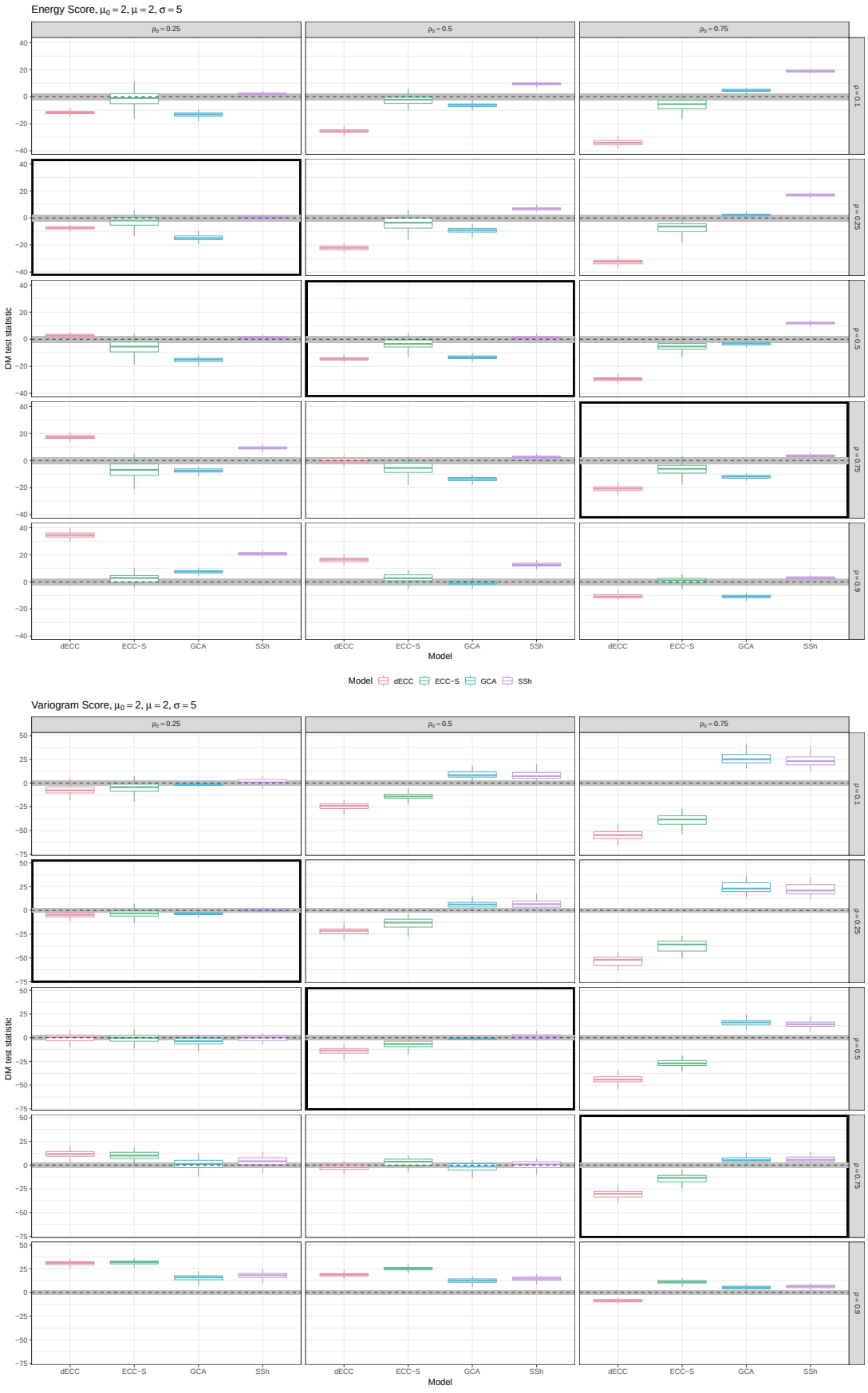


Figure 17: As Figure 13, but for $\mu_0 = 2, \epsilon = 2$, and $\sigma = 5$.

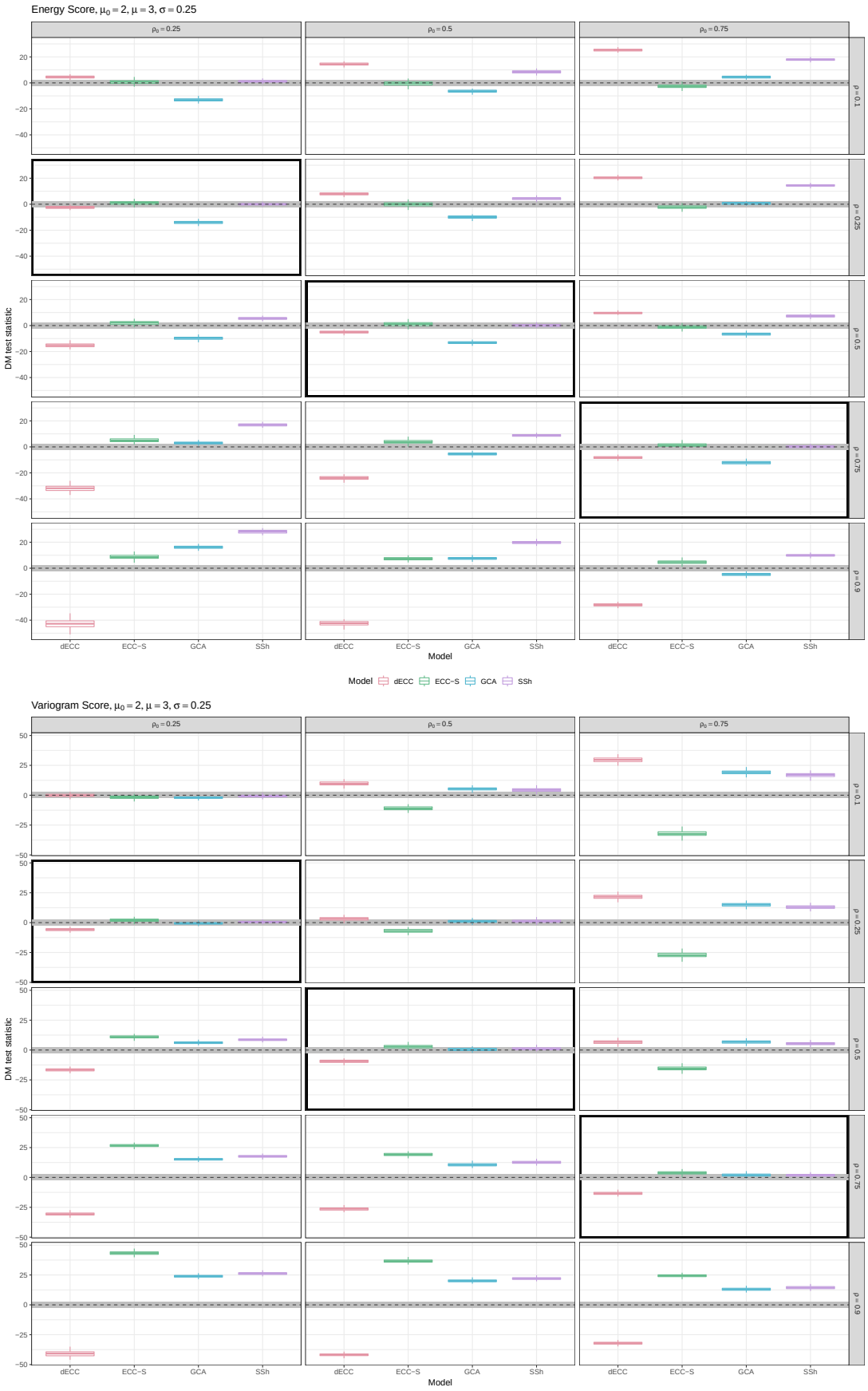


Figure 18: As Figure 13, but for $\mu_0 = 2, \epsilon = 3$, and $\sigma = 0.5$.

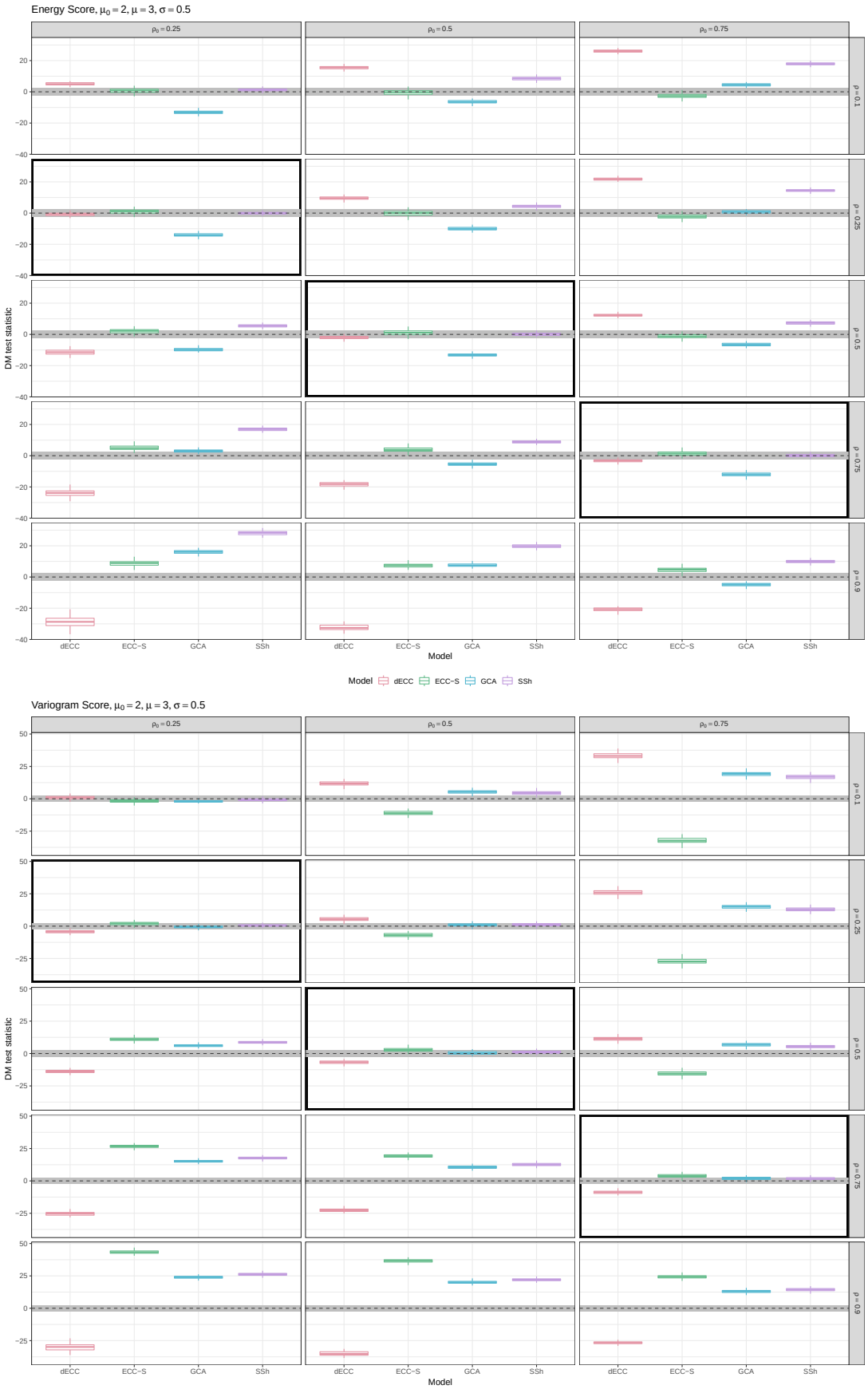


Figure 19: As Figure 13, but for $\mu_0 = 2, \epsilon = 3$, and $\sigma = 0.5$.

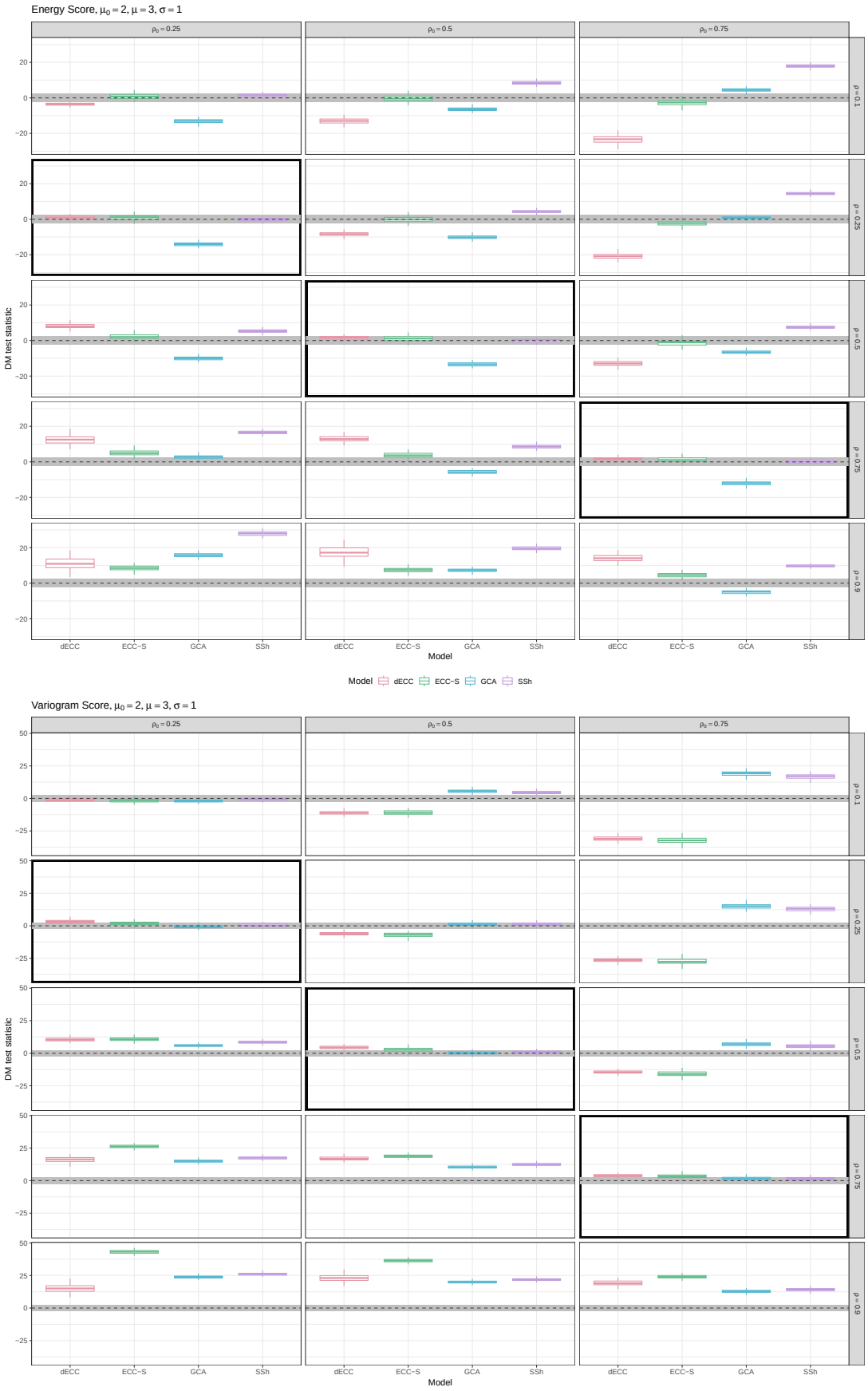


Figure 20: As Figure 13, but for $\mu_0 = 2, \epsilon = 3$, and $\sigma = 1$.

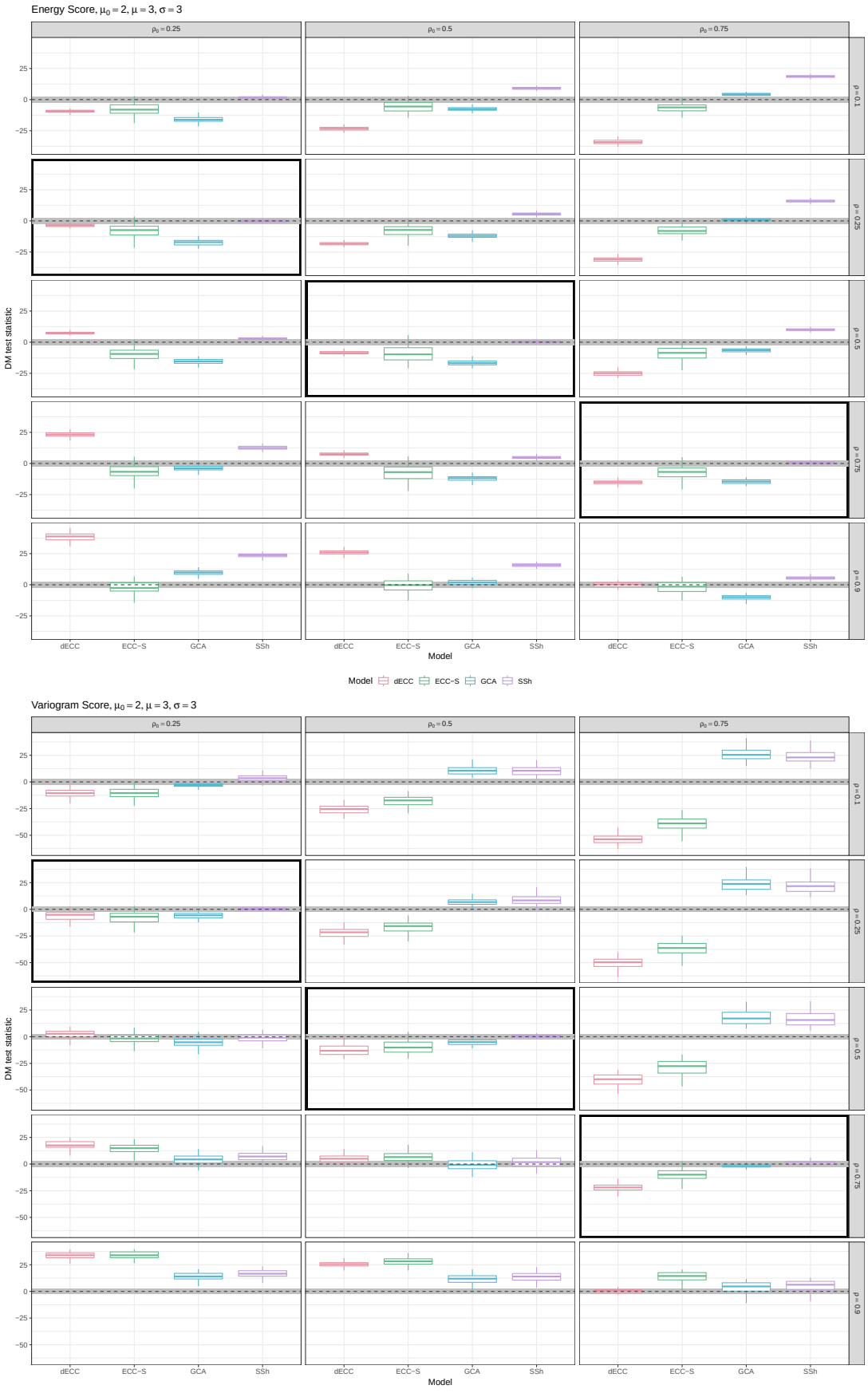


Figure 21: As Figure 13, but for $\mu_0 = 2, \epsilon = 3$, and $\sigma = 3$.

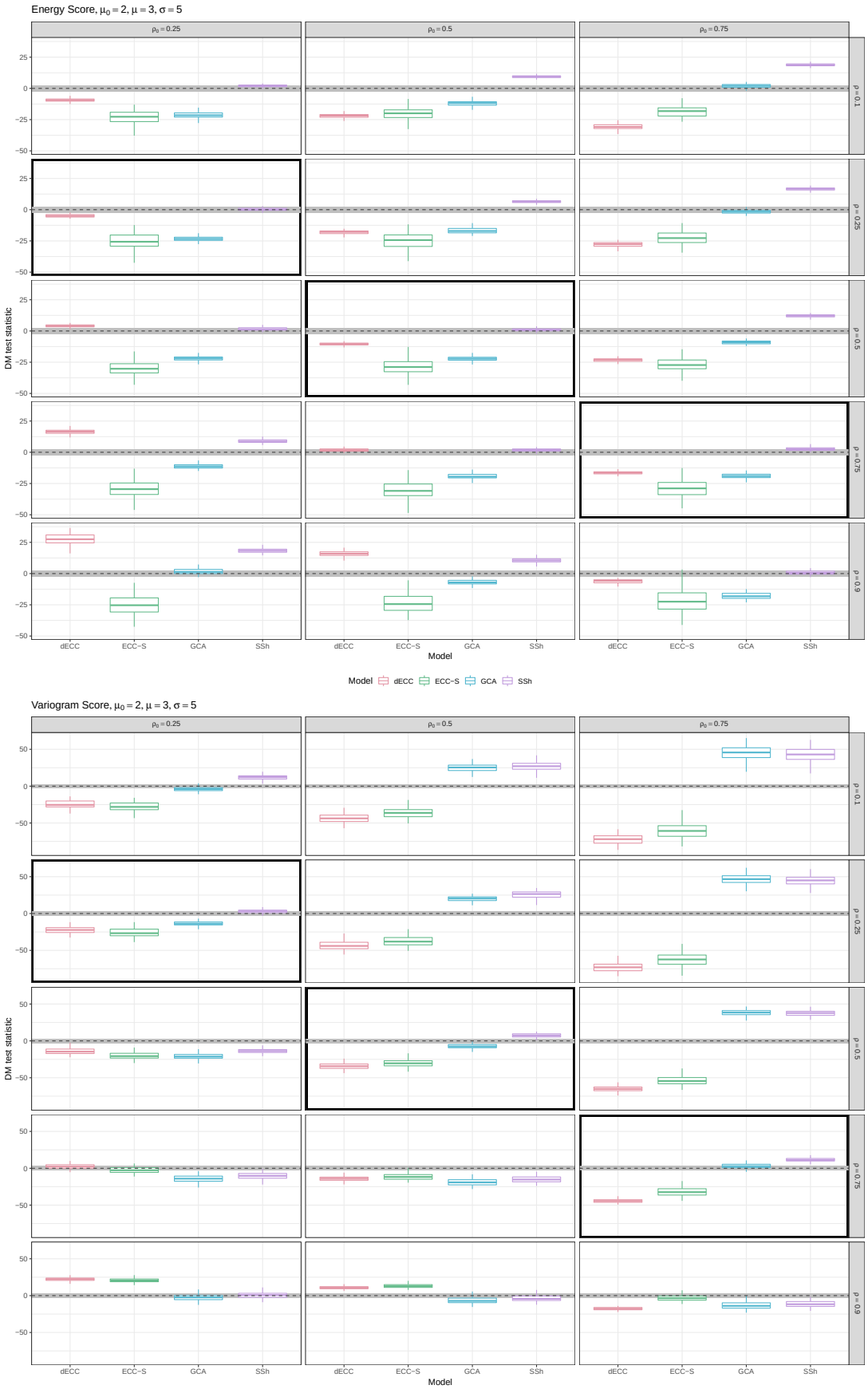
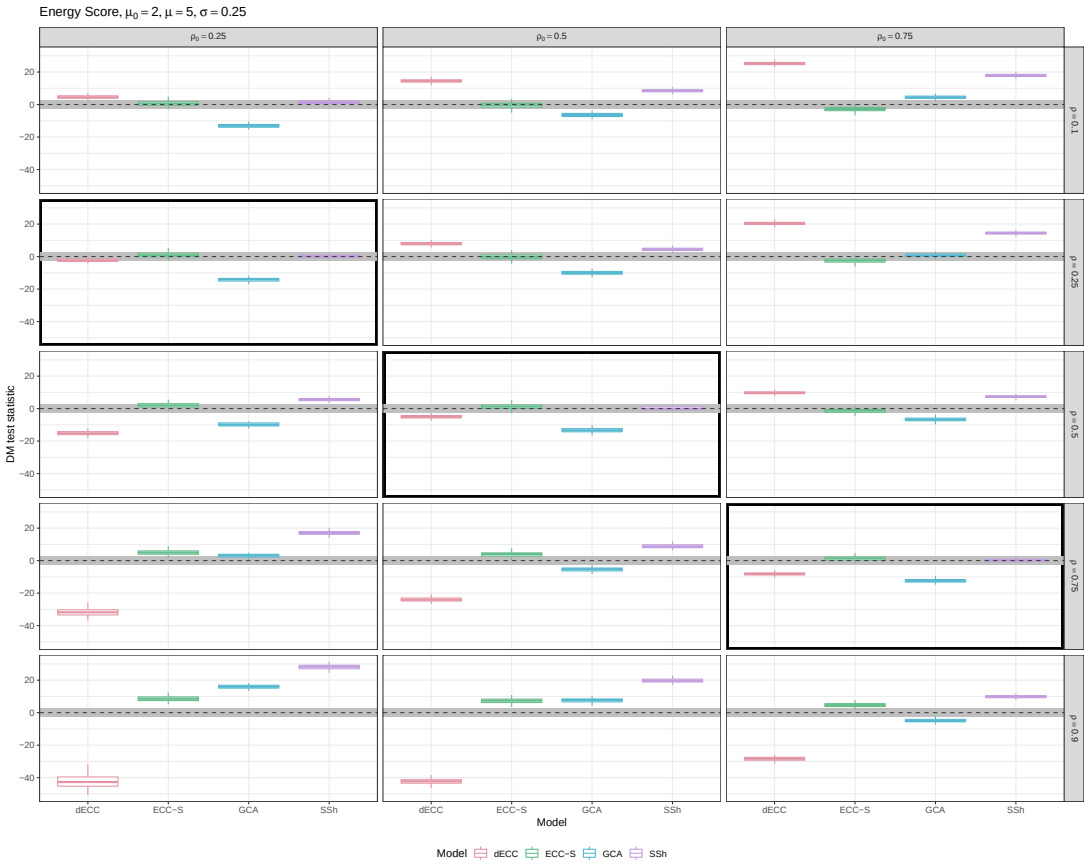


Figure 22: As Figure 13, but for $\mu_0 = 2, \epsilon = 3$, and $\sigma = 5$.



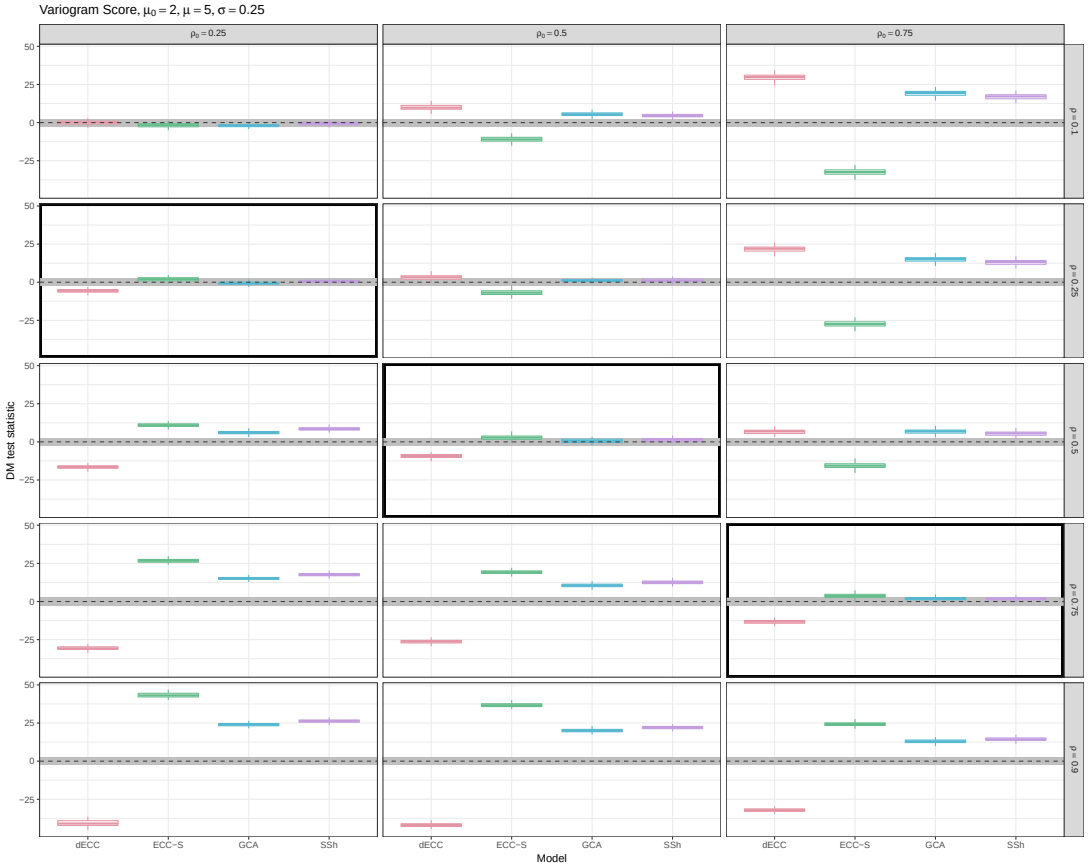


Figure 23: As Figure 13, but for $\mu_0 = 2, \epsilon = 5$, and $\sigma = 0.5$.

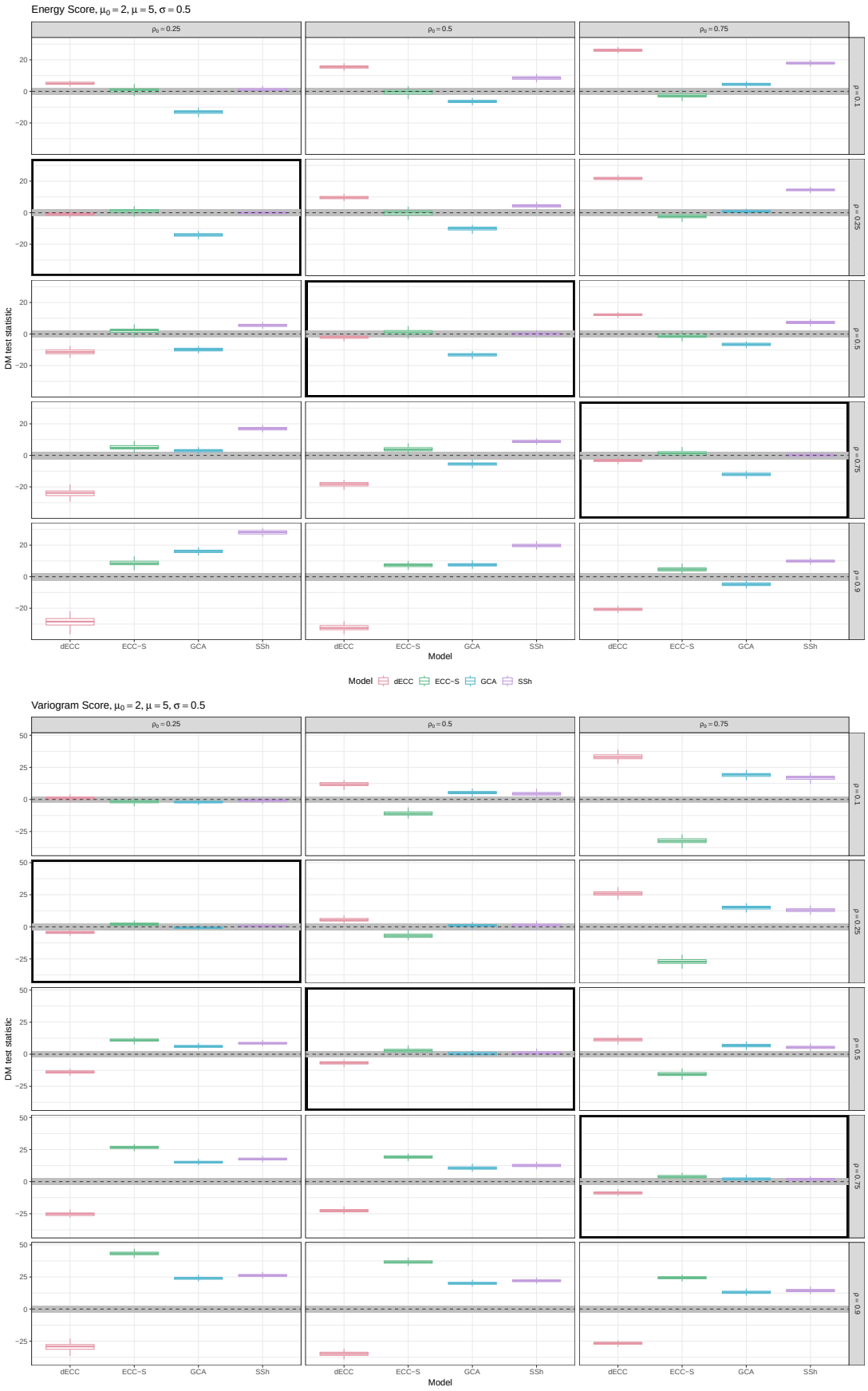


Figure 24: As Figure 13, but for $\mu_0 = 2, \epsilon = 5$, and $\sigma = 0.5$.

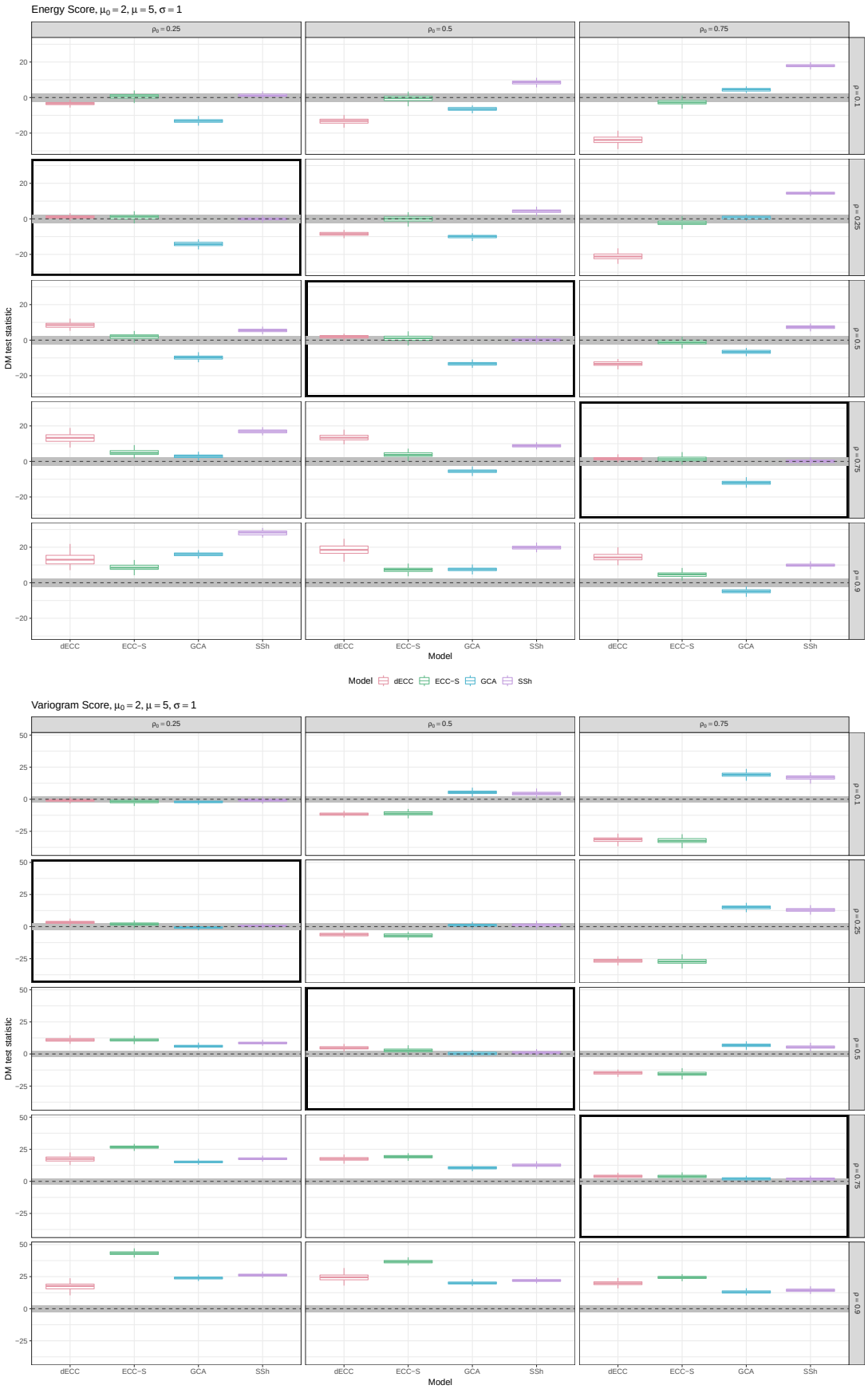


Figure 25: As Figure 13, but for $\mu_0 = 2, \epsilon = 5$, and $\sigma = 1$.

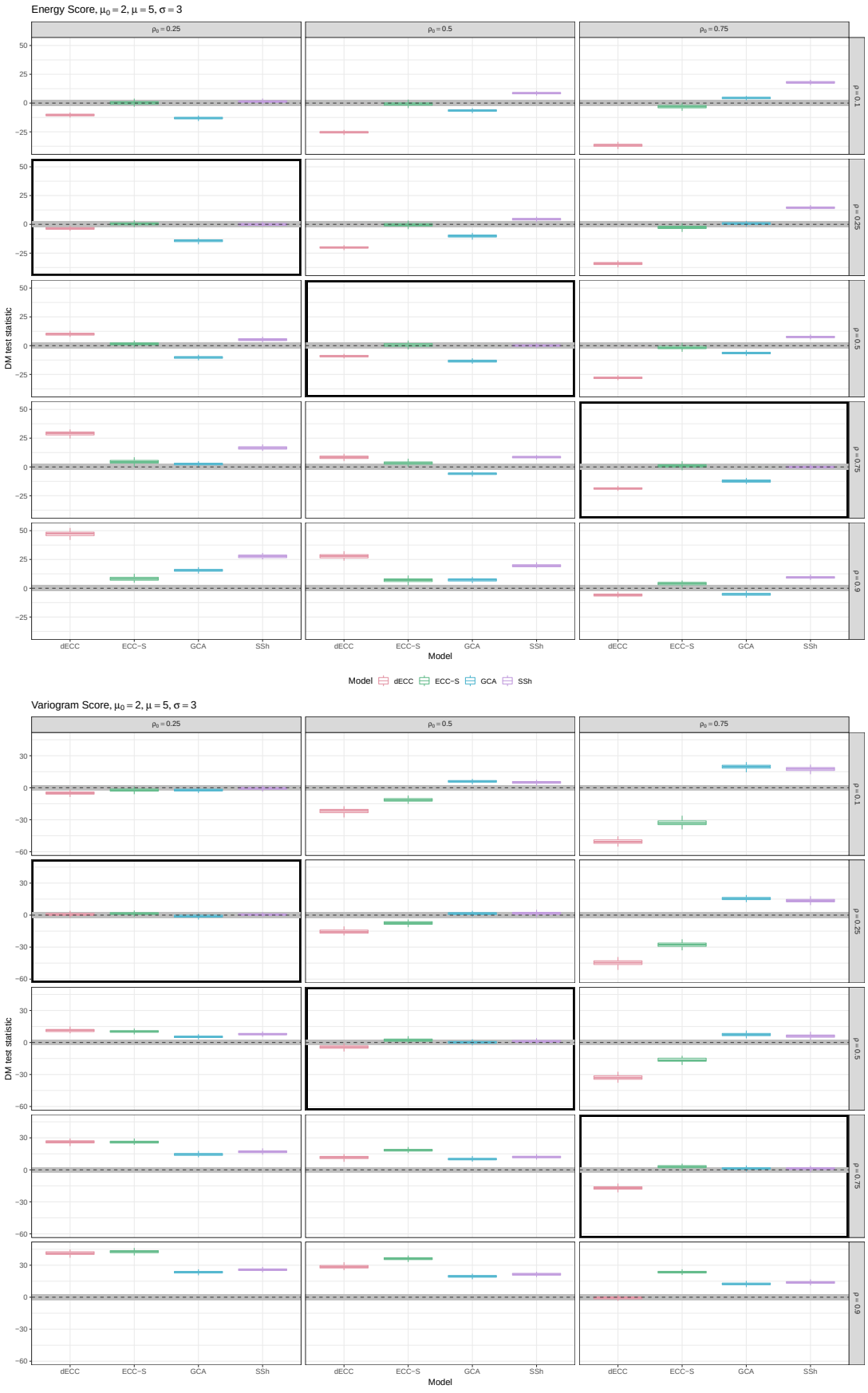


Figure 26: As Figure 13, but for $\mu_0 = 2, \epsilon = 5$, and $\sigma = 3$.

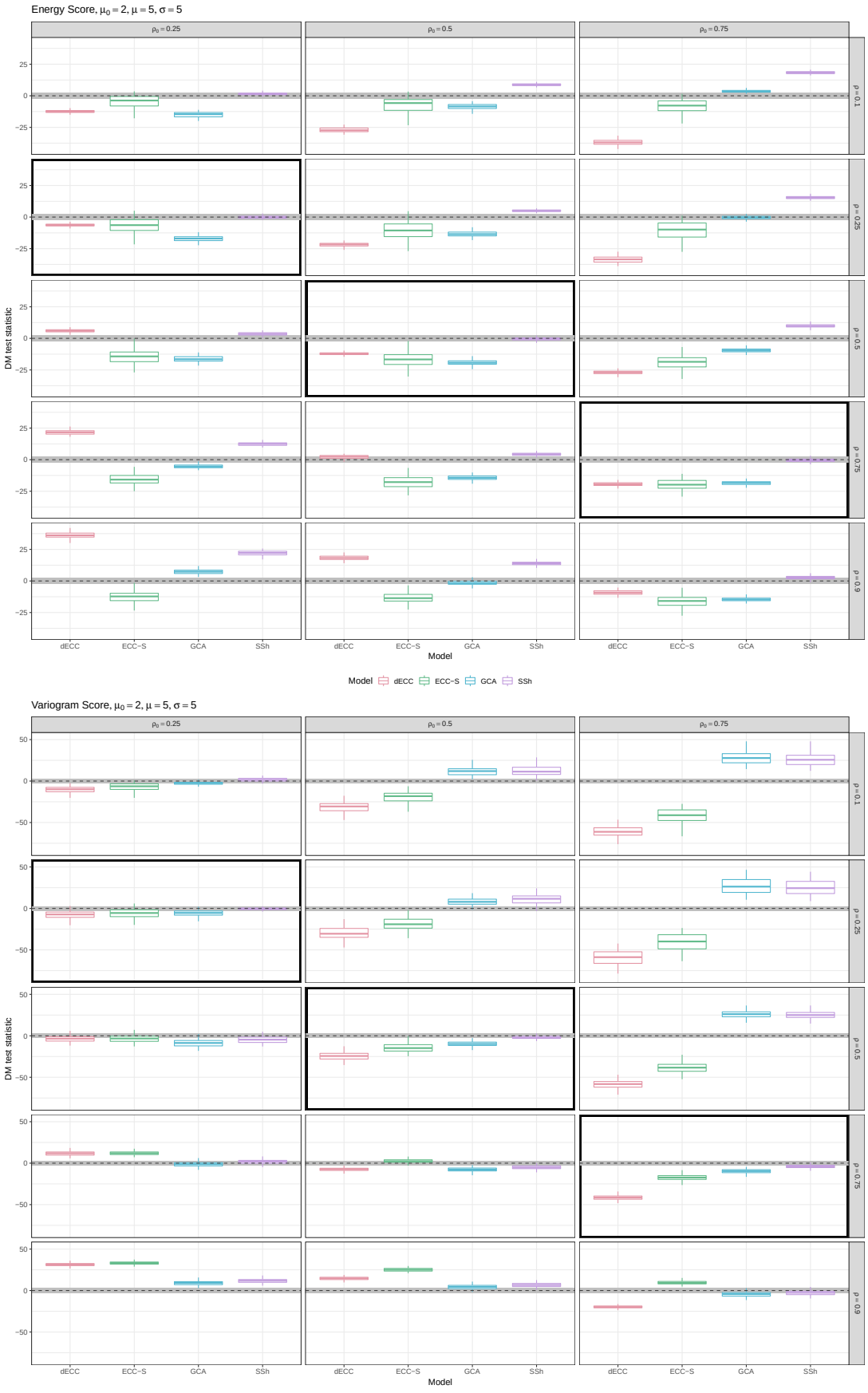


Figure 27: As Figure 13, but for $\mu_0 = 2, \epsilon = 5$, and $\sigma = 5$.

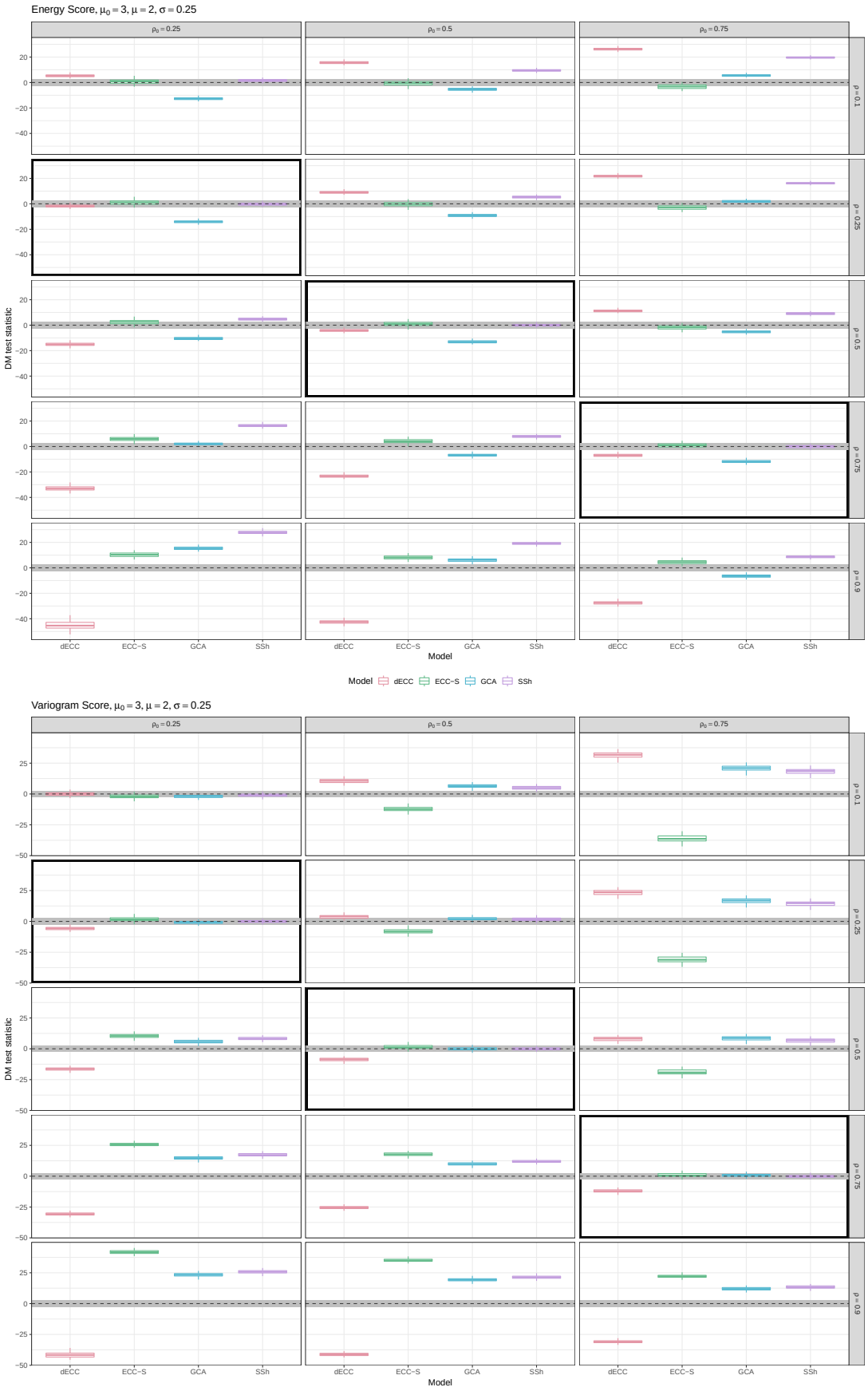


Figure 28: As Figure 13, but for $\mu_0 = 3, \epsilon = 2$, and $\sigma = 0.5$.

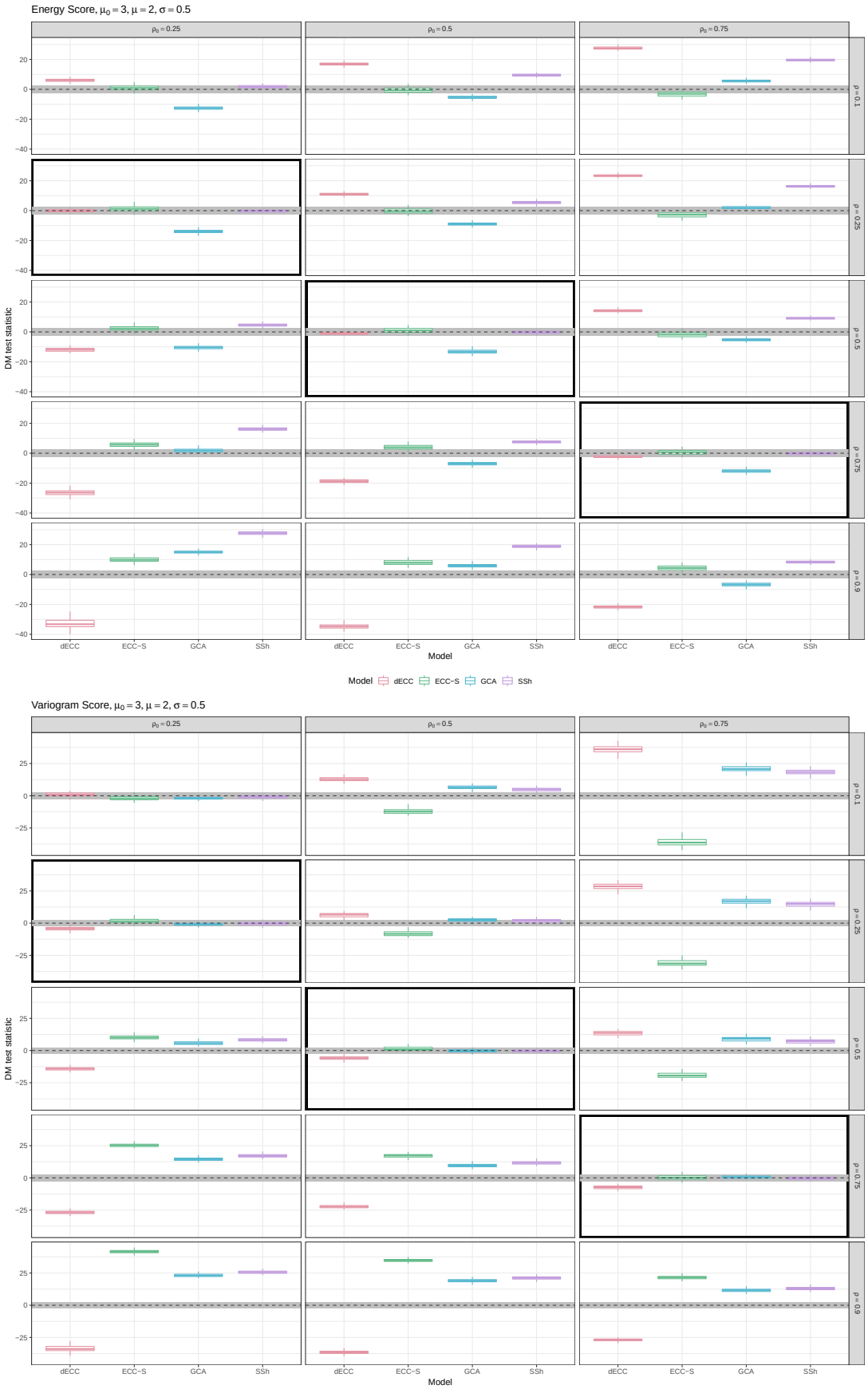


Figure 29: As Figure 13, but for $\mu_0 = 3, \epsilon = 2$, and $\sigma = 0.5$.

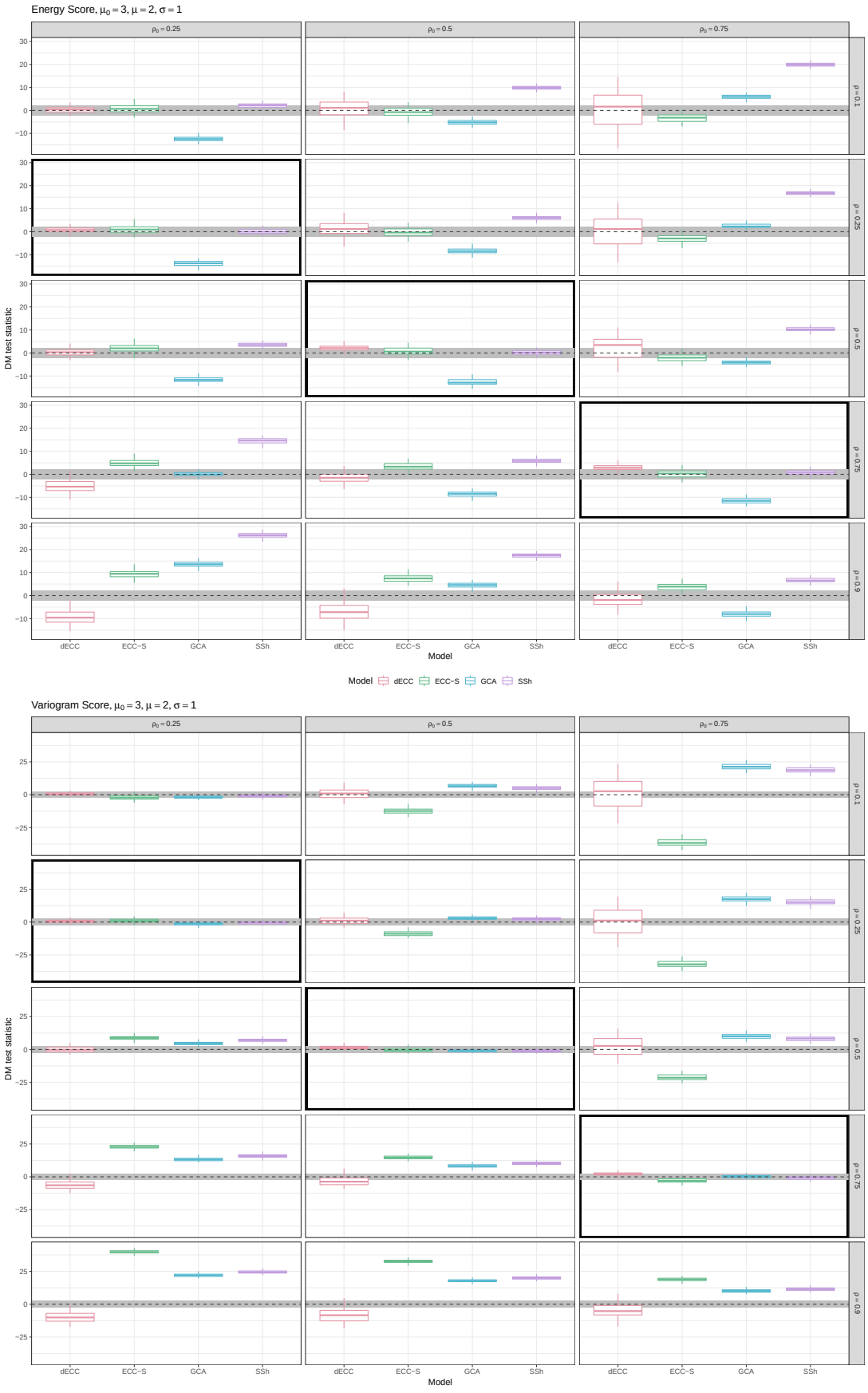


Figure 30: As Figure 13, but for $\mu_0 = 3, \epsilon = 2$, and $\sigma = 1$.

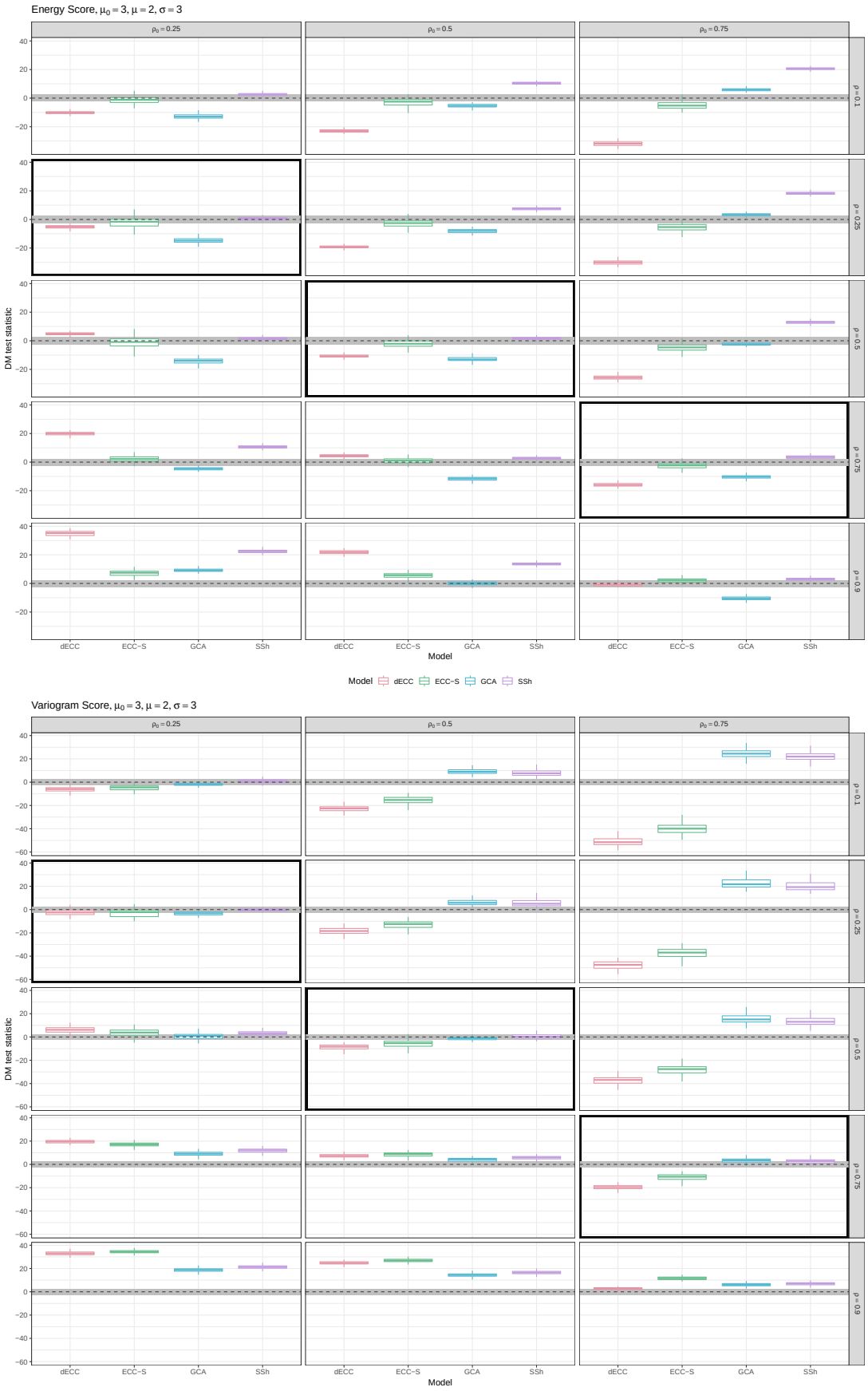


Figure 31: As Figure 13, but for $\mu_0 = 3, \epsilon = 2$, and $\sigma = 3$.

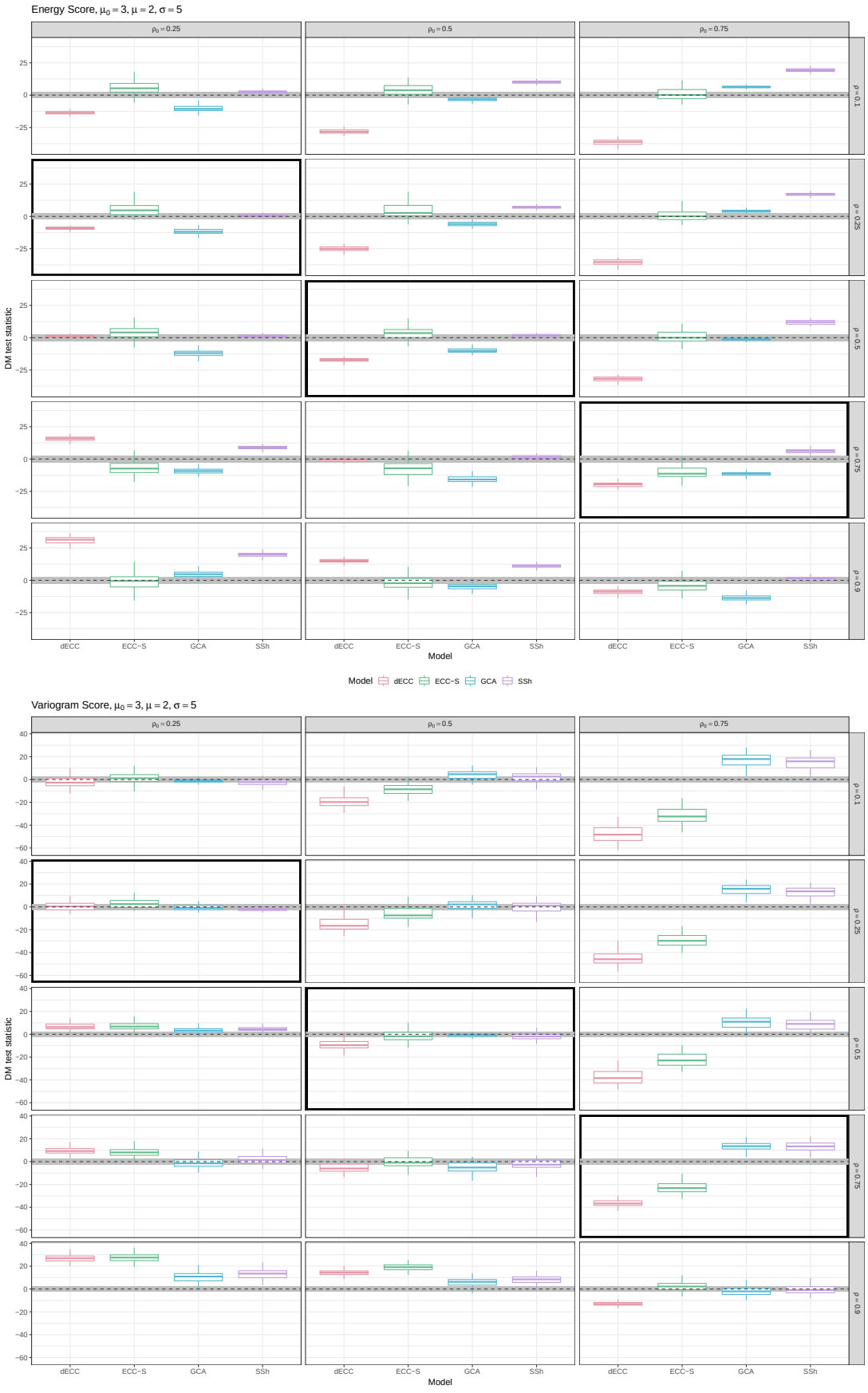


Figure 32: As Figure 13, but for $\mu_0 = 3, \epsilon = 2$, and $\sigma = 5$.

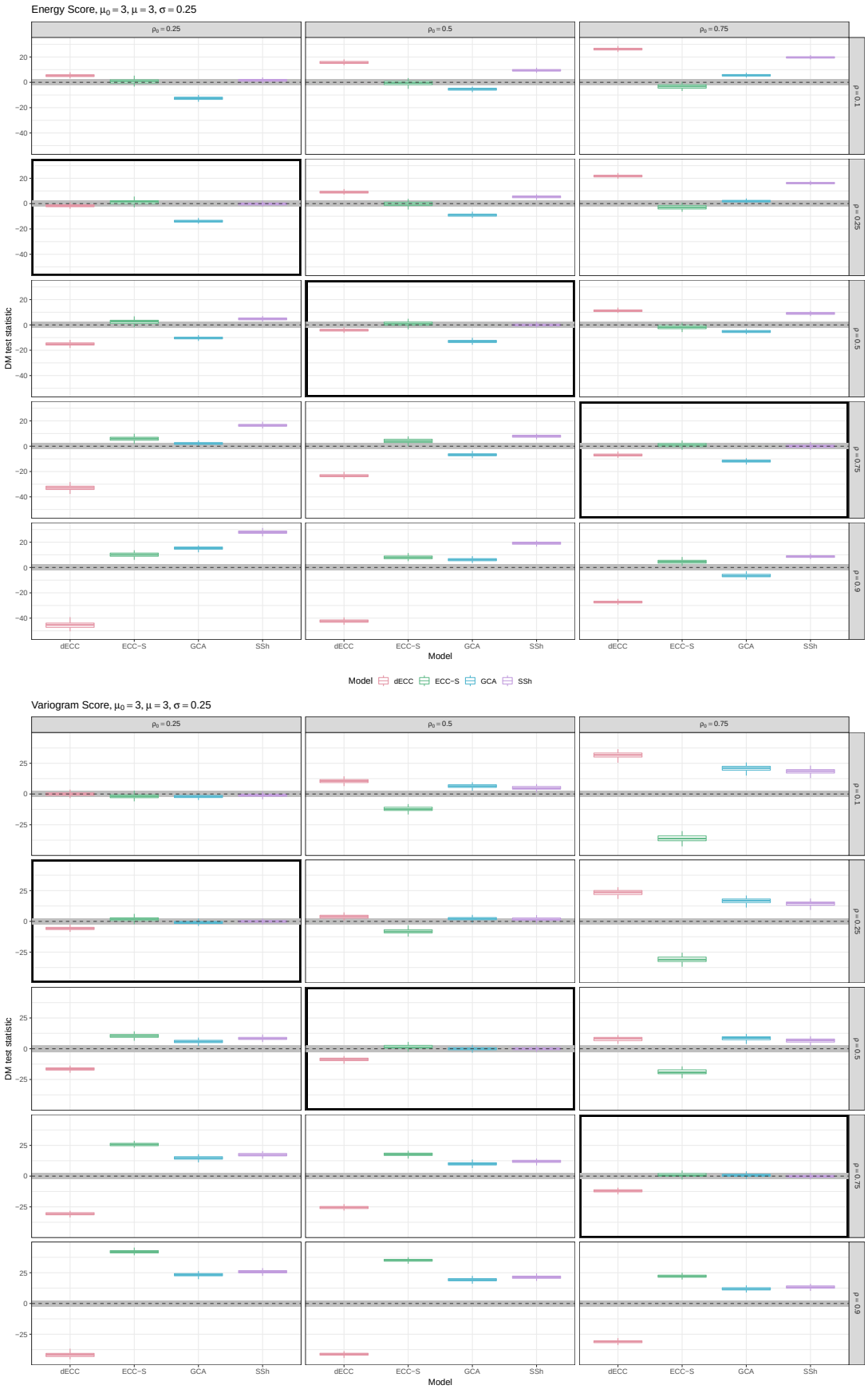


Figure 33: As Figure 13, but for $\mu_0 = 3, \epsilon = 3$, and $\sigma = 0.5$.

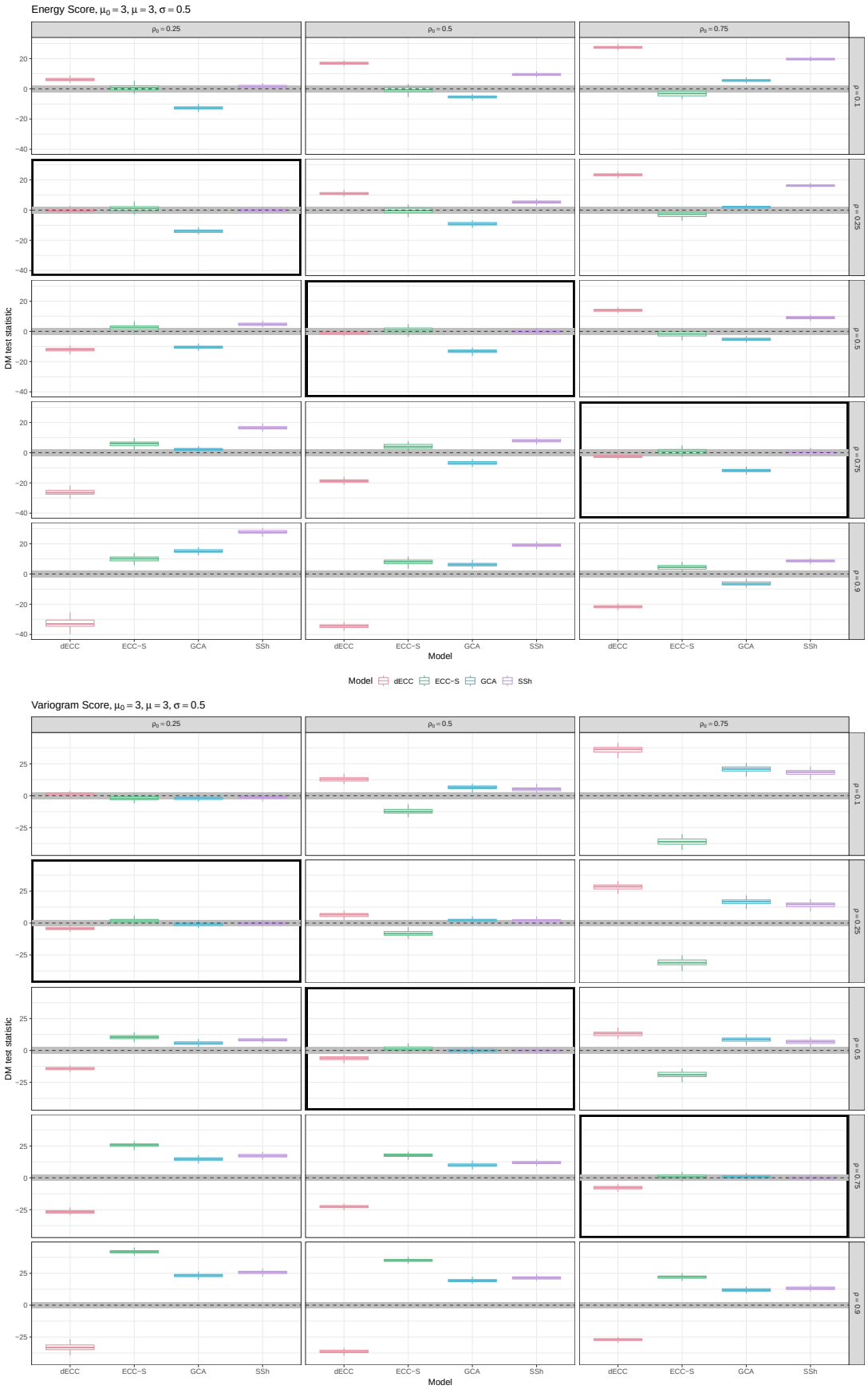


Figure 34: As Figure 13, but for $\mu_0 = 3, \epsilon = 3$, and $\sigma = 0.5$.

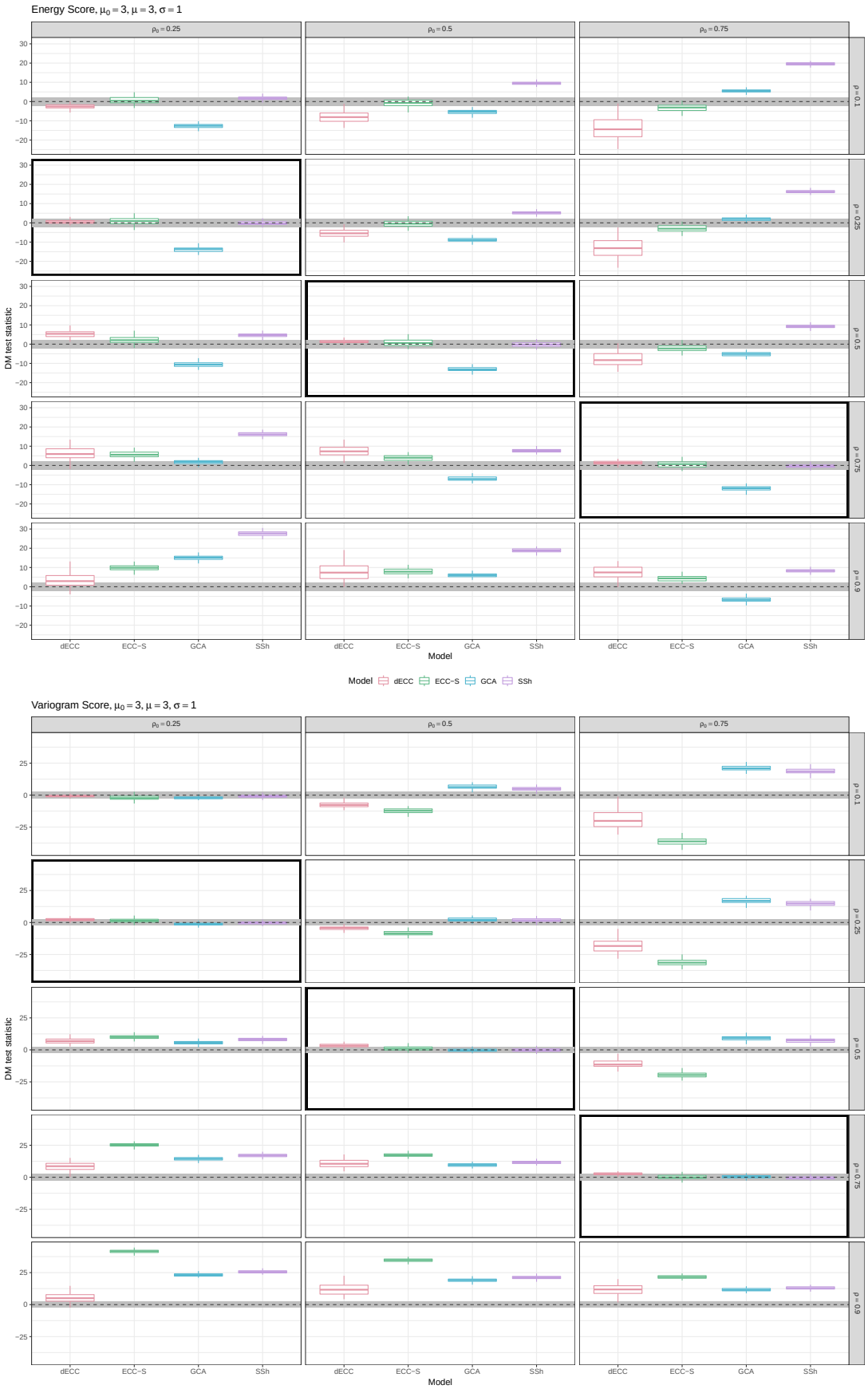


Figure 35: As Figure 13, but for $\mu_0 = 3, \epsilon = 3$, and $\sigma = 1$.

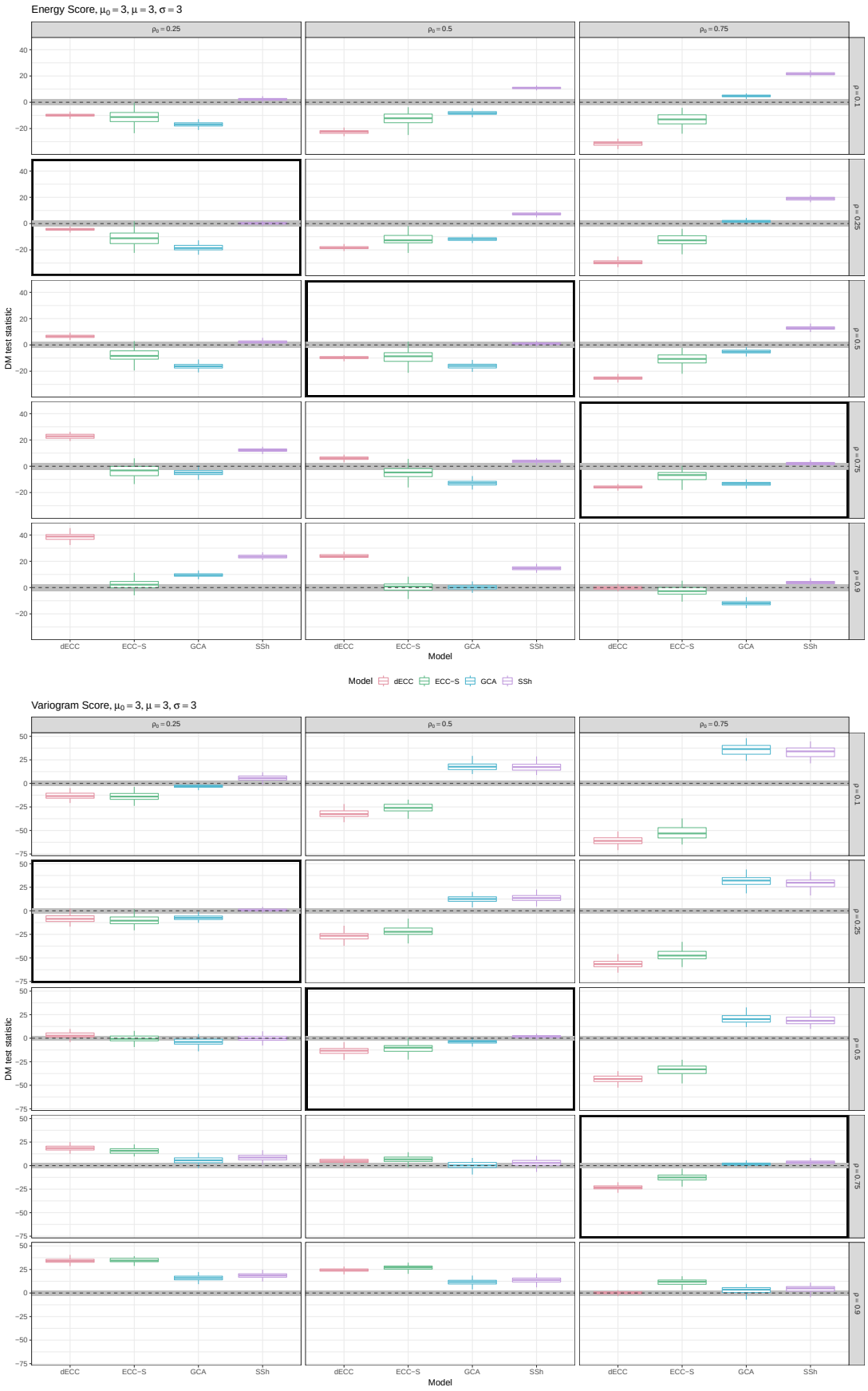


Figure 36: As Figure 13, but for $\mu_0 = 3, \epsilon = 3$, and $\sigma = 3$.

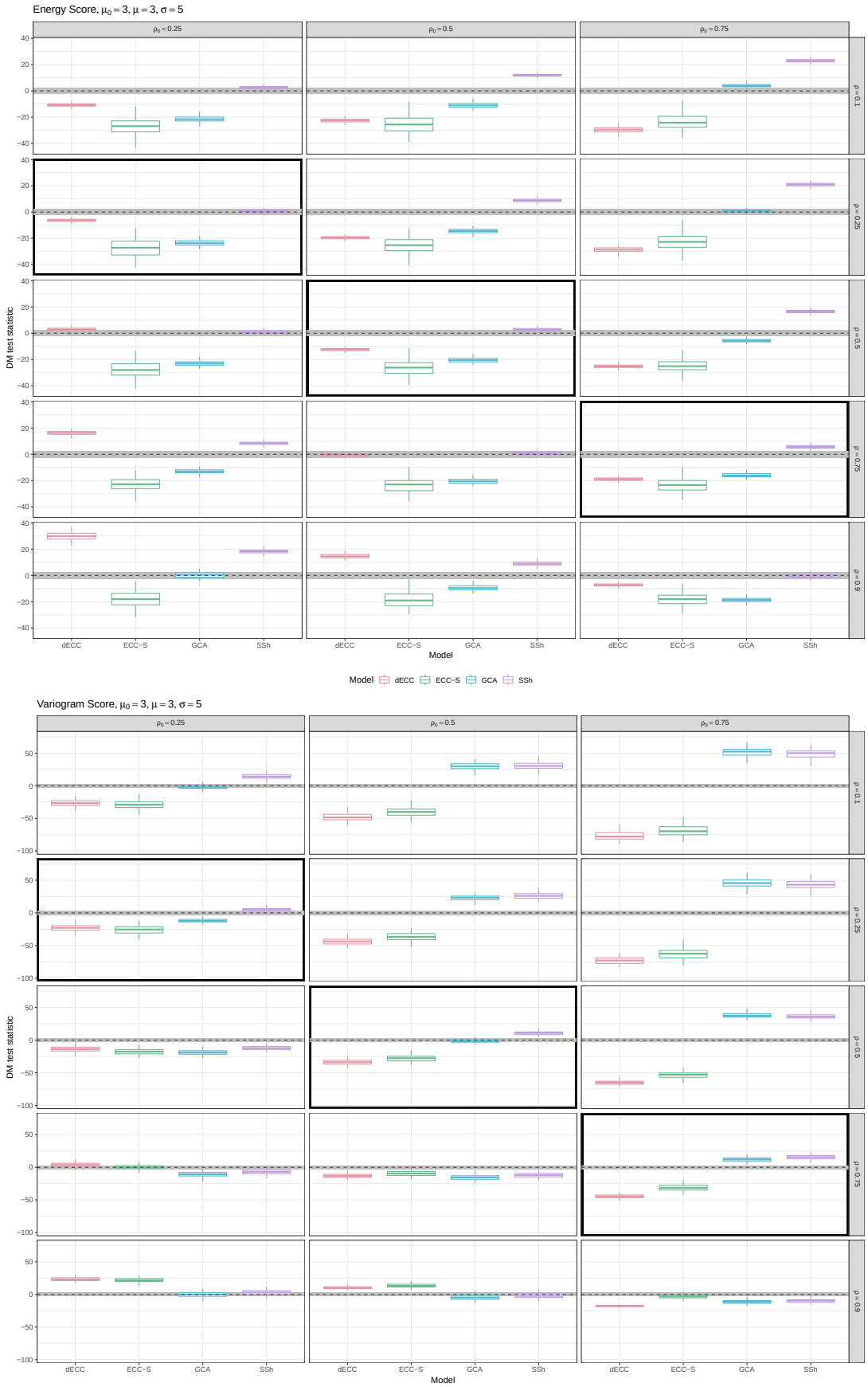


Figure 37: As Figure 13, but for $\mu_0 = 3, \epsilon = 3$, and $\sigma = 5$.

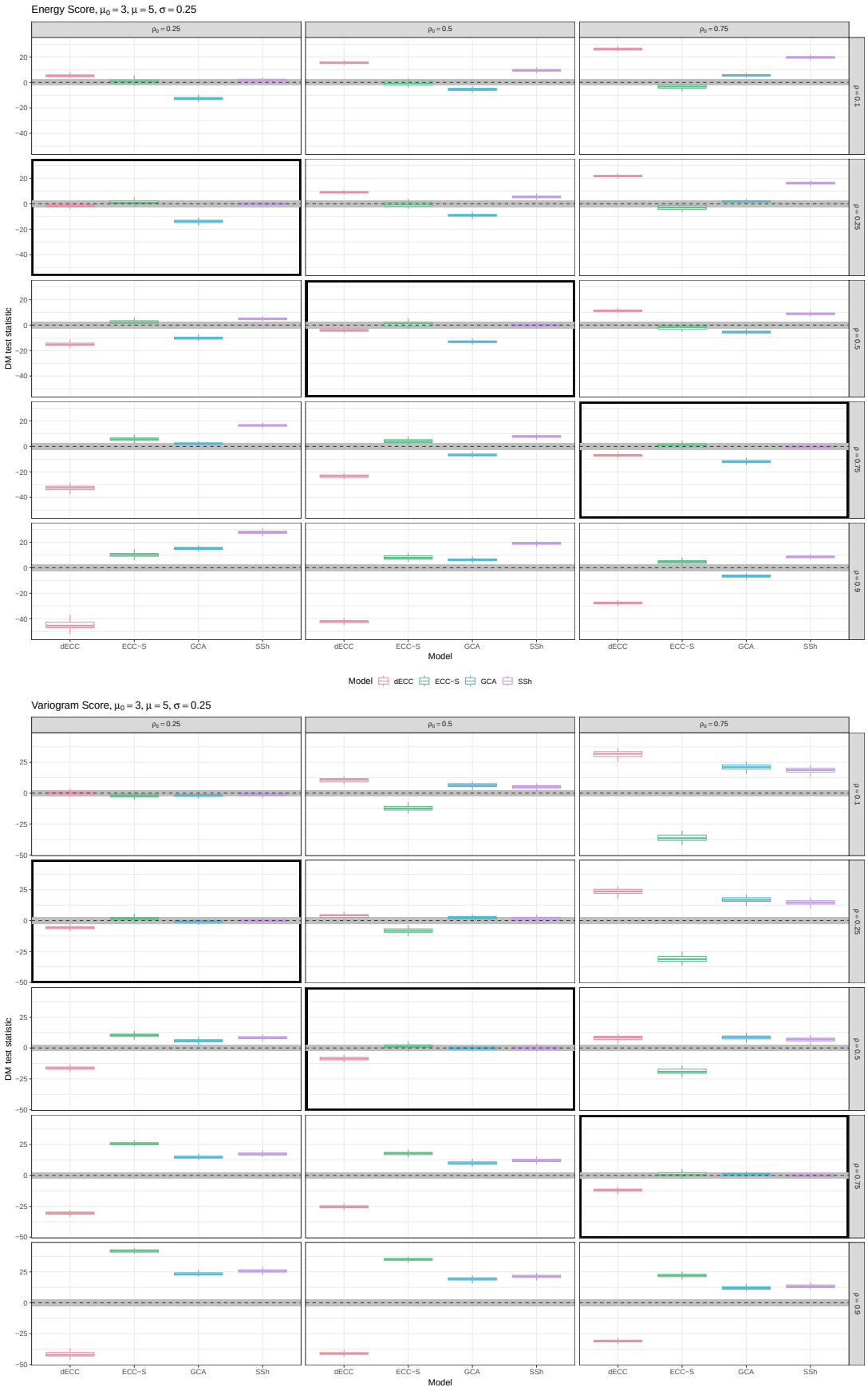


Figure 38: As Figure 13, but for $\mu_0 = 3, \epsilon = 5$, and $\sigma = 0.5$.

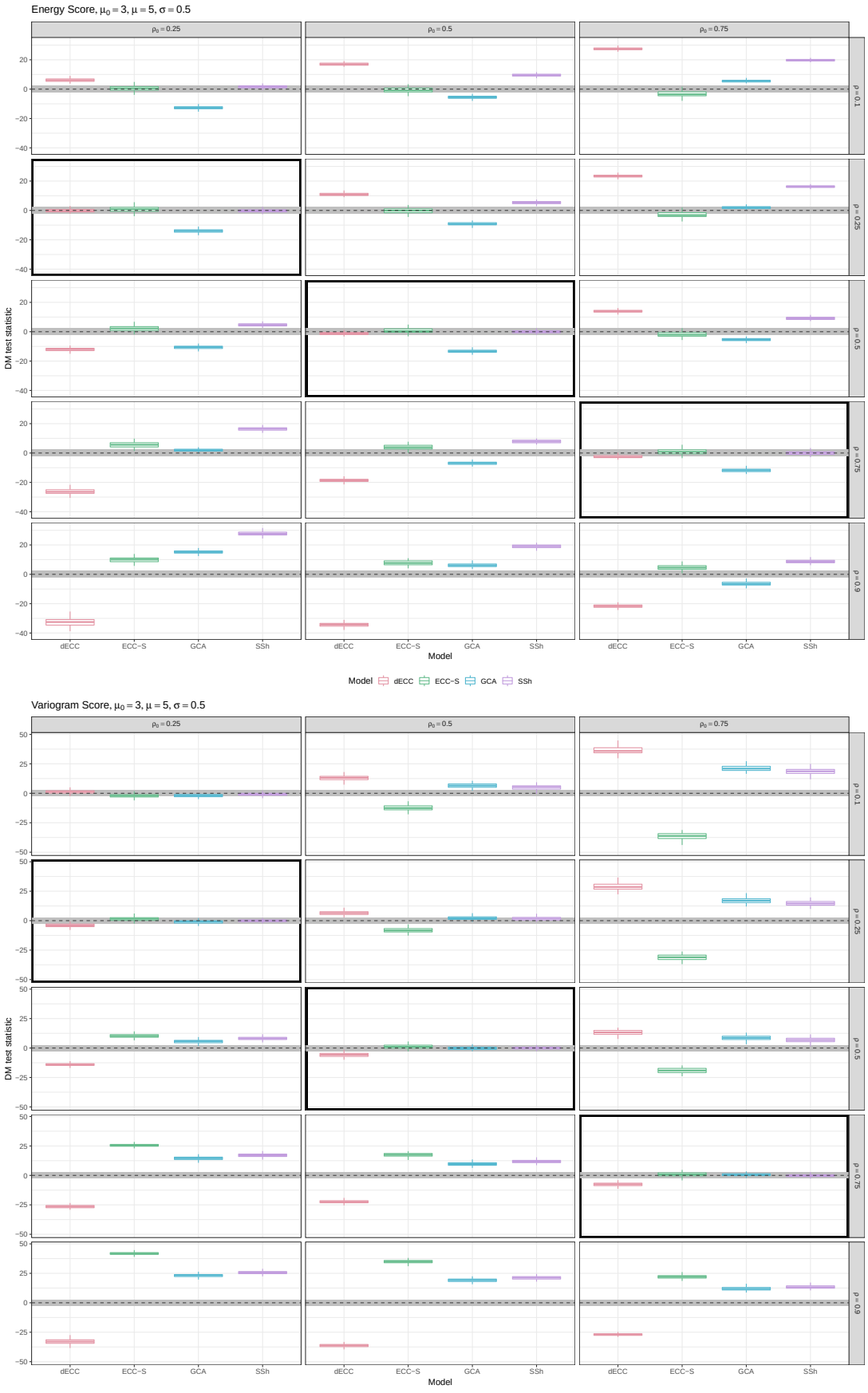


Figure 39: As Figure 13, but for $\mu_0 = 3, \epsilon = 5$, and $\sigma = 0.5$.

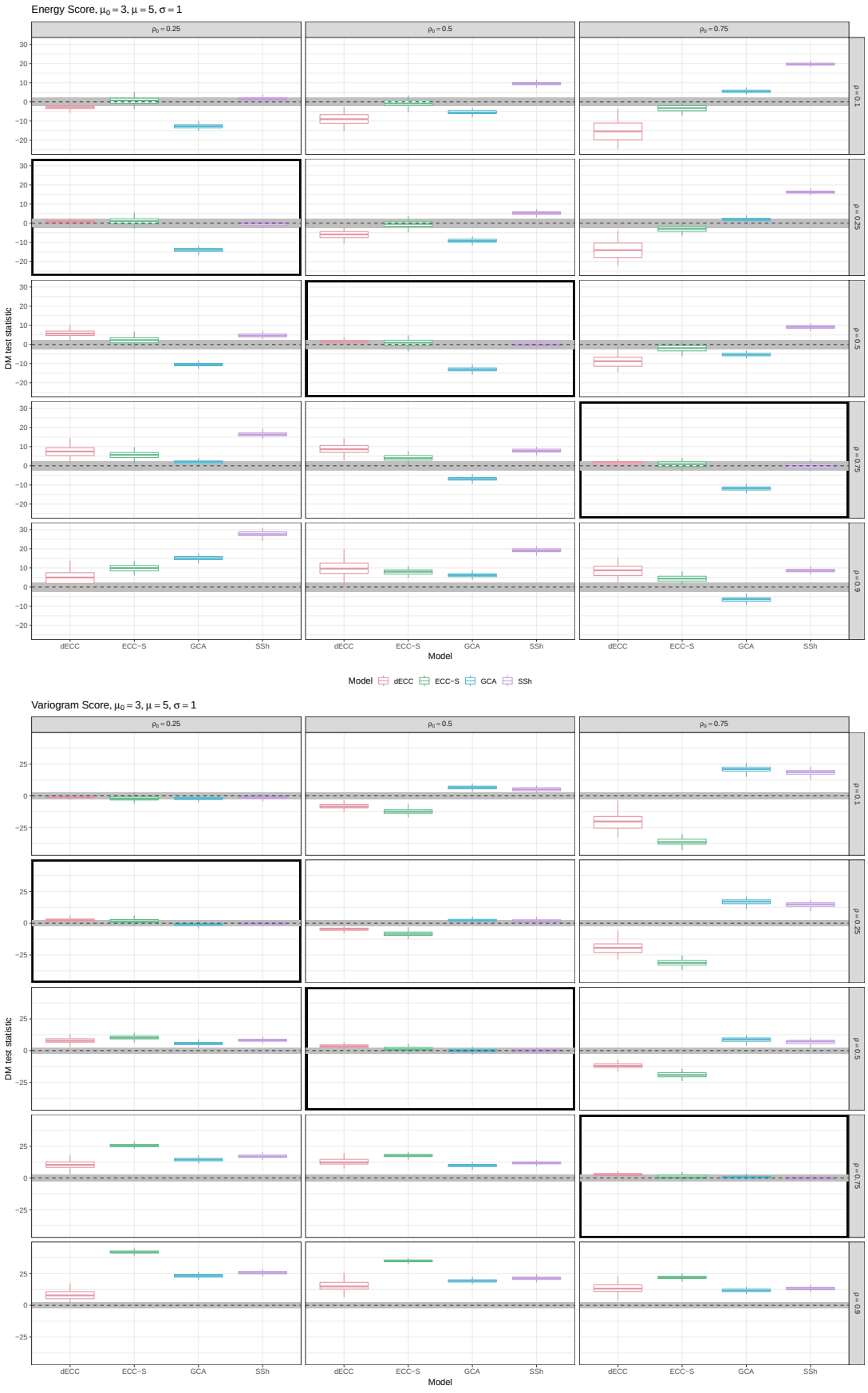


Figure 40: As Figure 13, but for $\mu_0 = 3, \epsilon = 5$, and $\sigma = 1$.

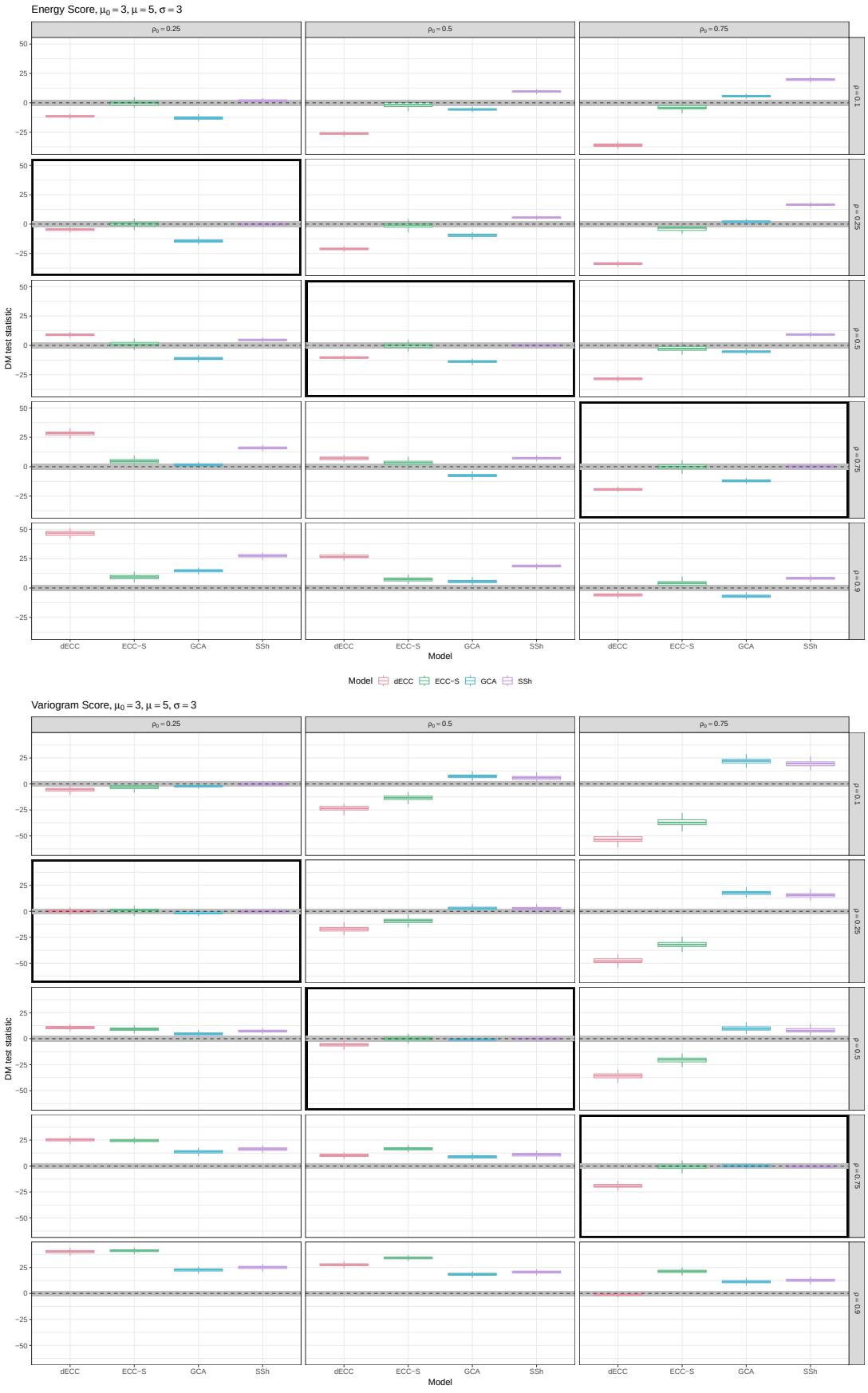


Figure 41: As Figure 13, but for $\mu_0 = 3, \epsilon = 5$, and $\sigma = 3$.

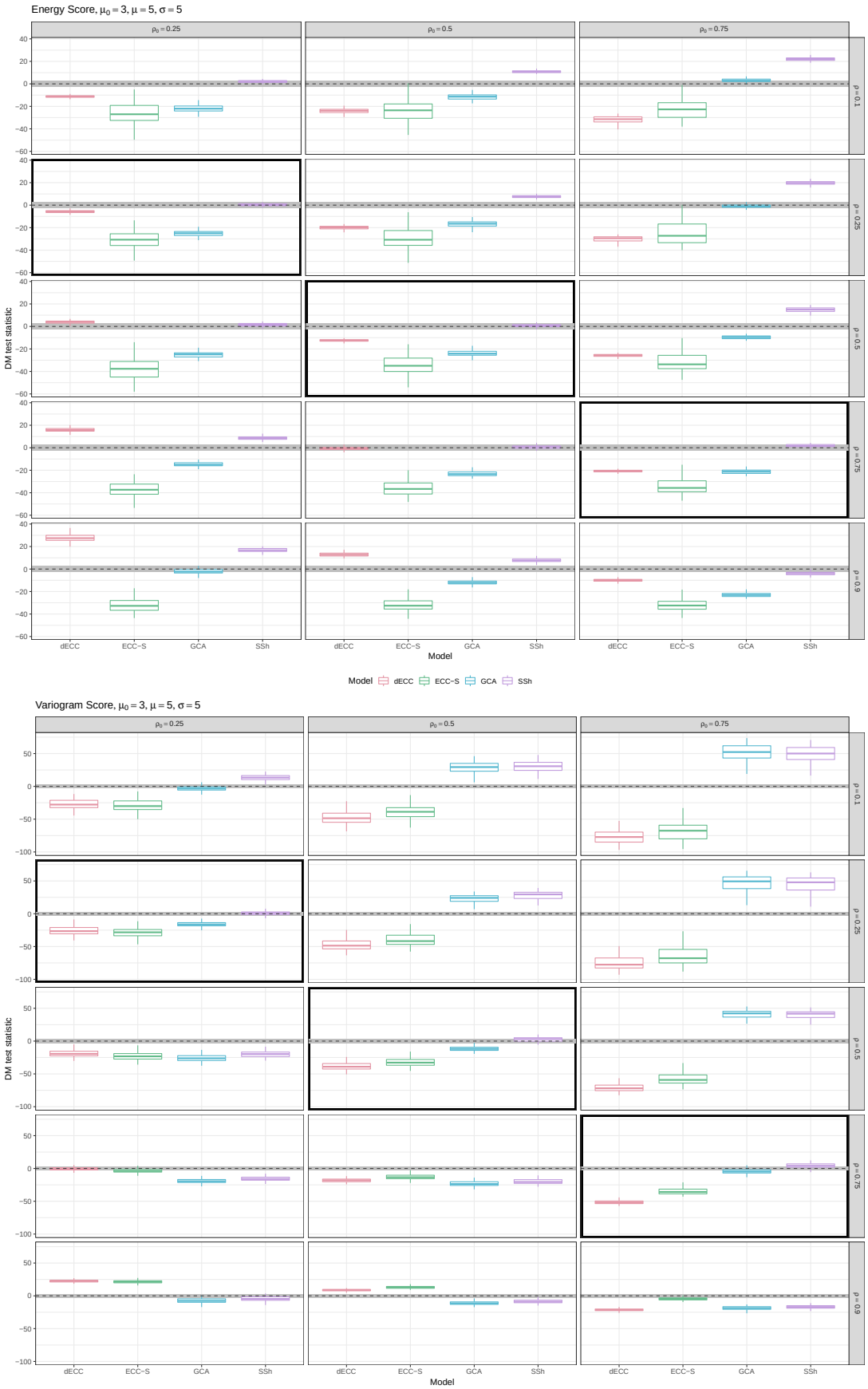


Figure 42: As Figure 13, but for $\mu_0 = 3, \epsilon = 5$, and $\sigma = 5$.

3 Additional results for Setting 3

Results for scenarios A, C, D from the following table.

	μ_y	ξ_y	σ_y	μ_x	ξ_x	σ_x
A	0.0	-0.1	1.0	1.0	0.0	0.2
B	0.0	-0.1	1.0	0.0	0.0	2.0
C	1.0	0.3	1.0	0.0	0.0	2.0
D	0.0	0.0	1.0	0.0	0.0	1.0

Table 3: Different simulation scenarios for Setting 3.

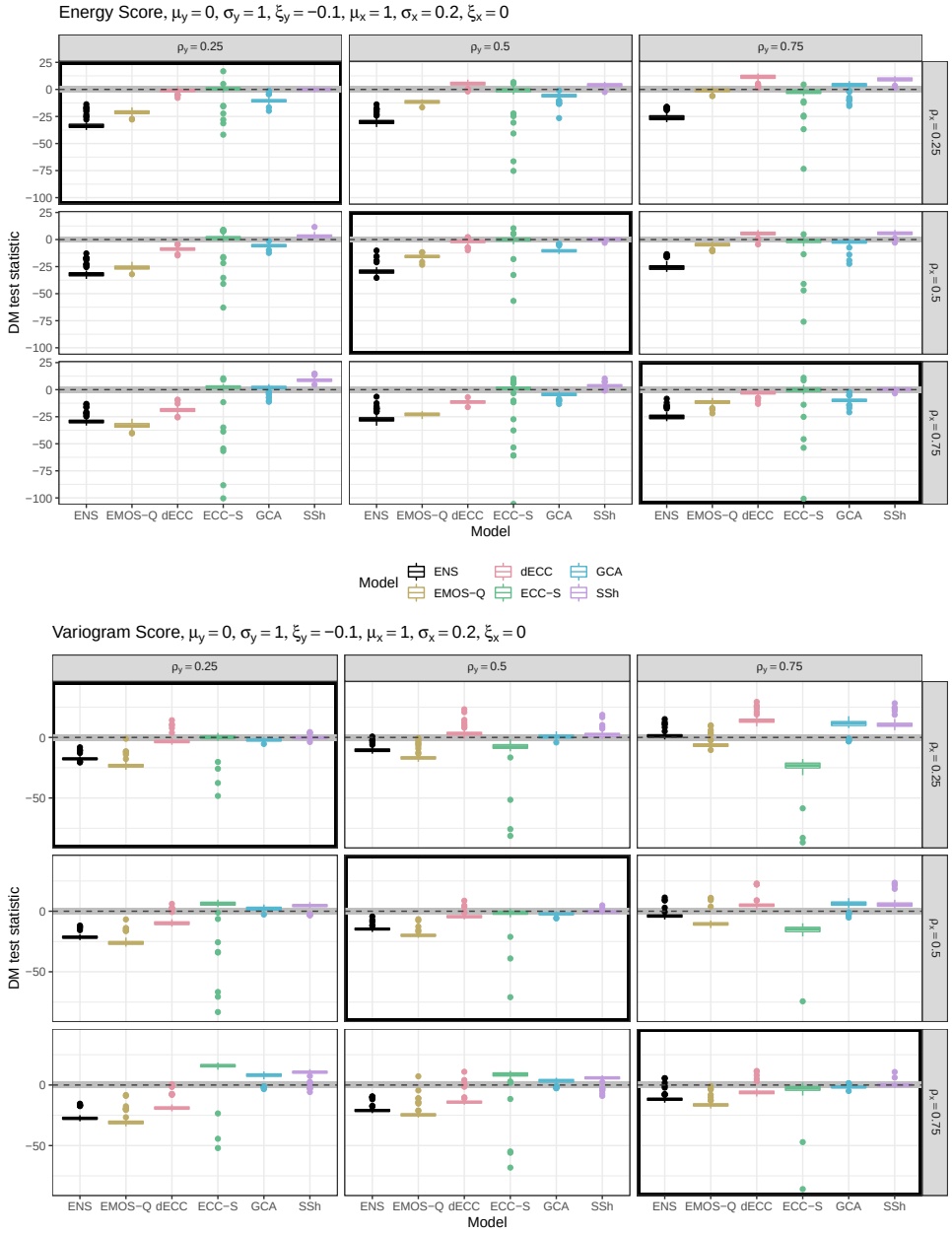


Figure 43: Summaries of DM test statistic values based on ES (top) and VS (bottom) for Setting 3, scenario A from Table 3. ECC-Q forecasts are used as reference model such that positive values of the test statistic indicate improvements over ECC-Q and negative value indicated deterioration of forecast skill. Boxplots summarize results of the 100 Monte Carlo repetitions of each individual experiment. The horizontal gray stripe indicates the acceptance region of the two-sided DM test under the null hypothesis of equal predictive performance at a level of 0.05. Simulation parameter choices where the correlation structure of the raw ensemble is correctly specified ($\rho_x = \rho_y$) are surrounded by black boxes.

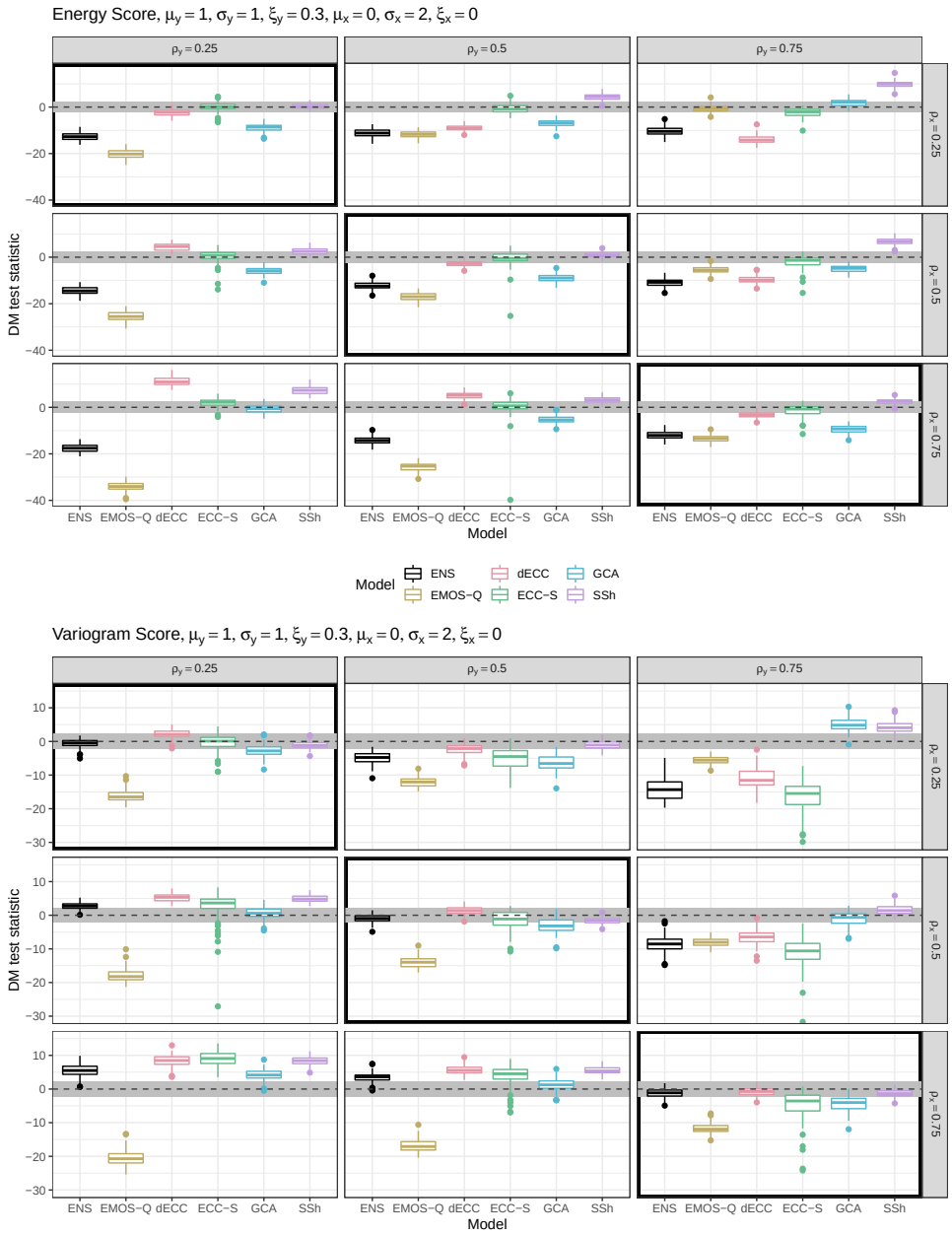


Figure 44: Summaries of DM test statistic values based on ES (top) and VS (bottom) for Setting 3, scenario C from Table 3. ECC-Q forecasts are used as reference model such that positive values of the test statistic indicate improvements over ECC-Q and negative value indicated deterioration of forecast skill. Boxplots summarize results of the 100 Monte Carlo repetitions of each individual experiment. The horizontal gray stripe indicates the acceptance region of the two-sided DM test under the null hypothesis of equal predictive performance at a level of 0.05. Simulation parameter choices where the correlation structure of the raw ensemble is correctly specified ($\rho_x = \rho_y$) are surrounded by black boxes.

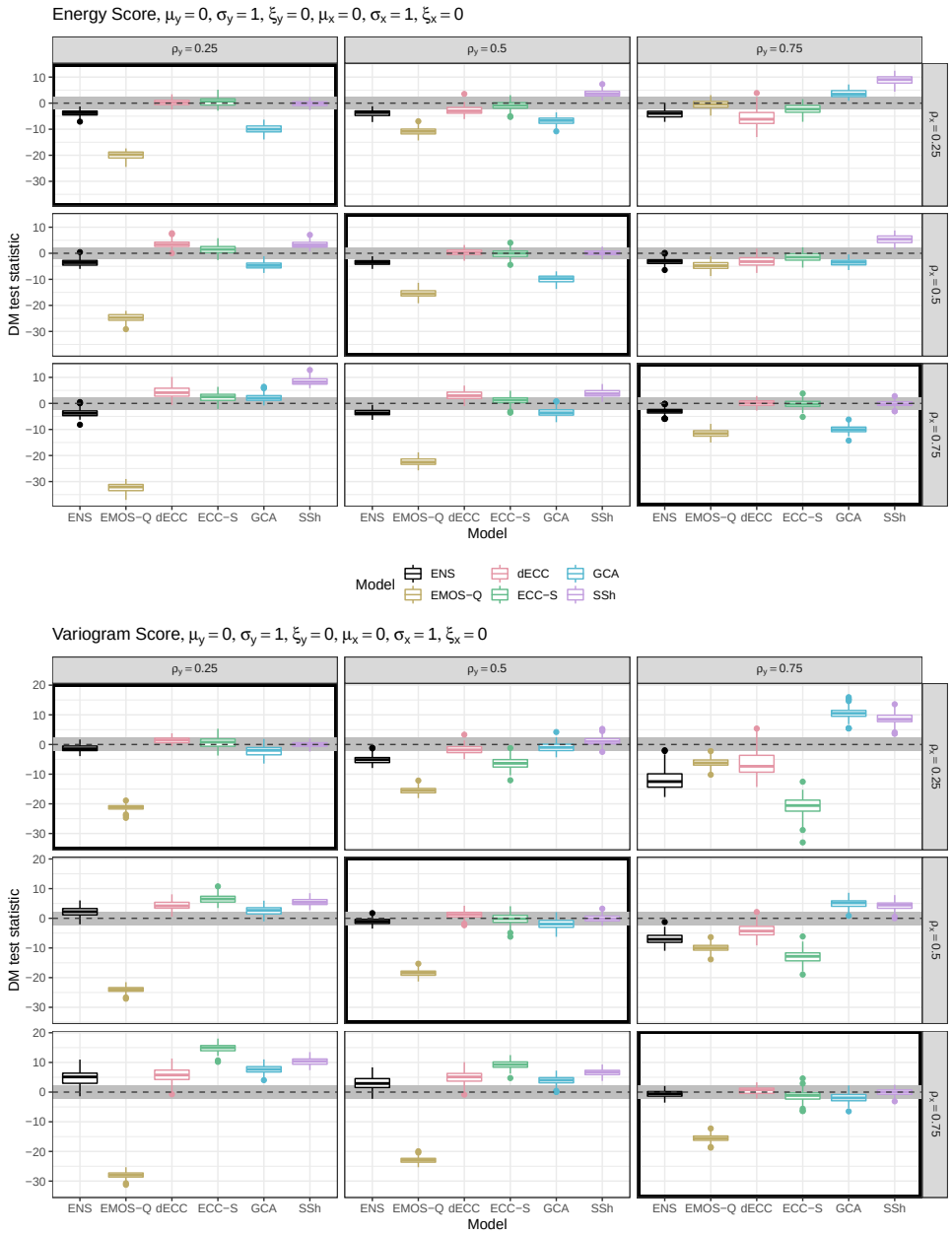


Figure 45: Summaries of DM test statistic values based on ES (top) and VS (bottom) for Setting 3, scenario D from Table 3. ECC-Q forecasts are used as reference model such that positive values of the test statistic indicate improvements over ECC-Q and negative value indicated deterioration of forecast skill. Boxplots summarize results of the 100 Monte Carlo repetitions of each individual experiment. The horizontal gray stripe indicates the acceptance region of the two-sided DM test under the null hypothesis of equal predictive performance at a level of 0.05. Simulation parameter choices where the correlation structure of the raw ensemble is correctly specified ($\rho_x = \rho_y$) are surrounded by black boxes.

4 Additional results for Setting 4

Results for all values of ρ and ρ_0 are shown in Figure 46.

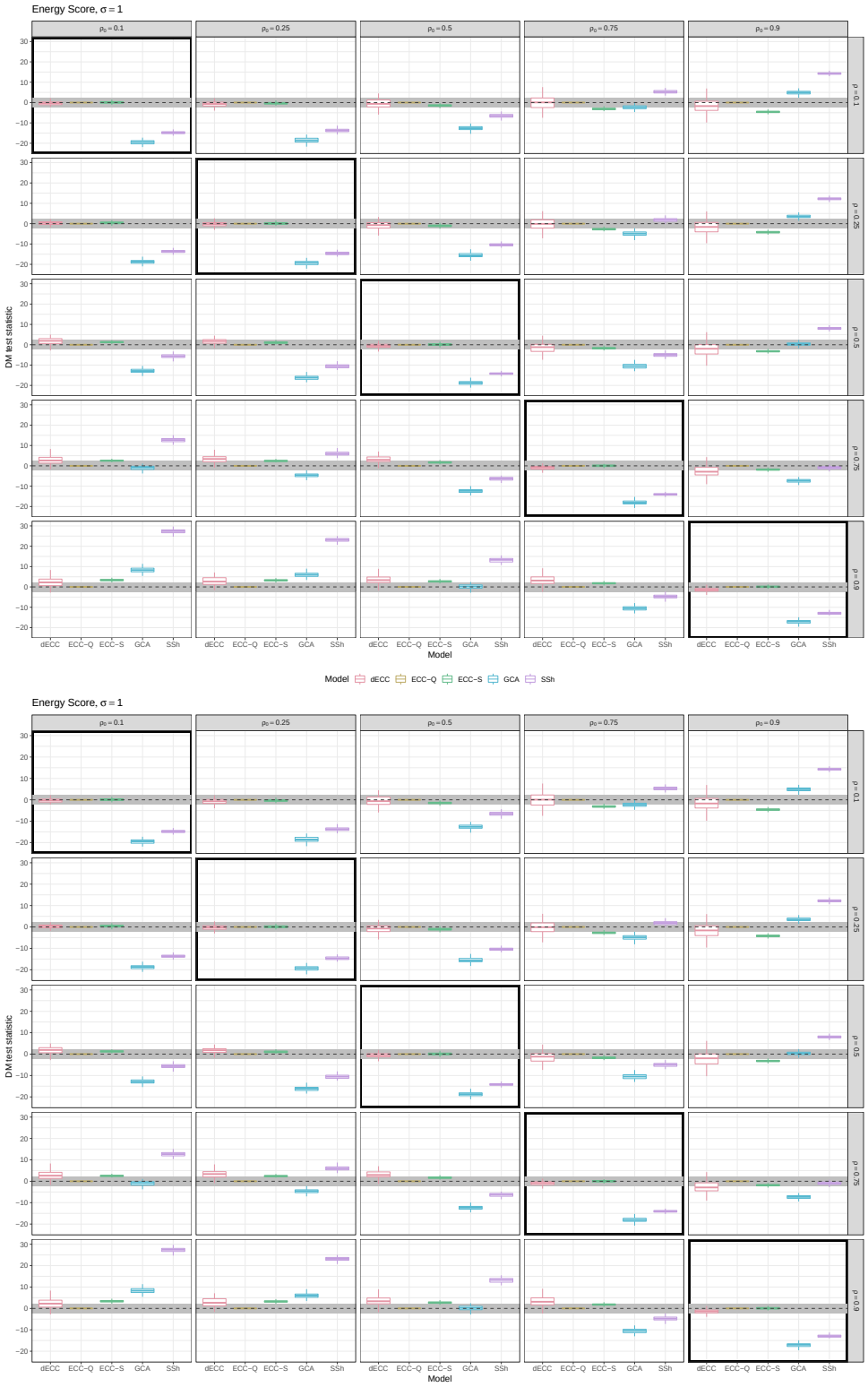


Figure 46: Summaries of DM test statistic values based on ES (top) and VS (bottom) for Setting 4. ECC-Q forecasts are used as reference model such that positive values of the test statistic indicate improvements over ECC-Q and negative value indicated deterioration of forecast skill. Boxplots summarize results of the 100 Monte Carlo repetitions of each individual experiment. The horizontal gray stripe indicates the acceptance region of the two-sided DM test under the null hypothesis of equal predictive performance at a level of 0.05. Simulation parameter choices where the correlation structure of the raw ensemble is correctly specified ($\rho = \rho_0$) are surrounded by black boxes.