Analysing Discretization Methods for Single-Cell RNA-Sequencing Data when Inferring Gene Regulatory Networks via Cartesian Genetic Programming

José Eduardo H. da Silva[1], Heder S. Bernardino[1], Itamar L. de Oliveira[1], Alex B. Vieira[1], Helio J.C. Barbosa[1,2]

[1]*Universidade Federal de Juiz de Fora*
Juiz de Fora, Brazil
jehenriques@ice.ufjf.br, heder@ice.ufjf.br, alex.borges@ice.ufjf.br, itamar.leite@ice.ufjf.br
[2]*Laboratório Nacional de Computação Científica*
Petrópolis, Brazil
hcbm@lncc.br

1 Introduction

This supplementary material presents additional results that were obtained by Cartesian Genetic Programming (CGP) when using several discretization approaches. Section 2 presents the comparison of the use of spline for smooth the data. This additional content is organized as follows: In Section 3, the parameter analysis for Top%X and Max -X%Max are performed.

2 Spline Analysis

Empty boxplots means that CGP did not found a feasible solution for that configuration.

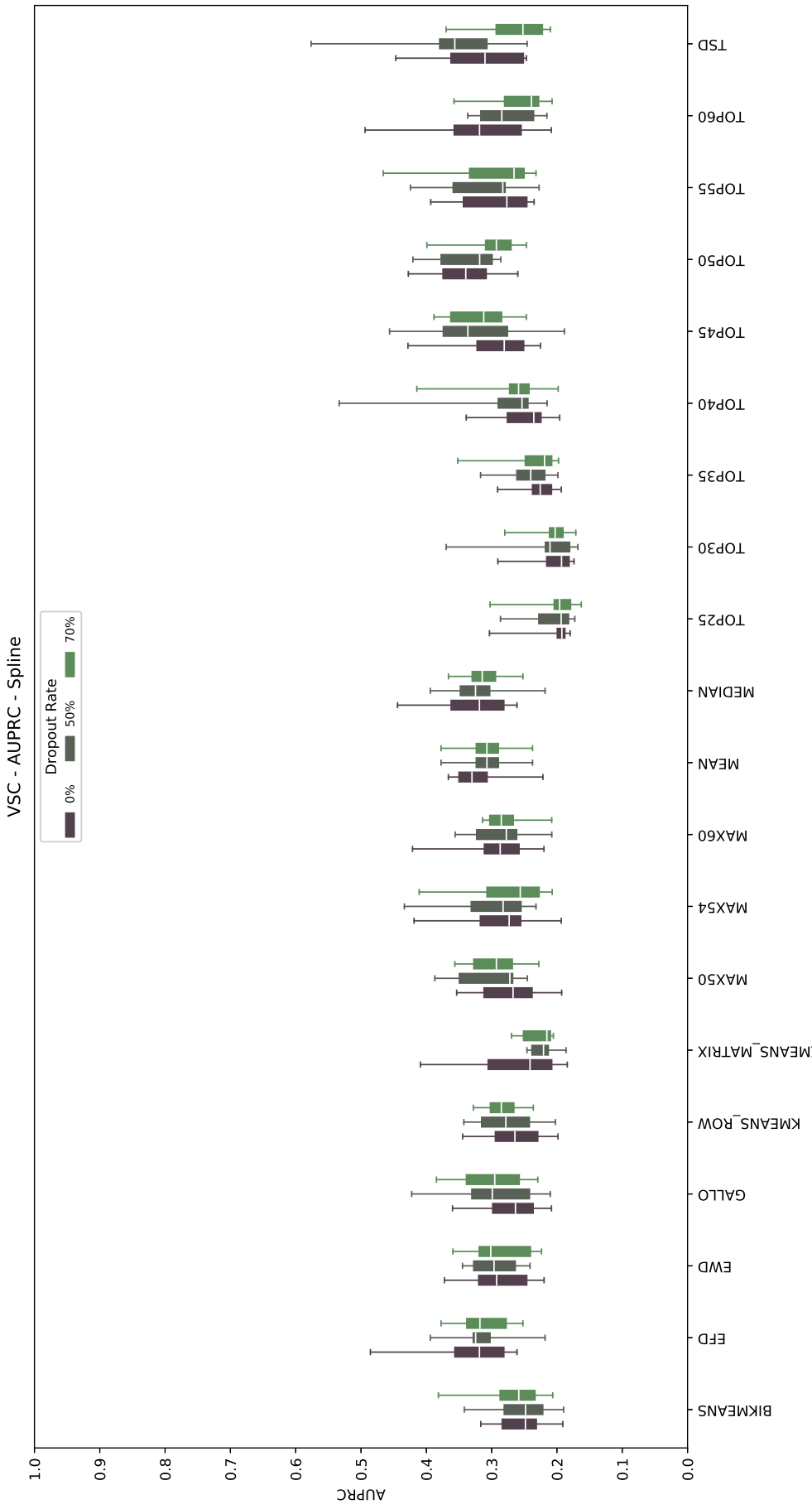

Figure 1. Results for problem VSC with Spline, considering AUPRC.

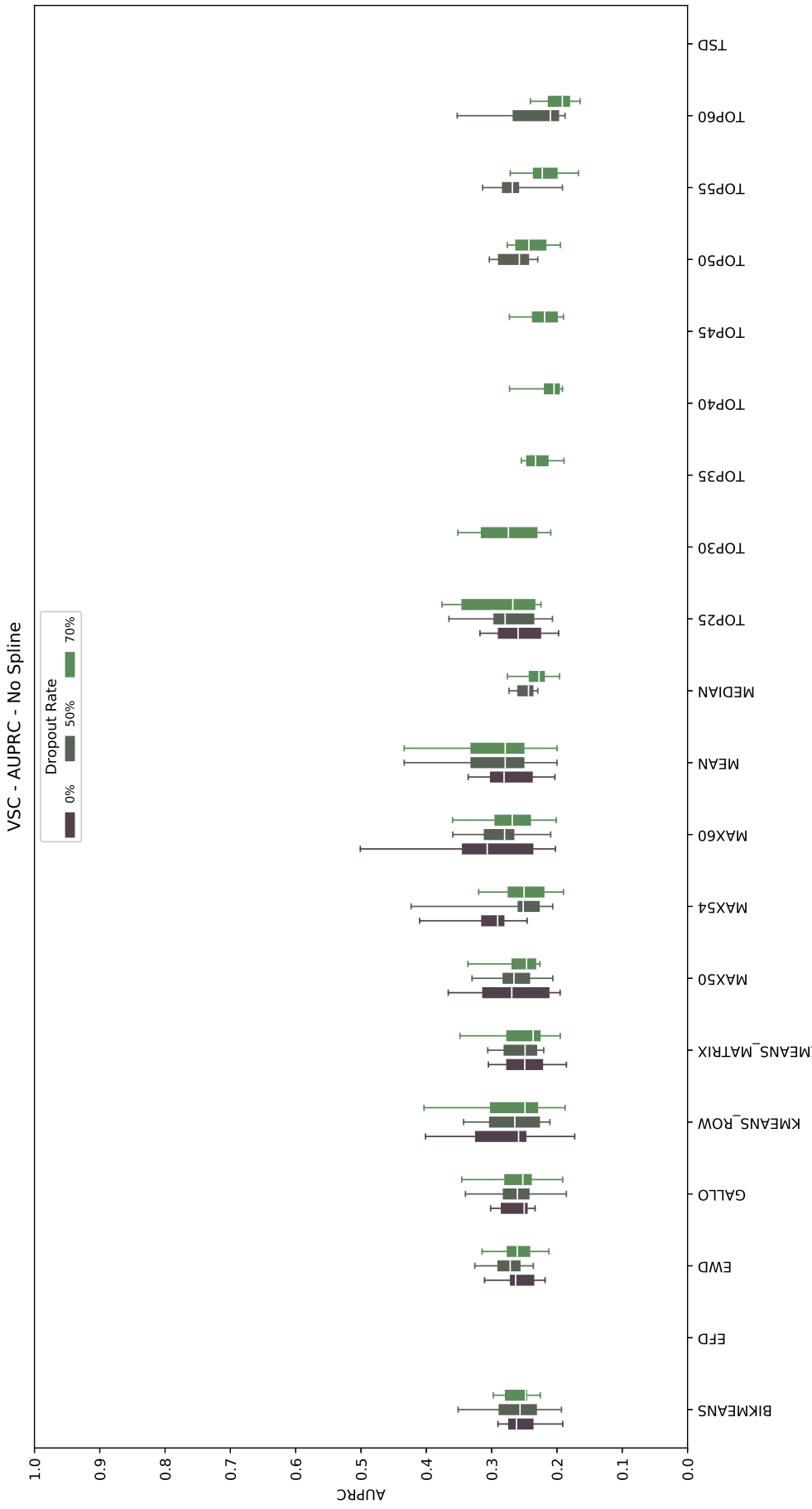

Figure 2. Results for problem VSC without Spline, considering AUPRC.

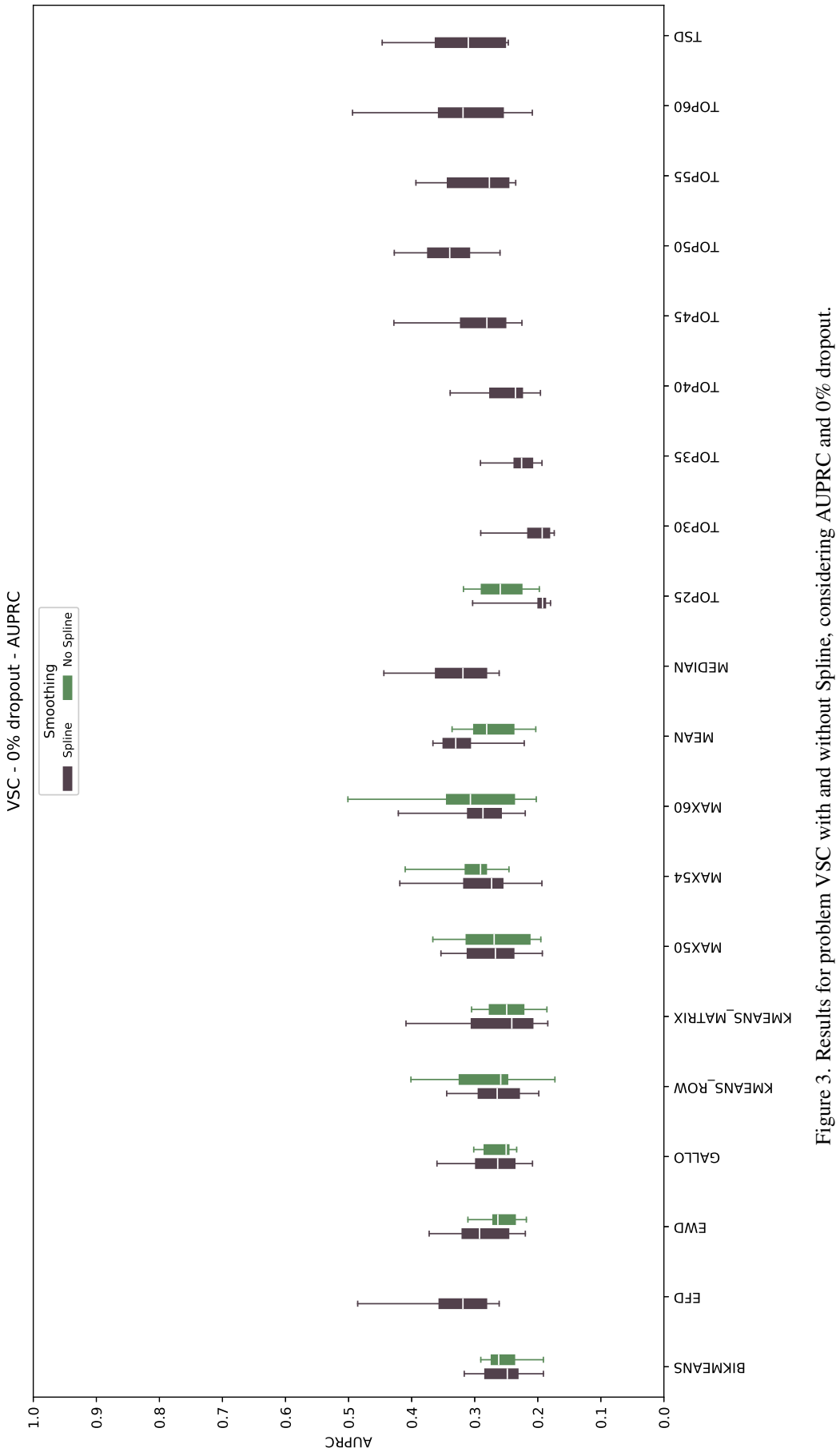

Figure 3. Results for problem VSC with and without Spline, considering AUPRC and 0% dropout.

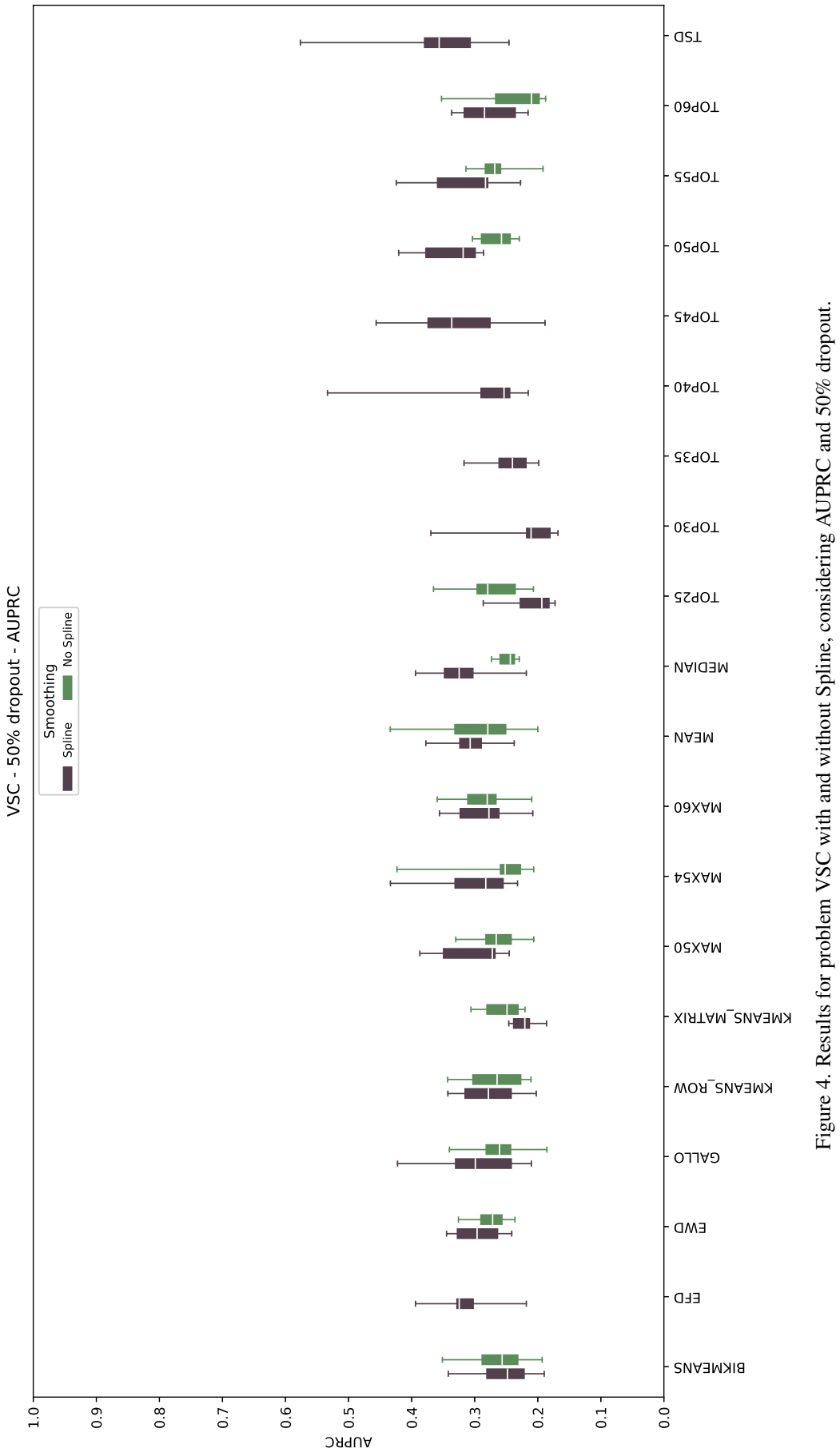

Figure 4. Results for problem VSC with and without Spline, considering AUPRC and 50% dropout.

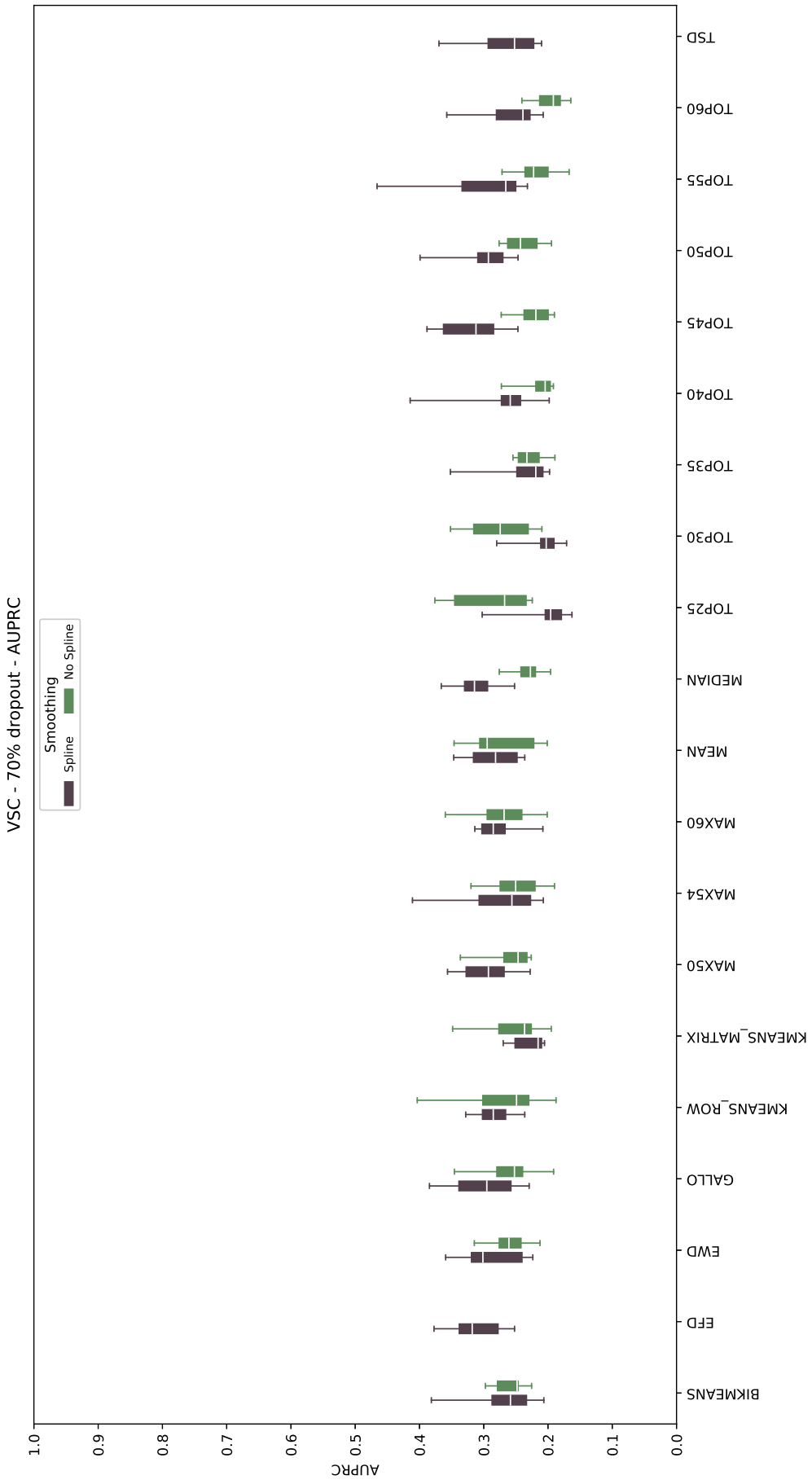

Figure 5. Results for problem VSC with and without Spline, considering AUPRC and 70% dropout.

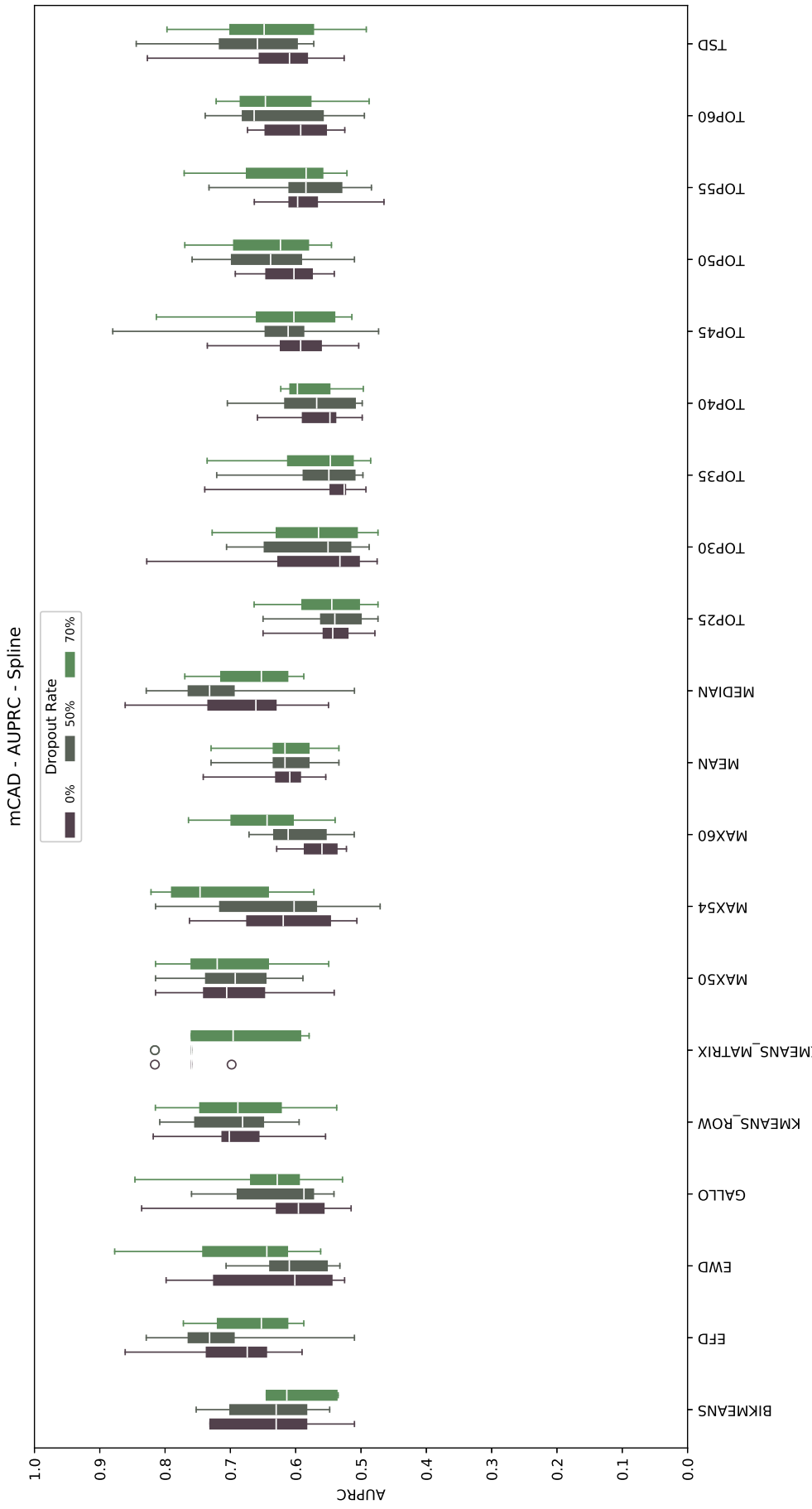

Figure 6. Results for problem mCAD with Spline, considering AUPRC.

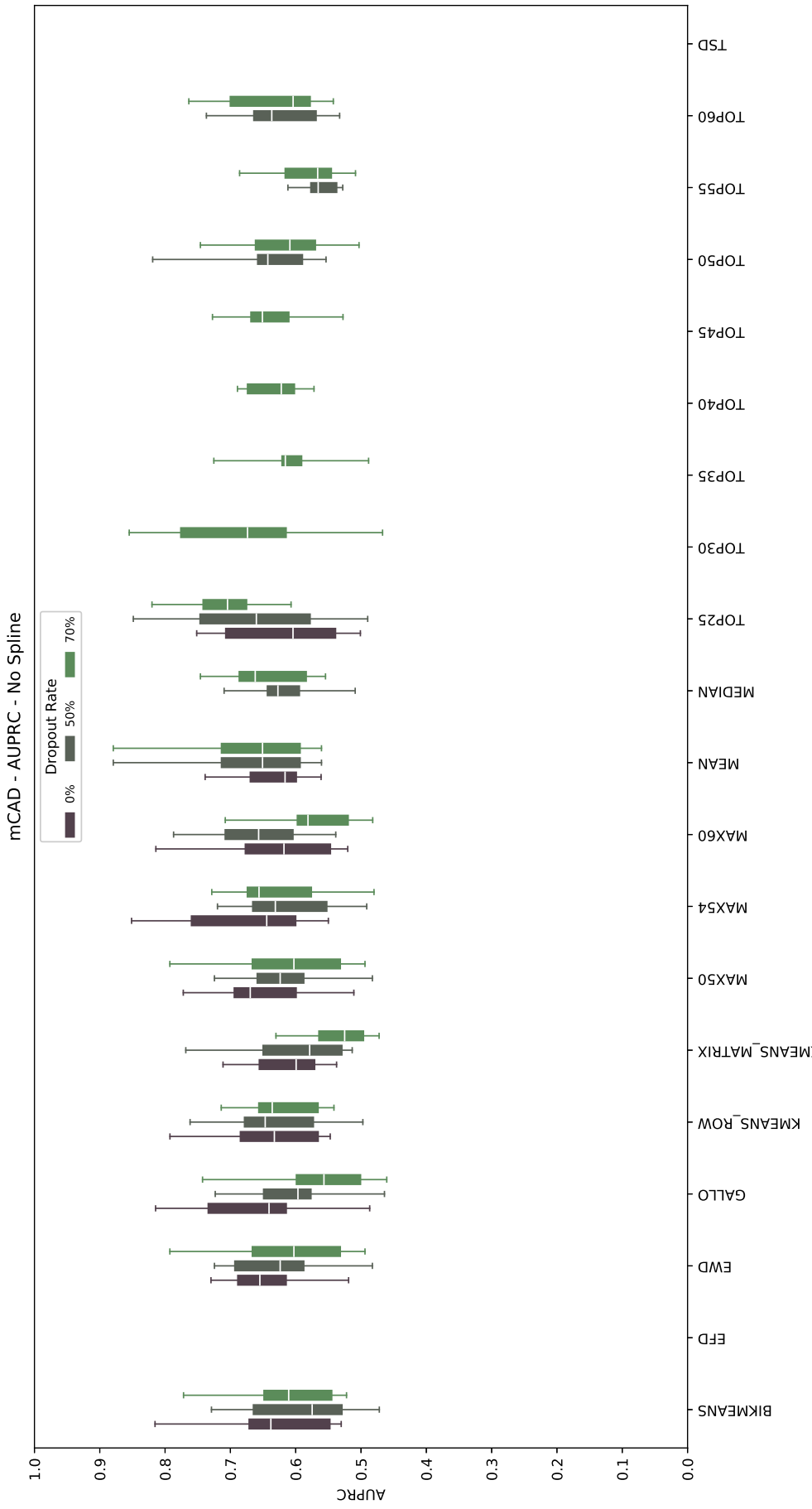

Figure 7. Results for problem mCAD without Spline, considering AUPRC.

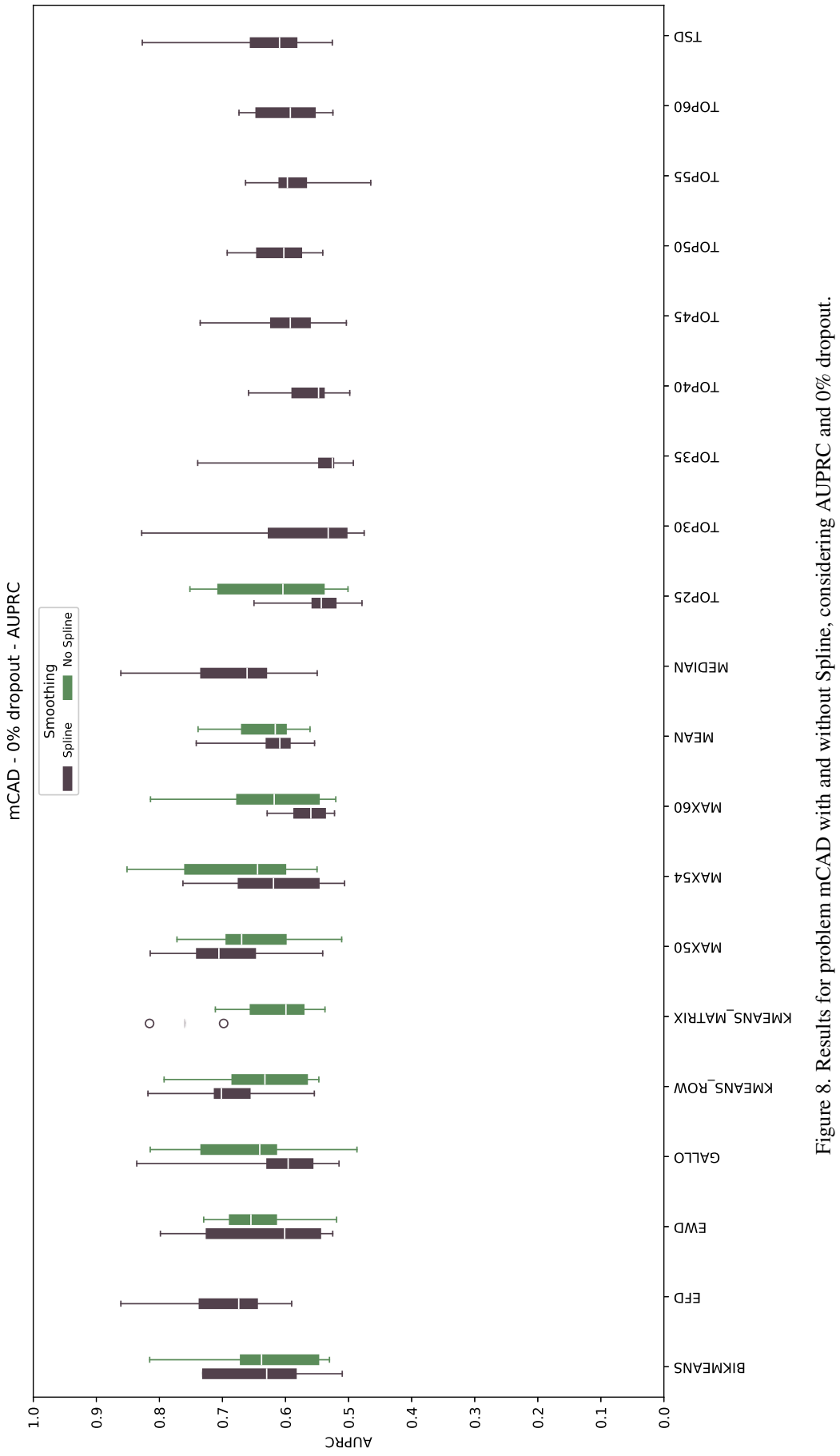

Figure 8. Results for problem mCAD with and without Spline, considering AUPRC and 0% dropout.

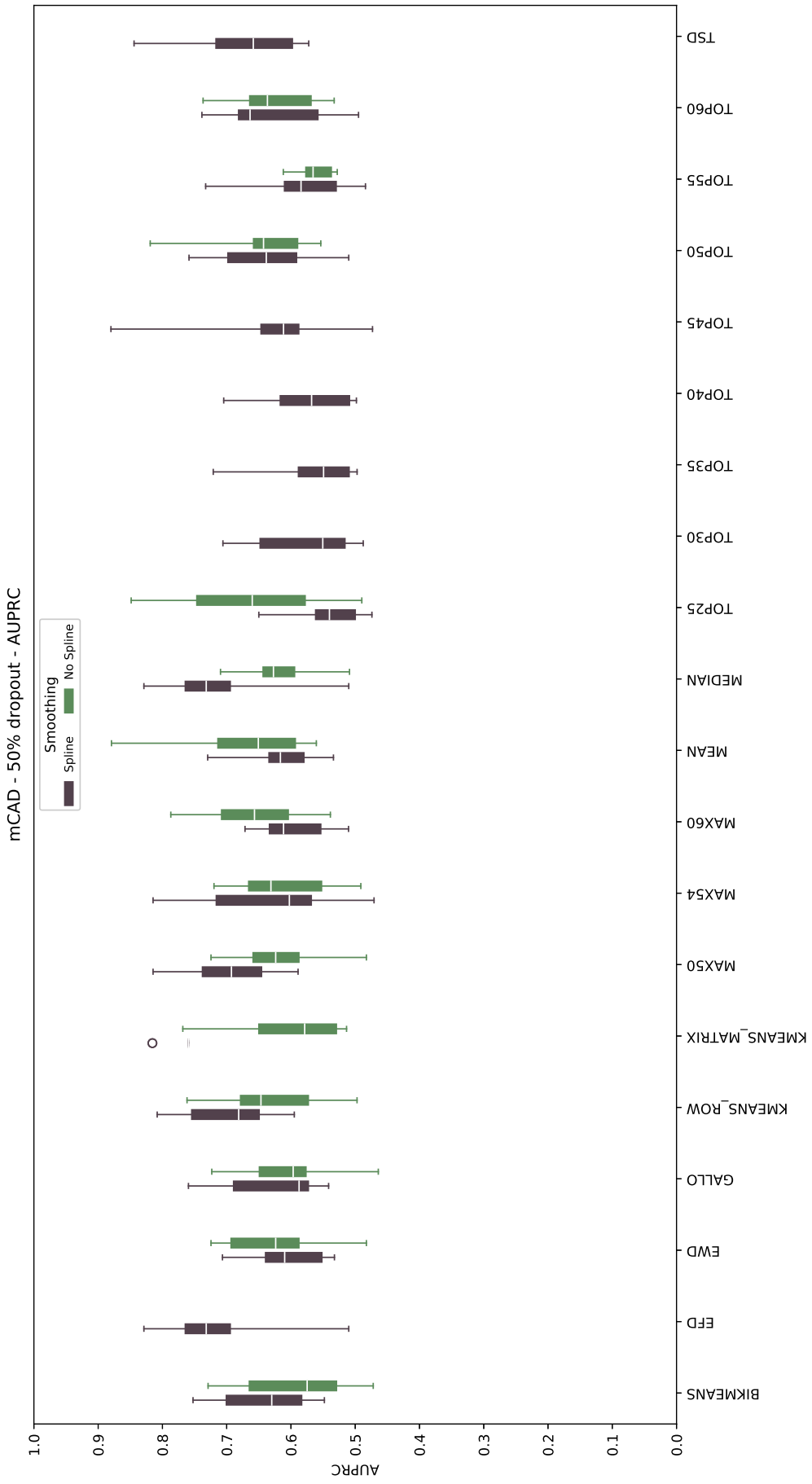

Figure 9. Results for problem mCAD with and without Spline, considering AUPRC and 50% dropout.

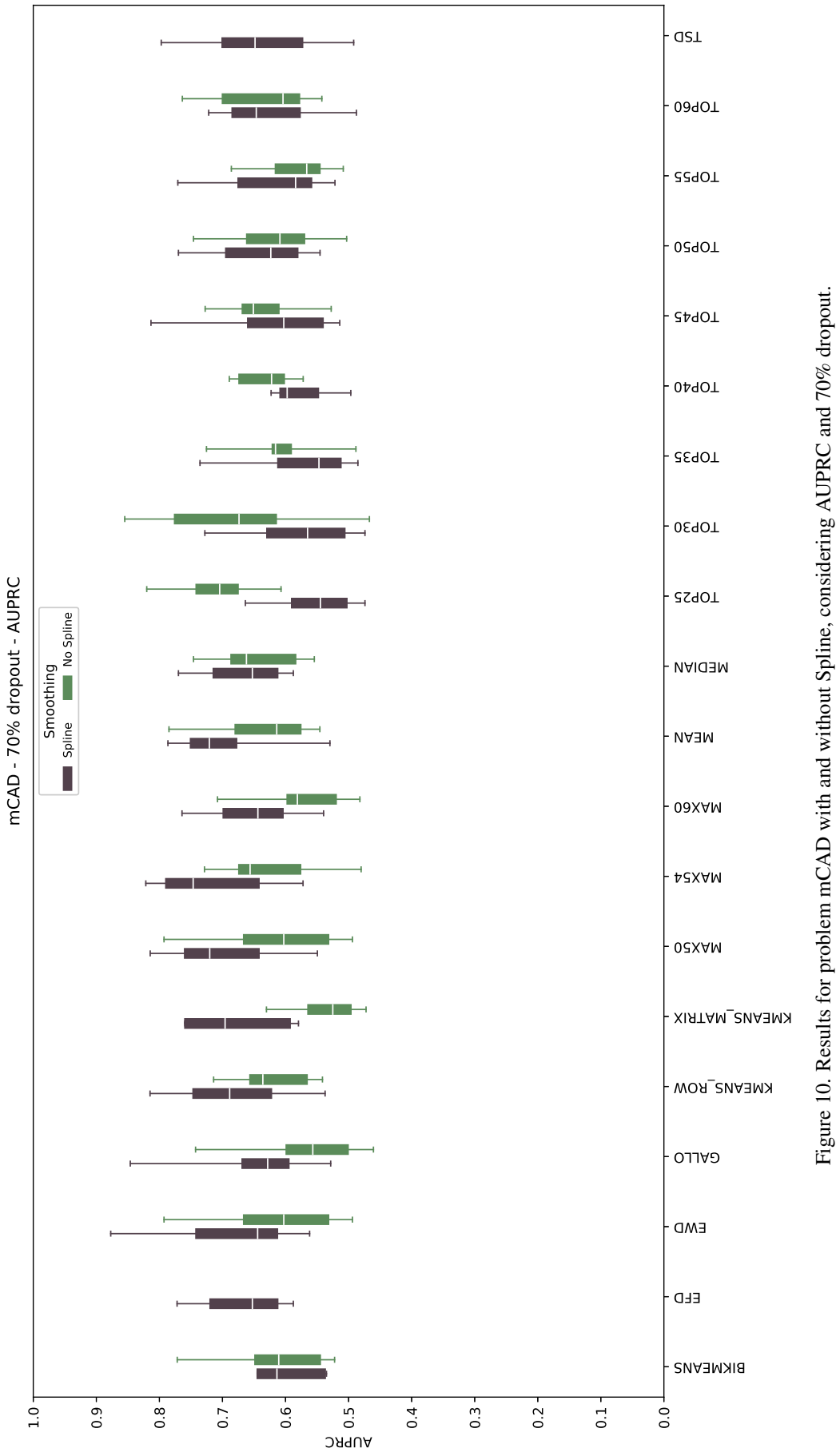

Figure 10. Results for problem mCAD with and without Spline, considering AUPRC and 70% dropout.

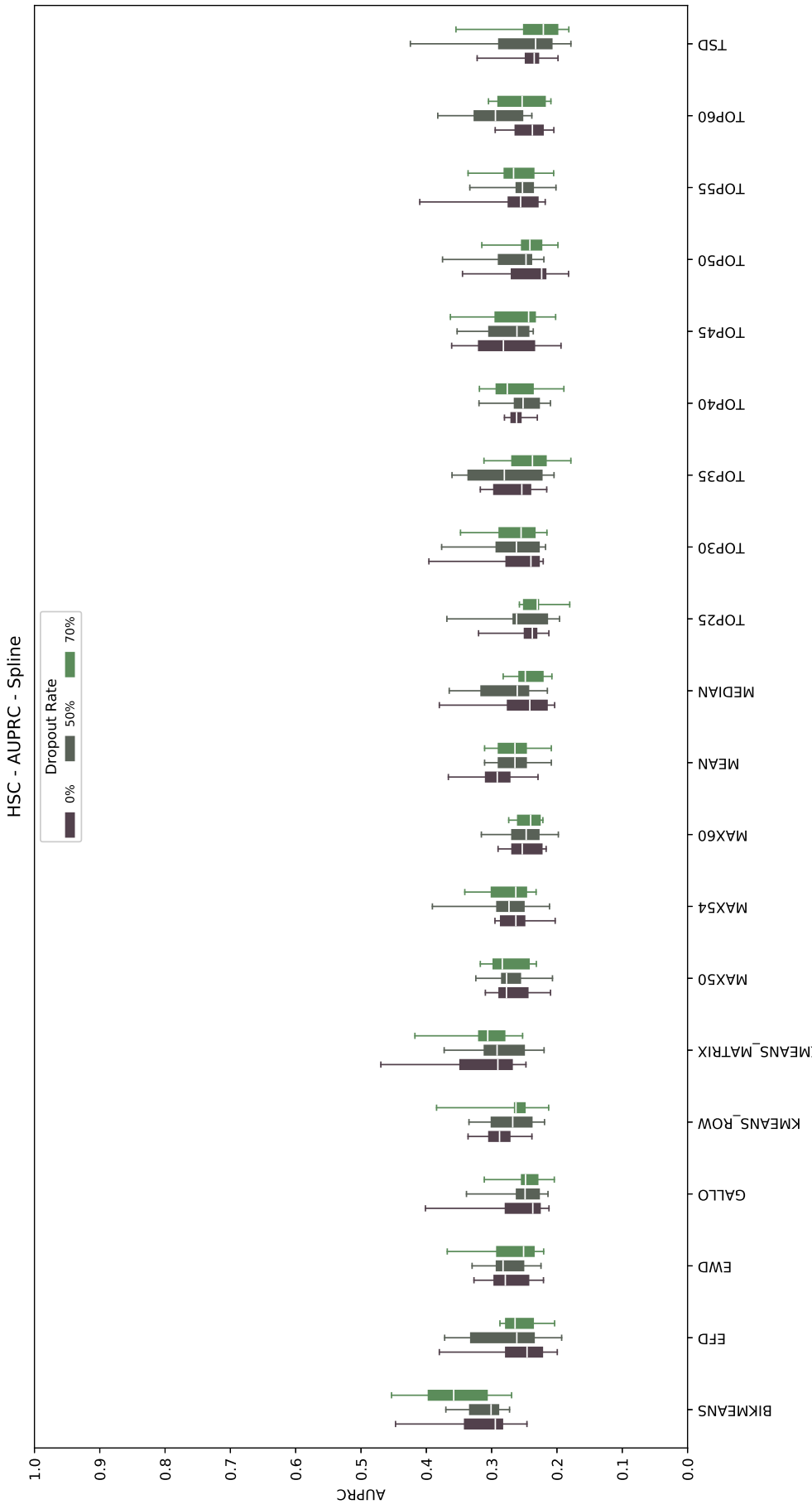

Figure 11. Results for problem HSC with Spline, considering AUPRC.

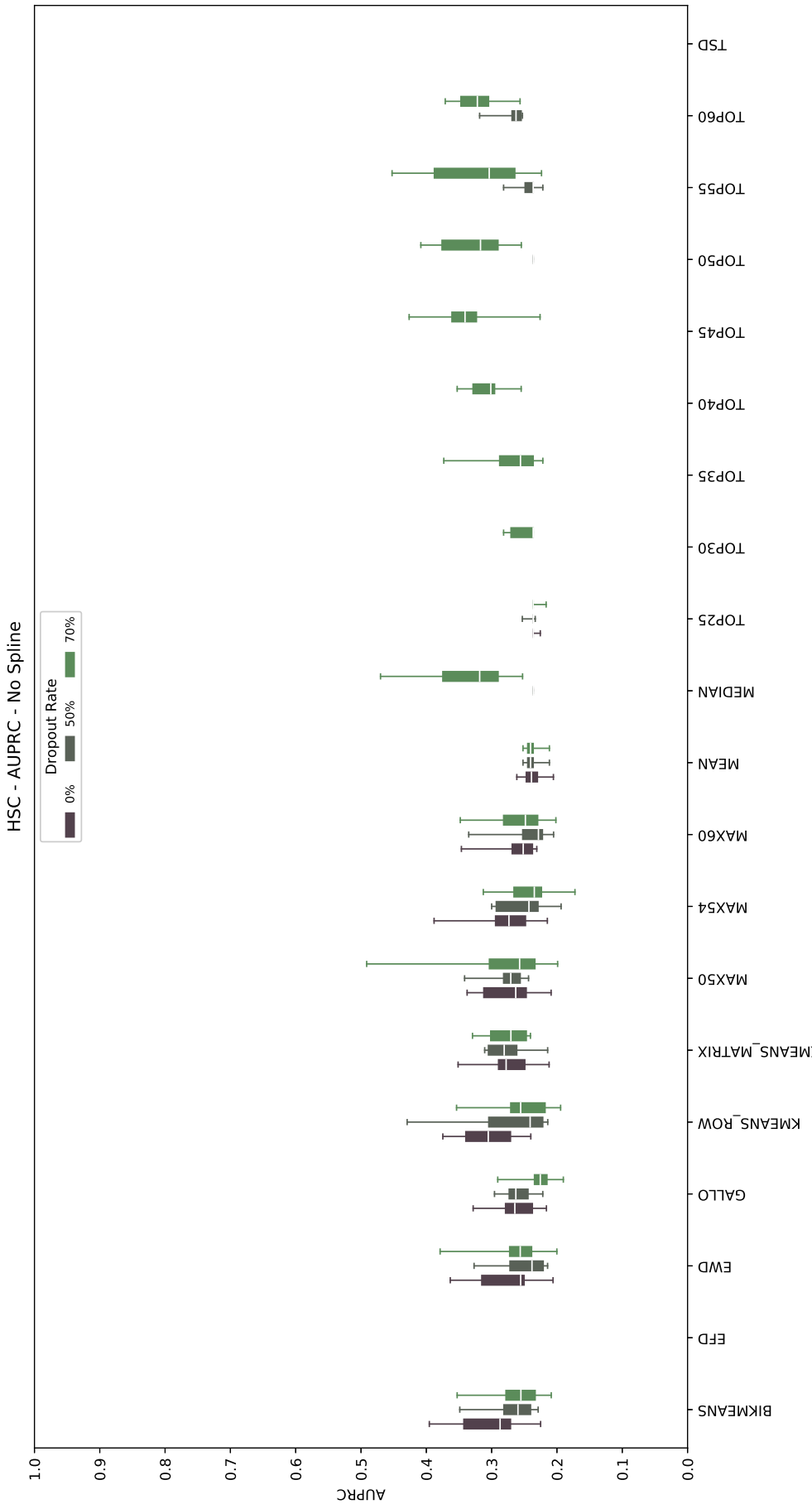

Figure 12. Results for problem HSC without Spline, considering AUPRC.

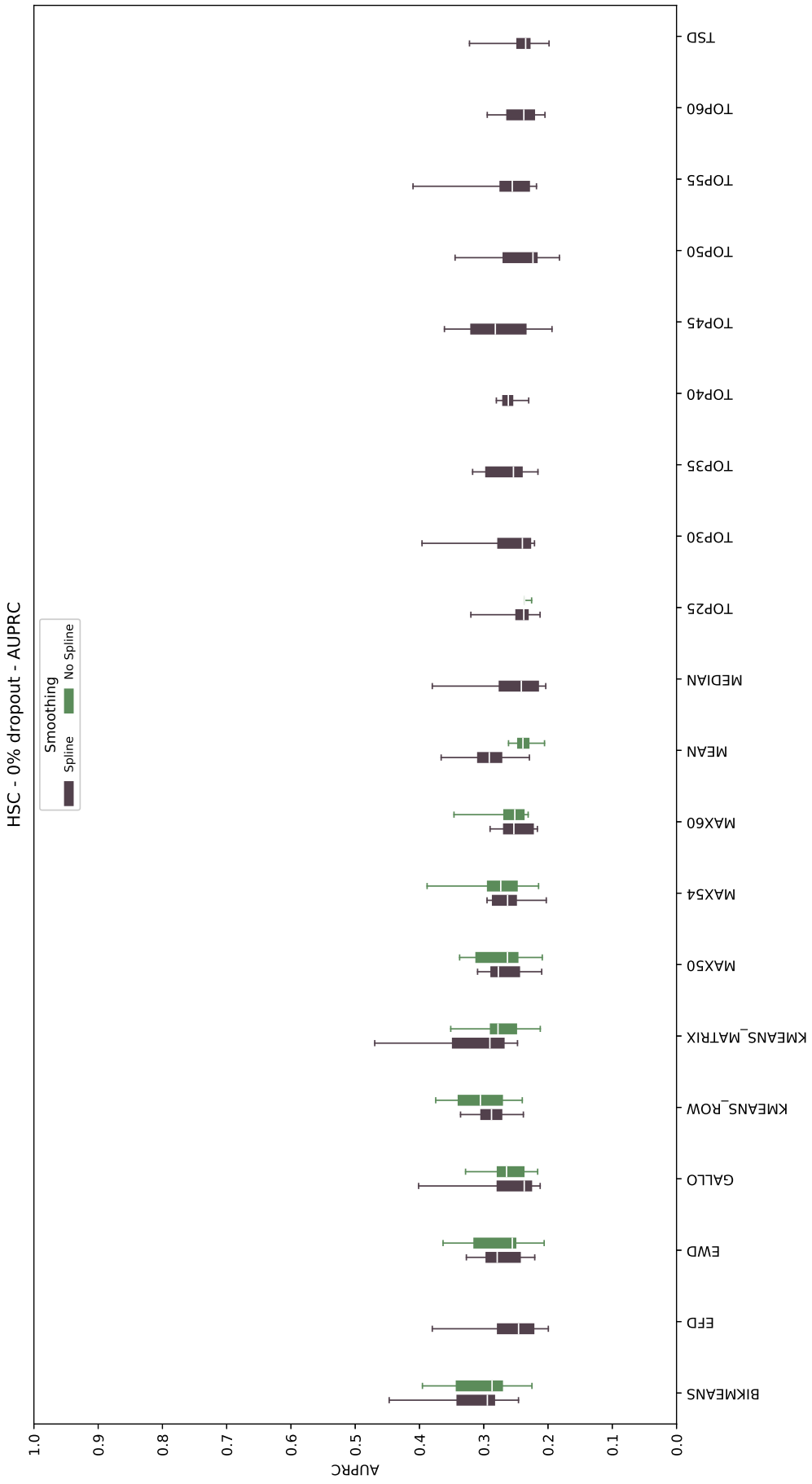

Figure 13. Results for problem HSC with and without Spline, considering AUPRC and 0% dropout.

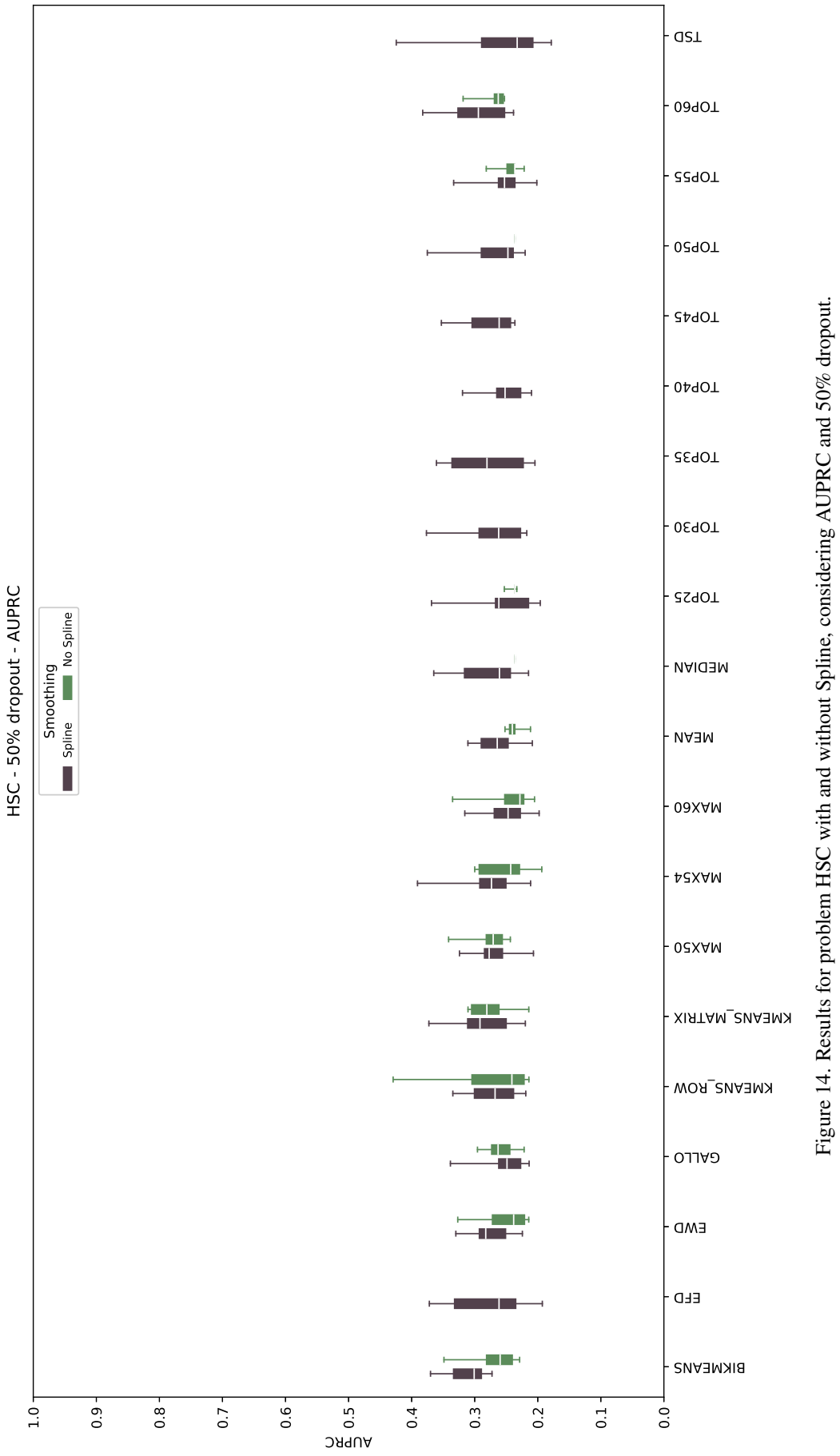

Figure 14. Results for problem HSC with and without Spline, considering AUPRC and 50% dropout.

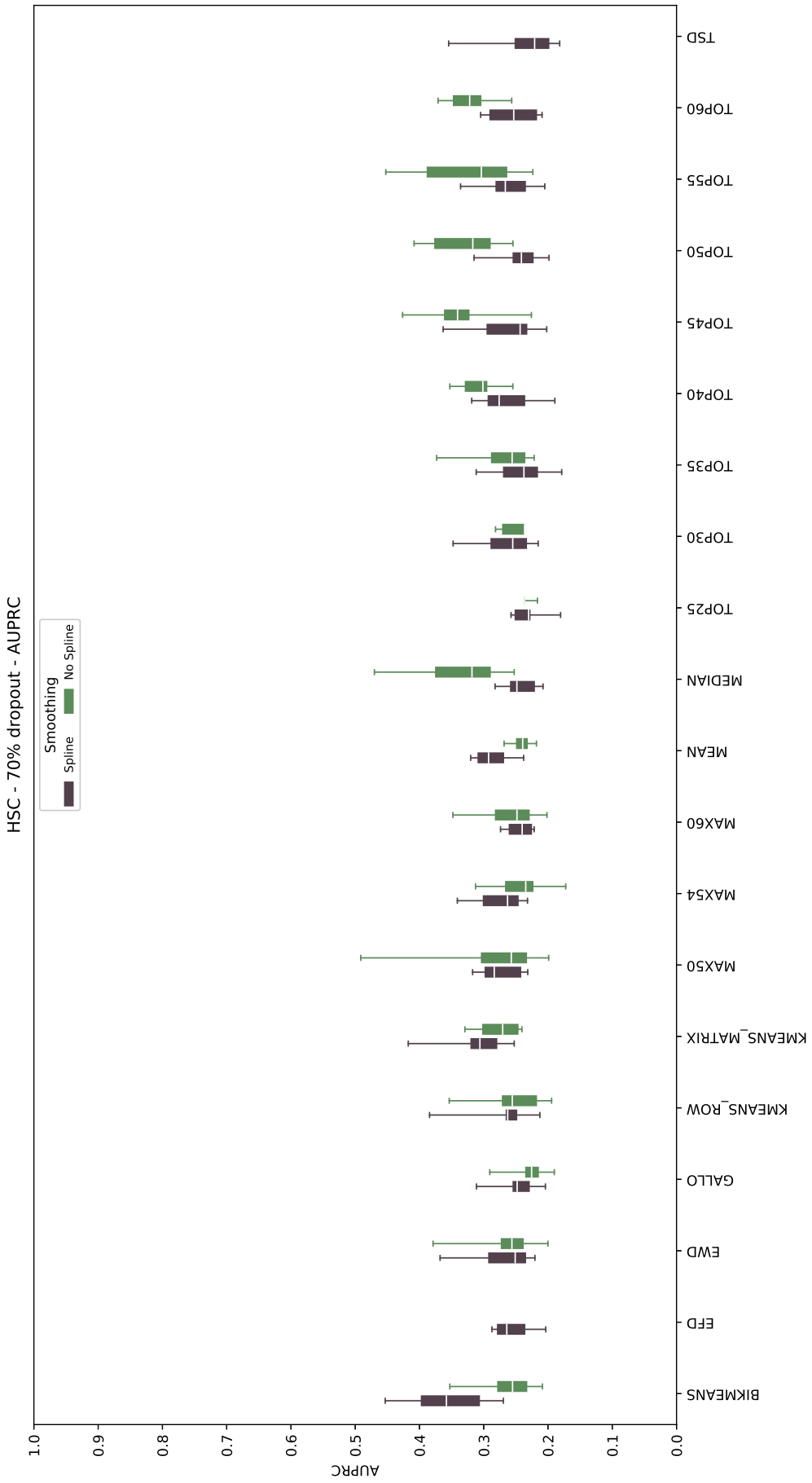

Figure 15. Results for problem HSC with and without Spline, considering AUPRC and 70% dropout.

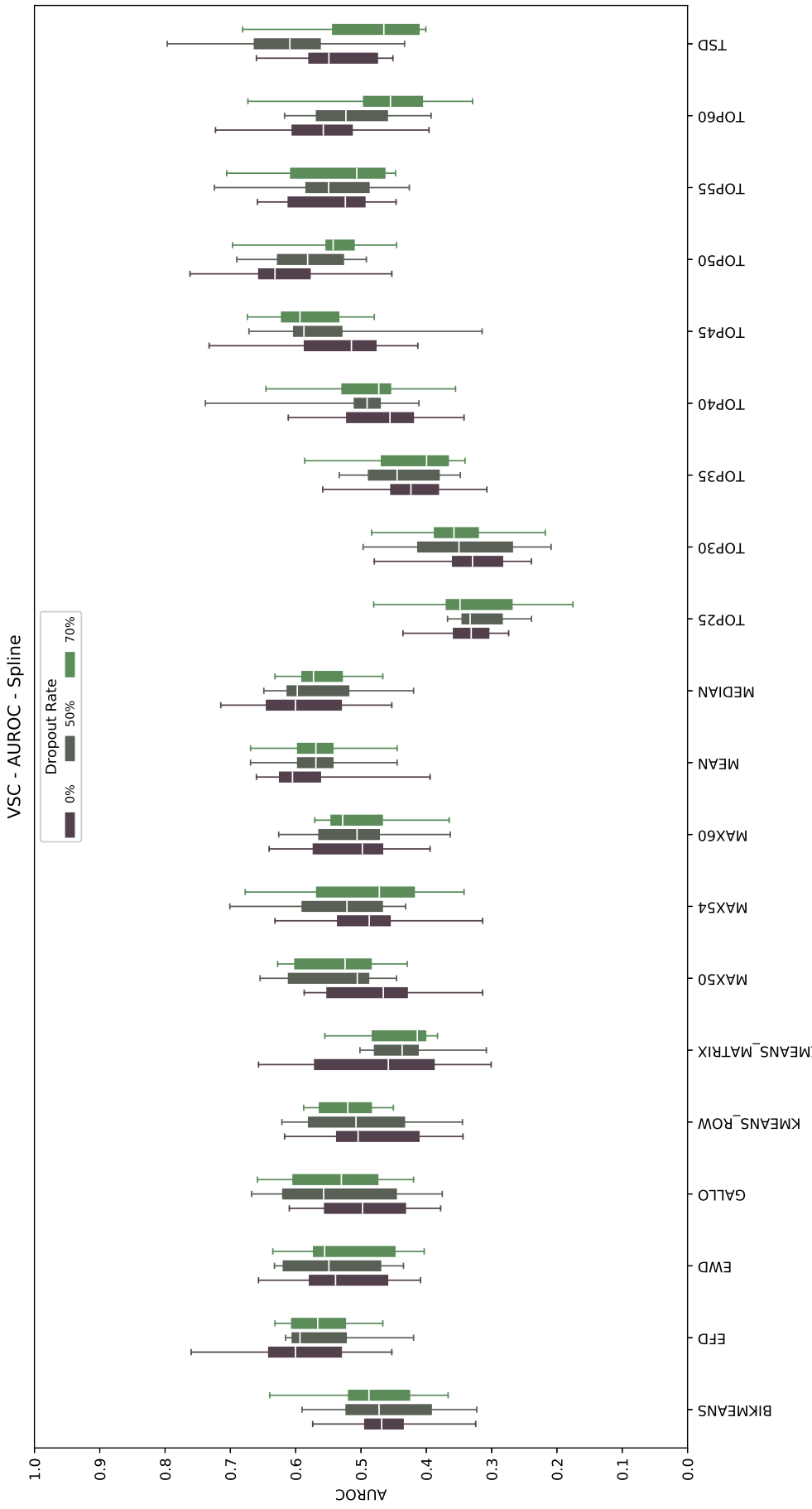

Figure 16. Results for problem VSC with Spline, considering AUROC.

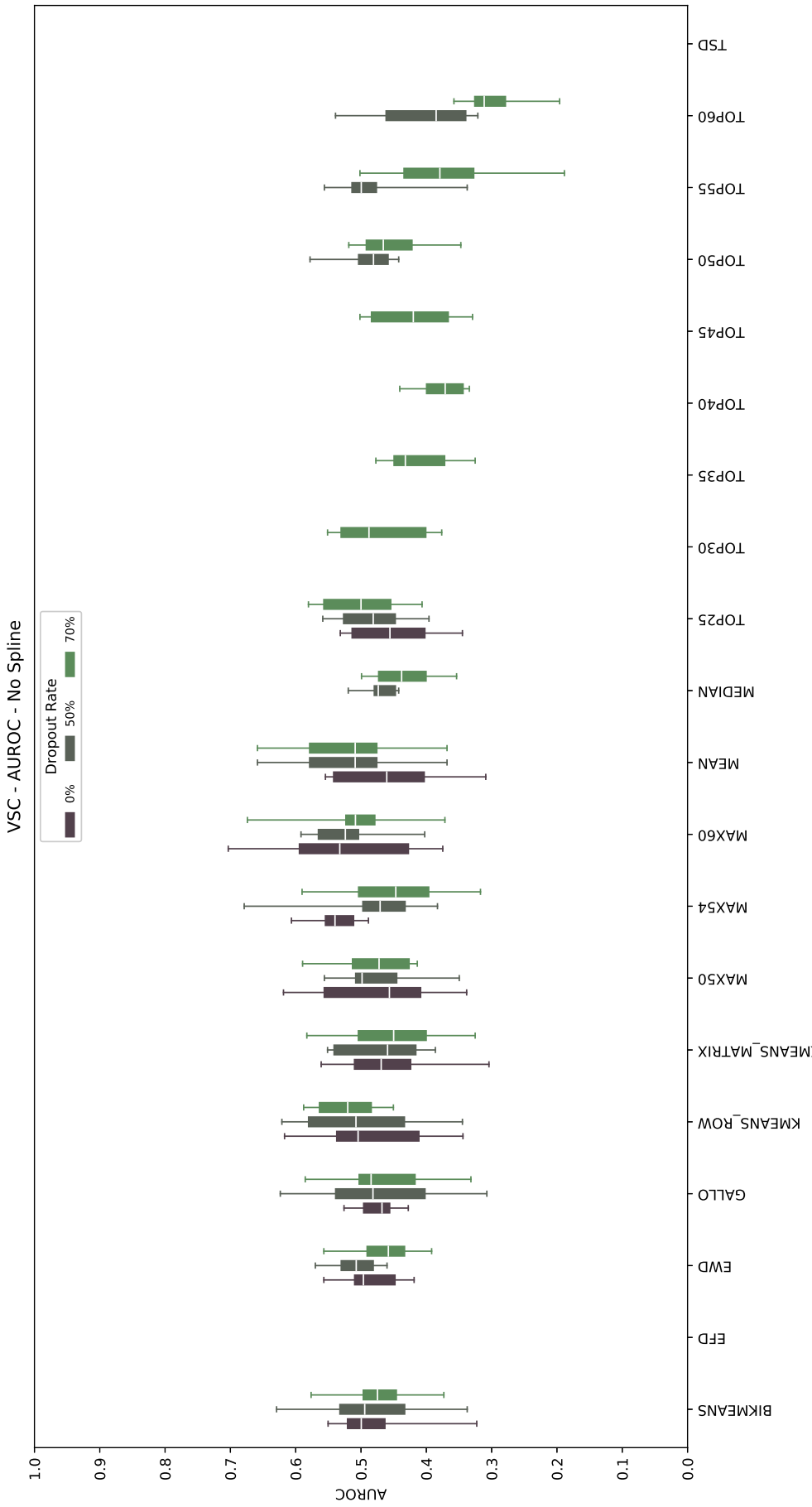

Figure 17. Results for problem VSC without Spline, considering AUROC.

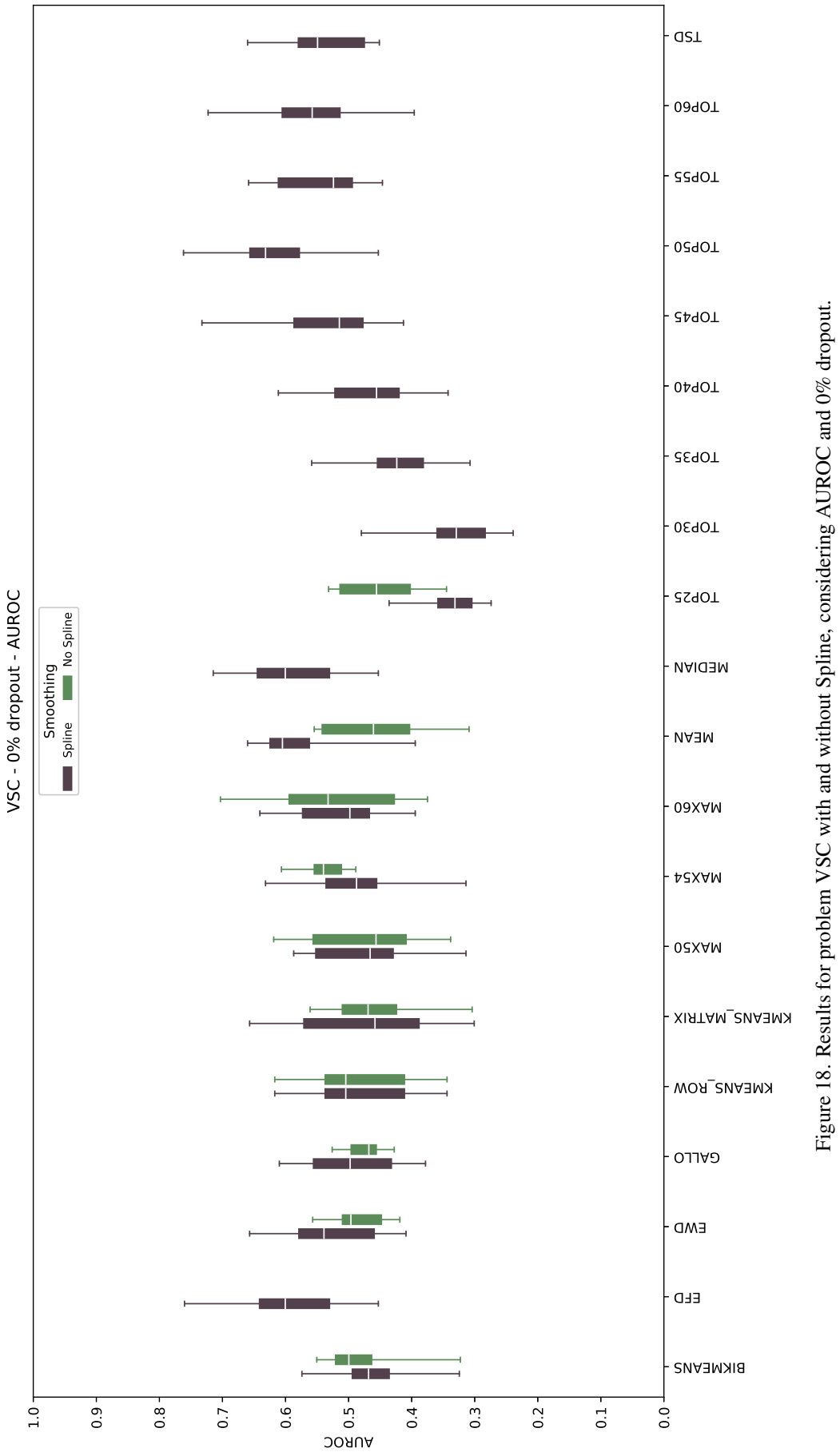

Figure 18. Results for problem VSC with and without Spline, considering AUROC and 0% dropout.

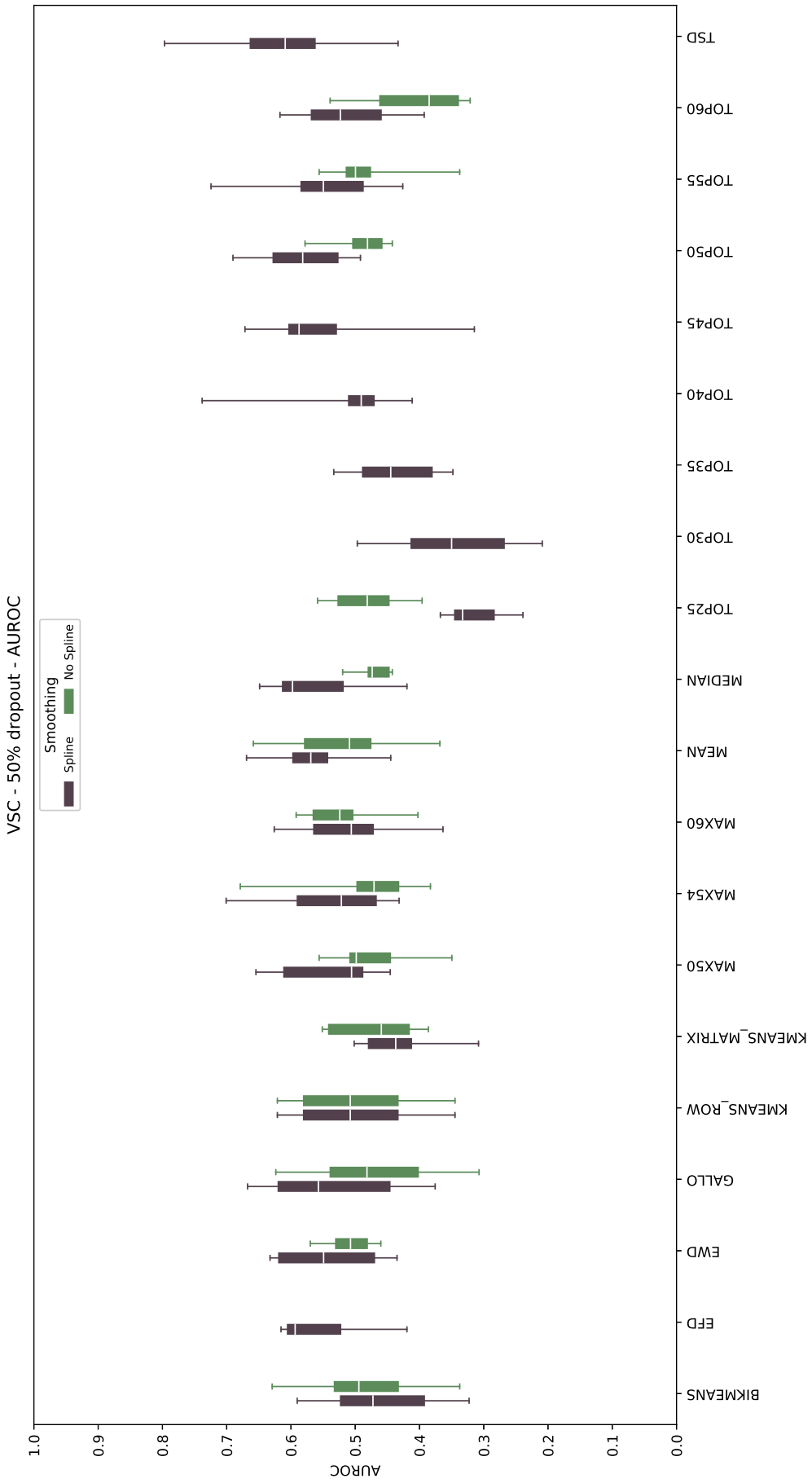

Figure 19. Results for problem VSC with and without Spline, considering AUROC and 50% dropout.

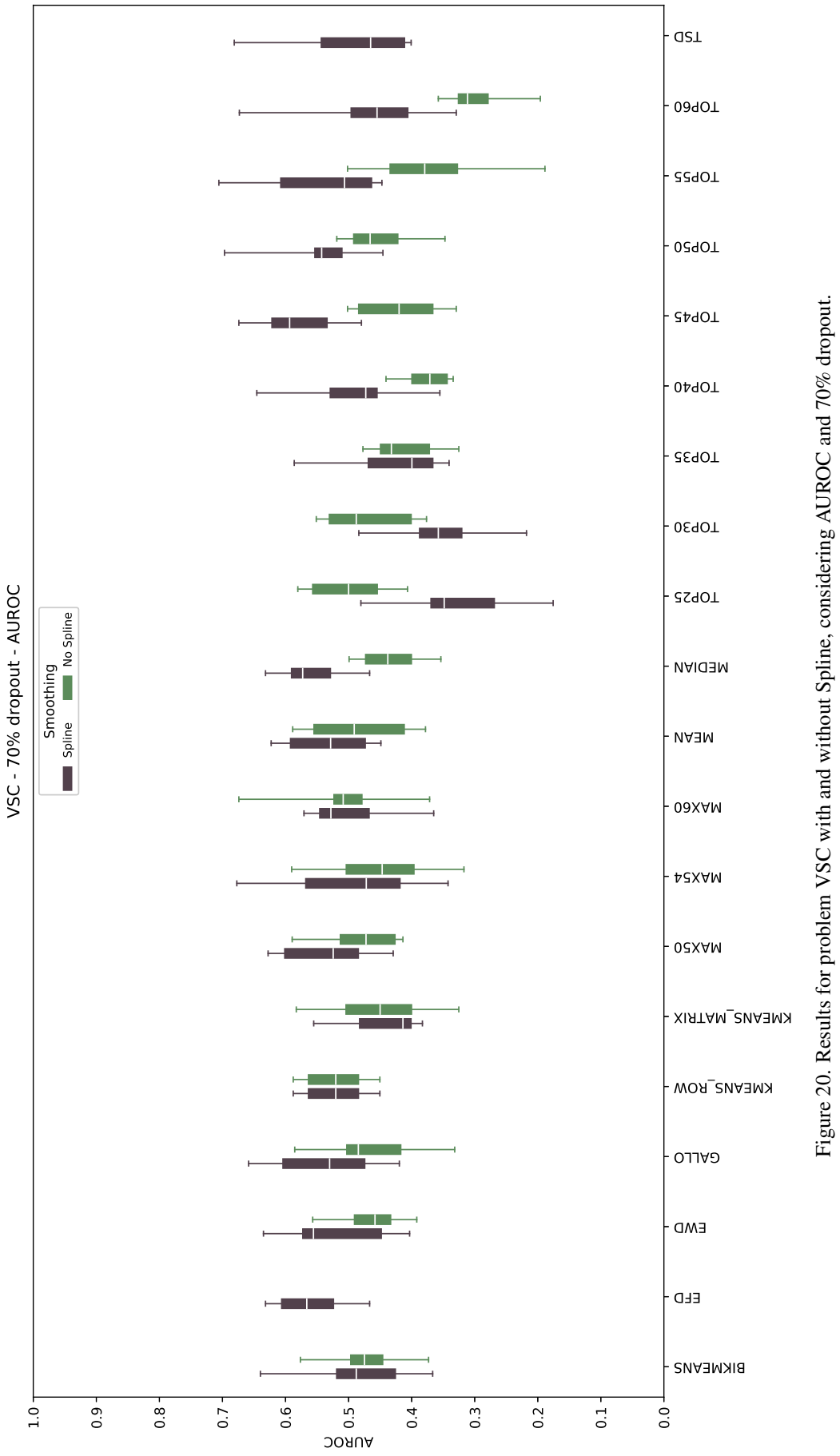

Figure 20. Results for problem VSC with and without Spline, considering AUROC and 70% dropout.

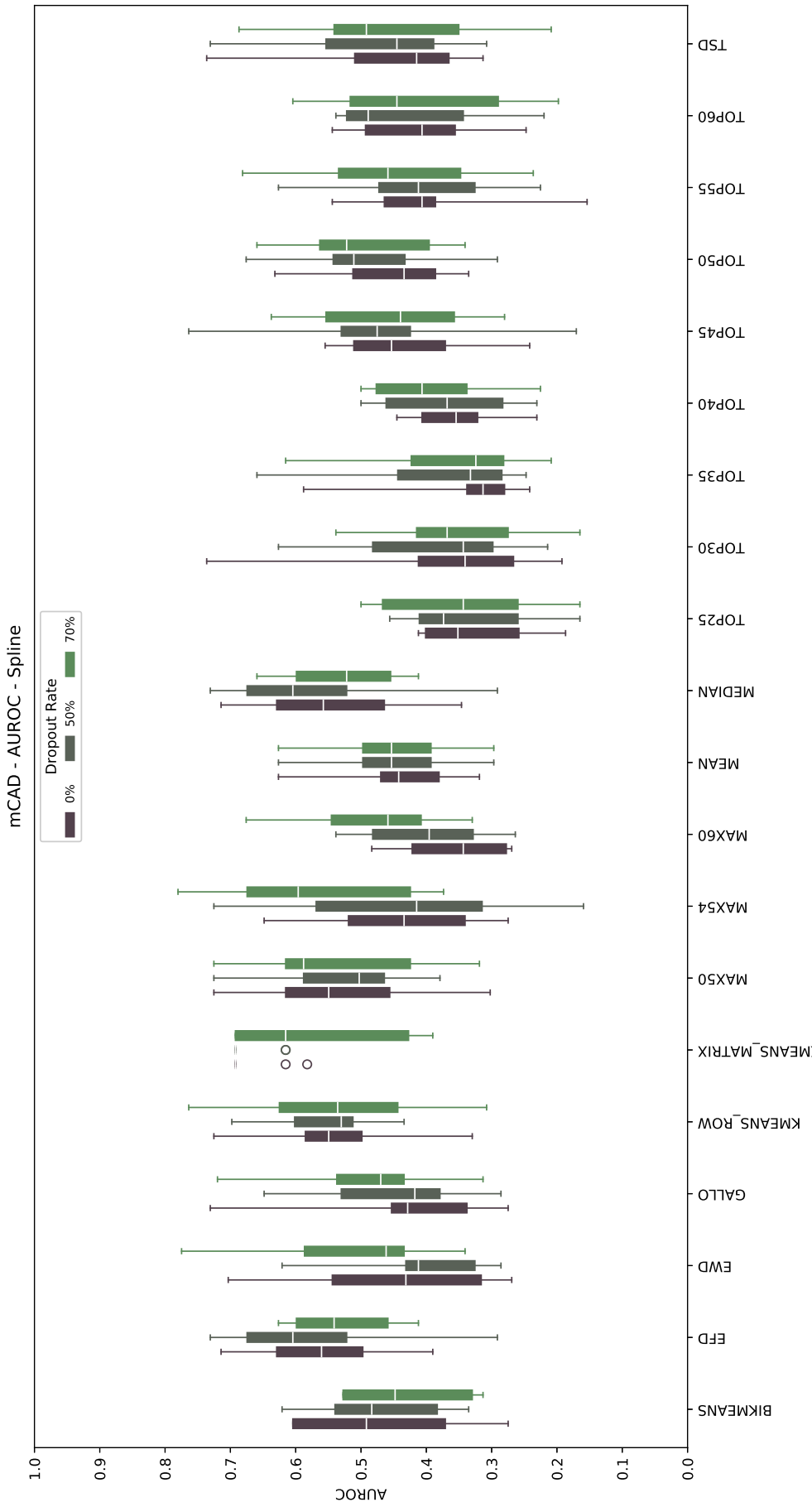

Figure 21. Results for problem mCAD with Spline, considering AUROC.

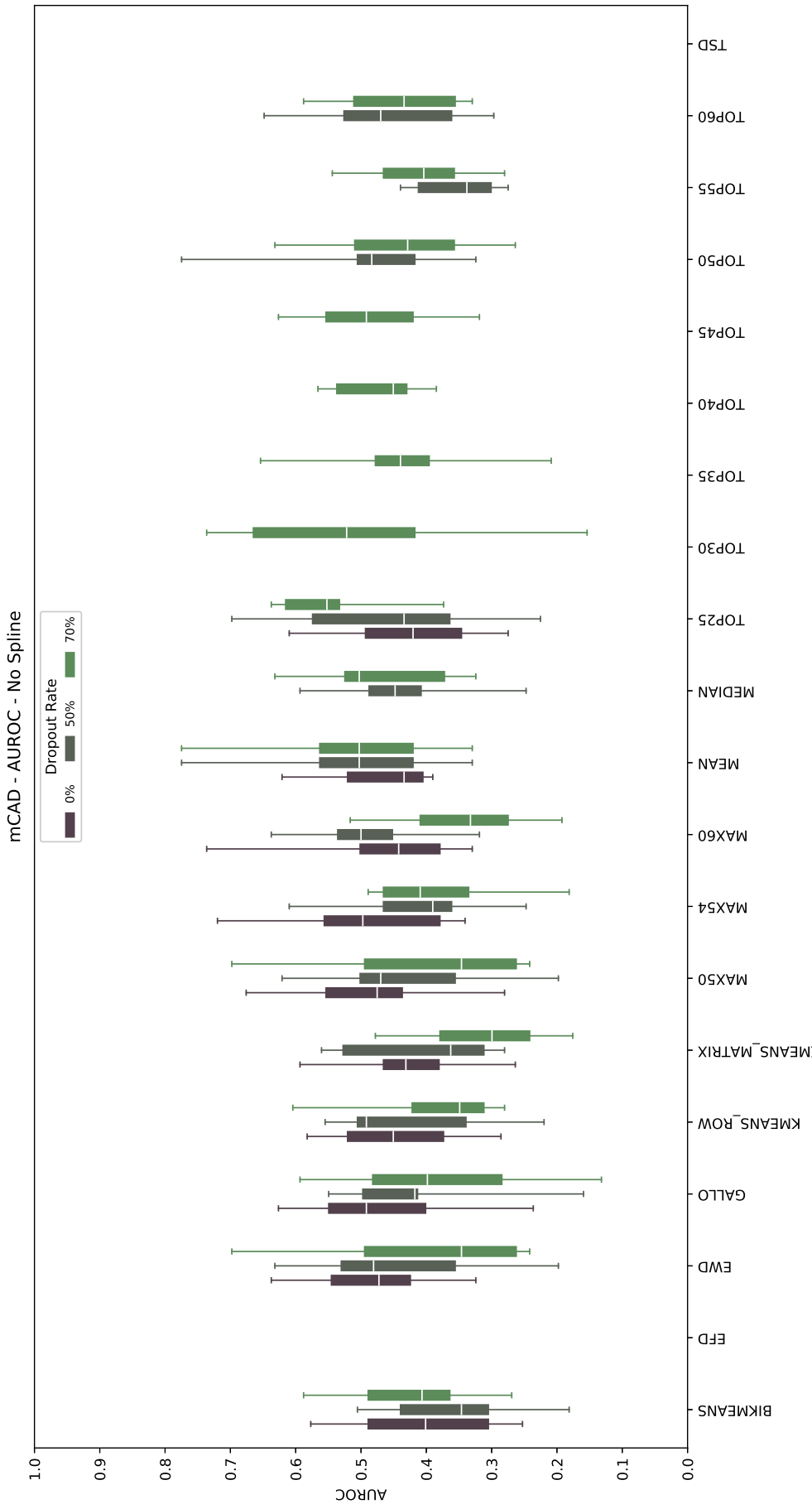

Figure 22. Results for problem mCAD without Spline, considering AUROC.

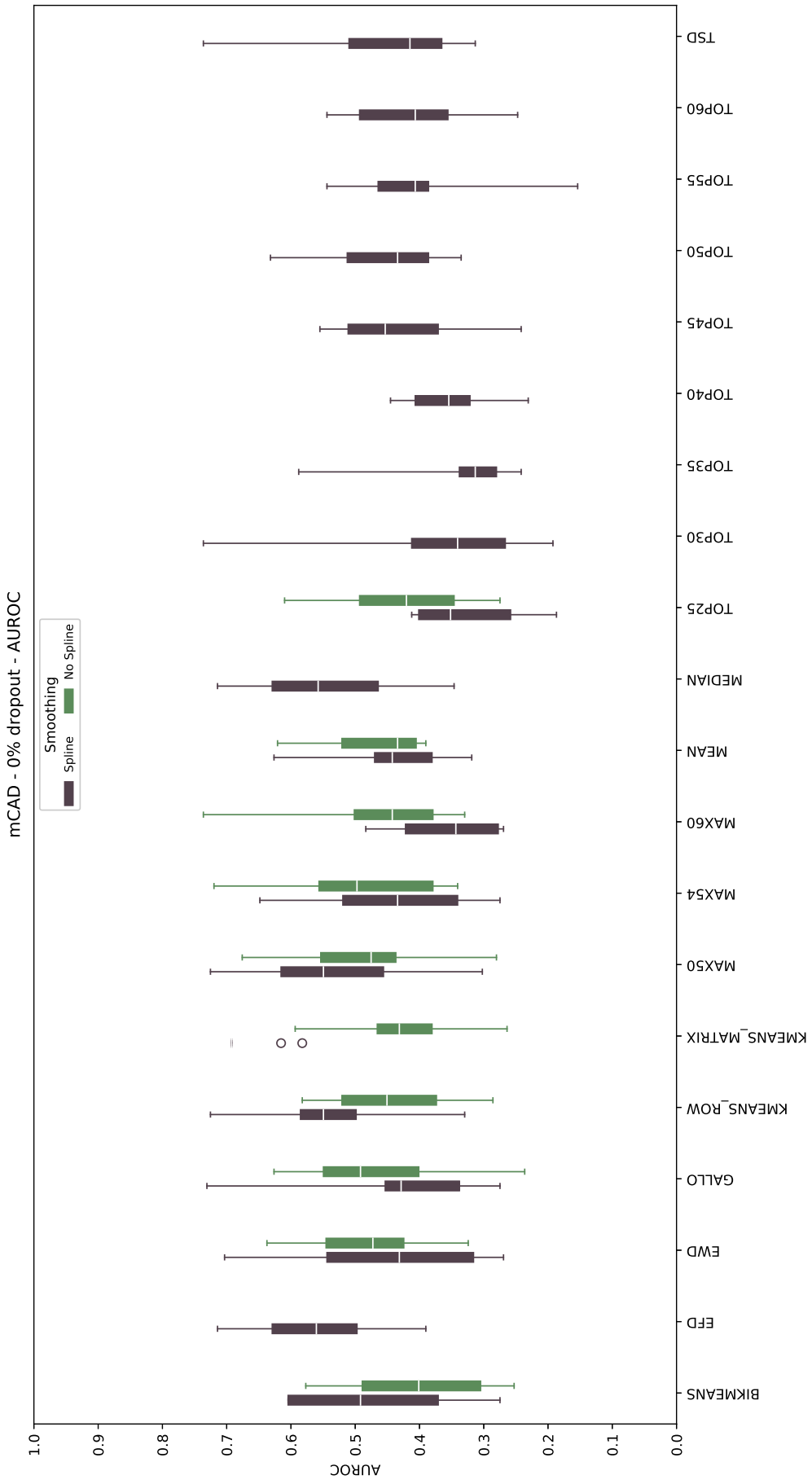

Figure 23. Results for problem mCAD with and without Spline, considering AUROC and 0% dropout.

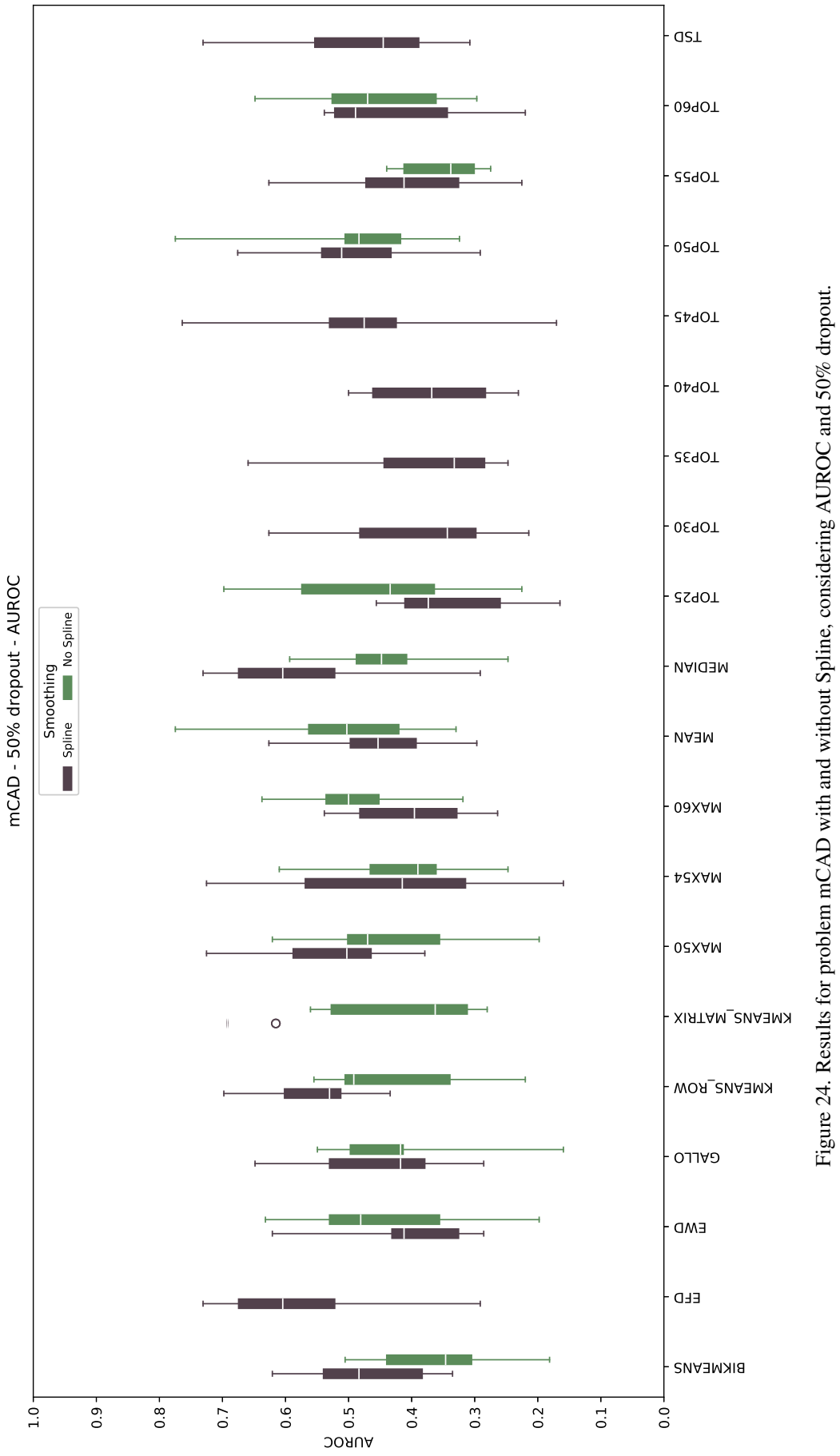

Figure 24. Results for problem mCAD with and without Spline, considering AUROC and 50% dropout.

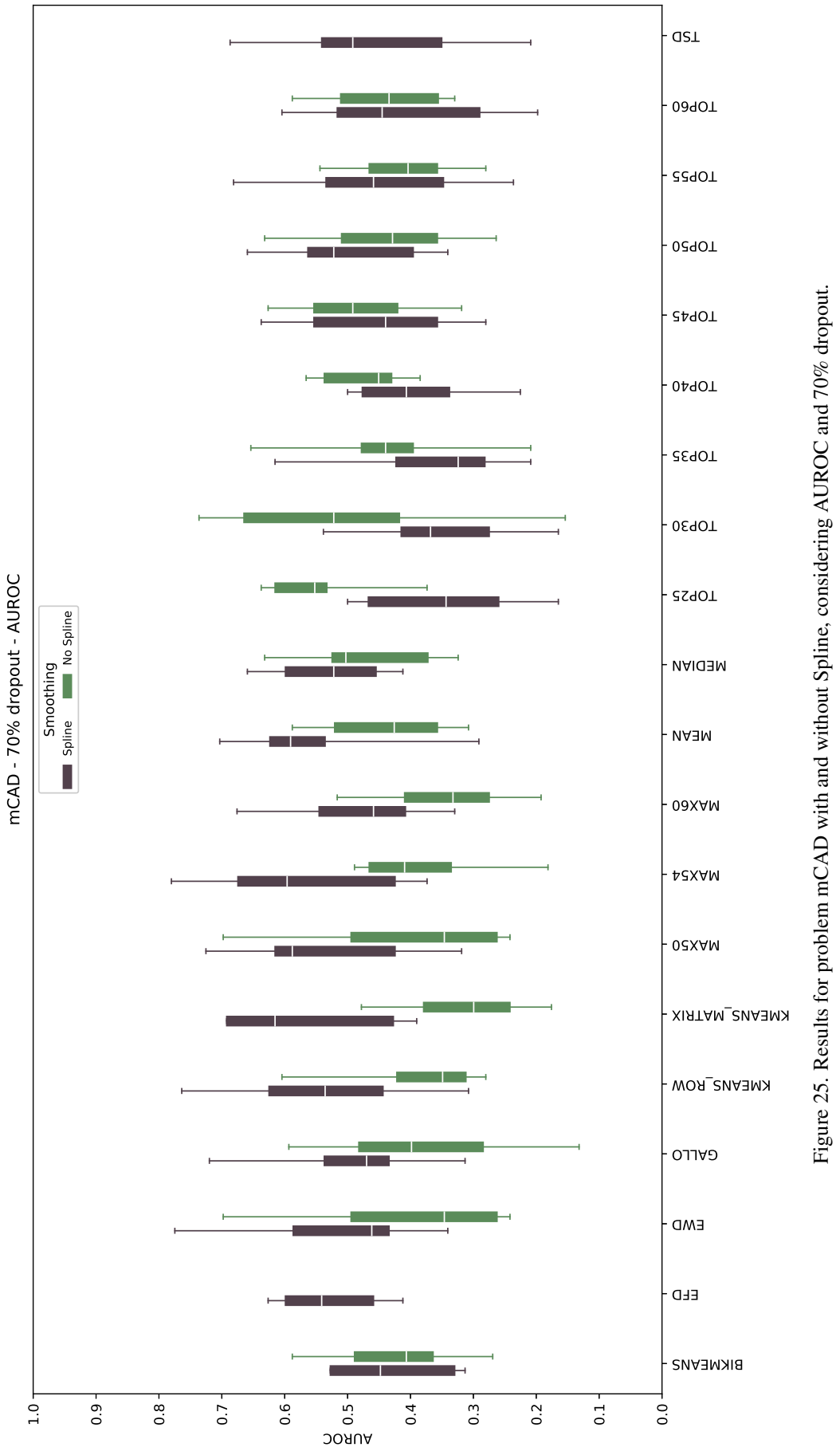

Figure 25. Results for problem mCAD with and without Spline, considering AUROC and 70% dropout.

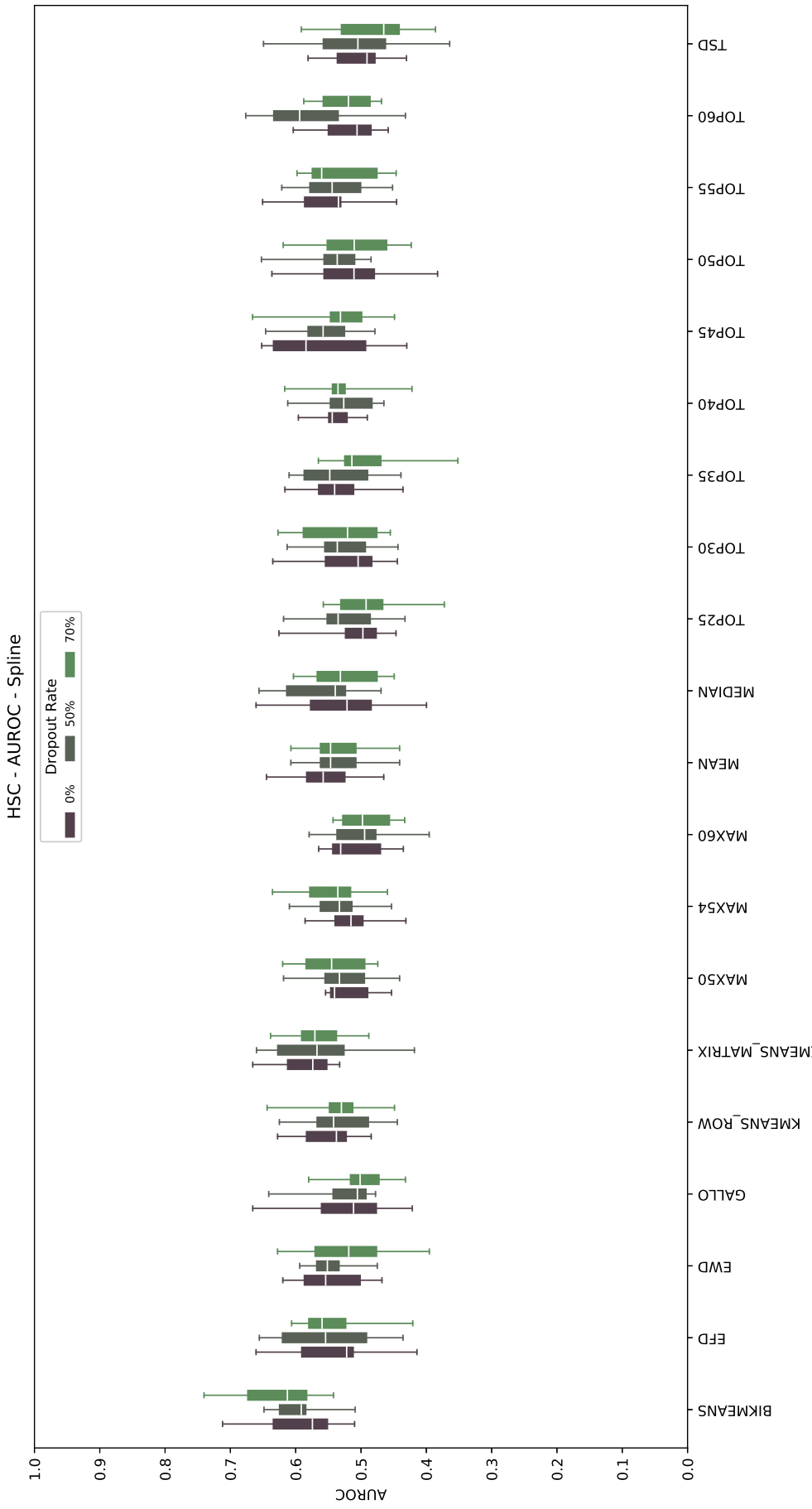

Figure 26. Results for problem HSC with Spline, considering AUROC.

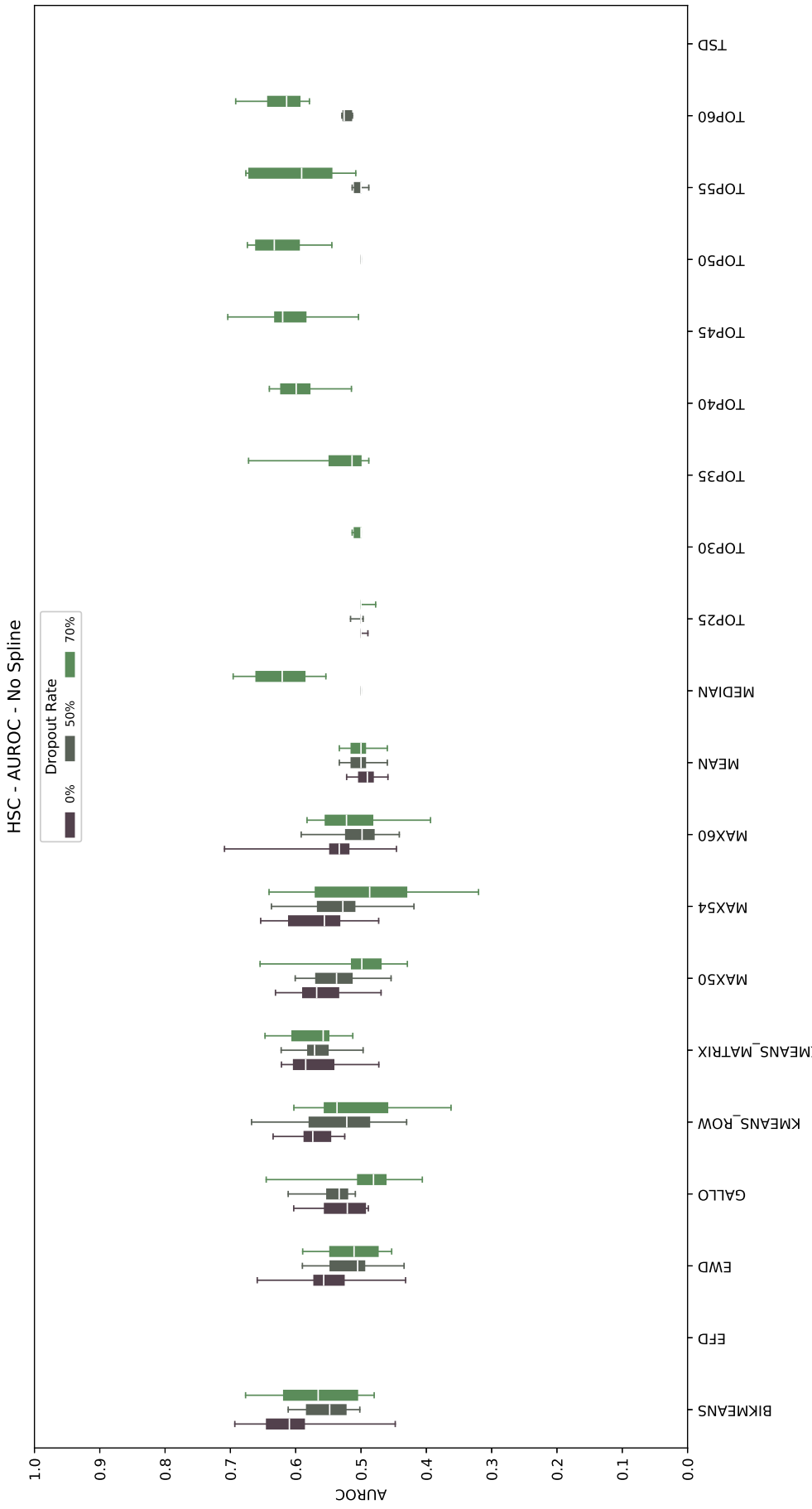

Figure 27. Results for problem HSC without Spline, considering AUROC.

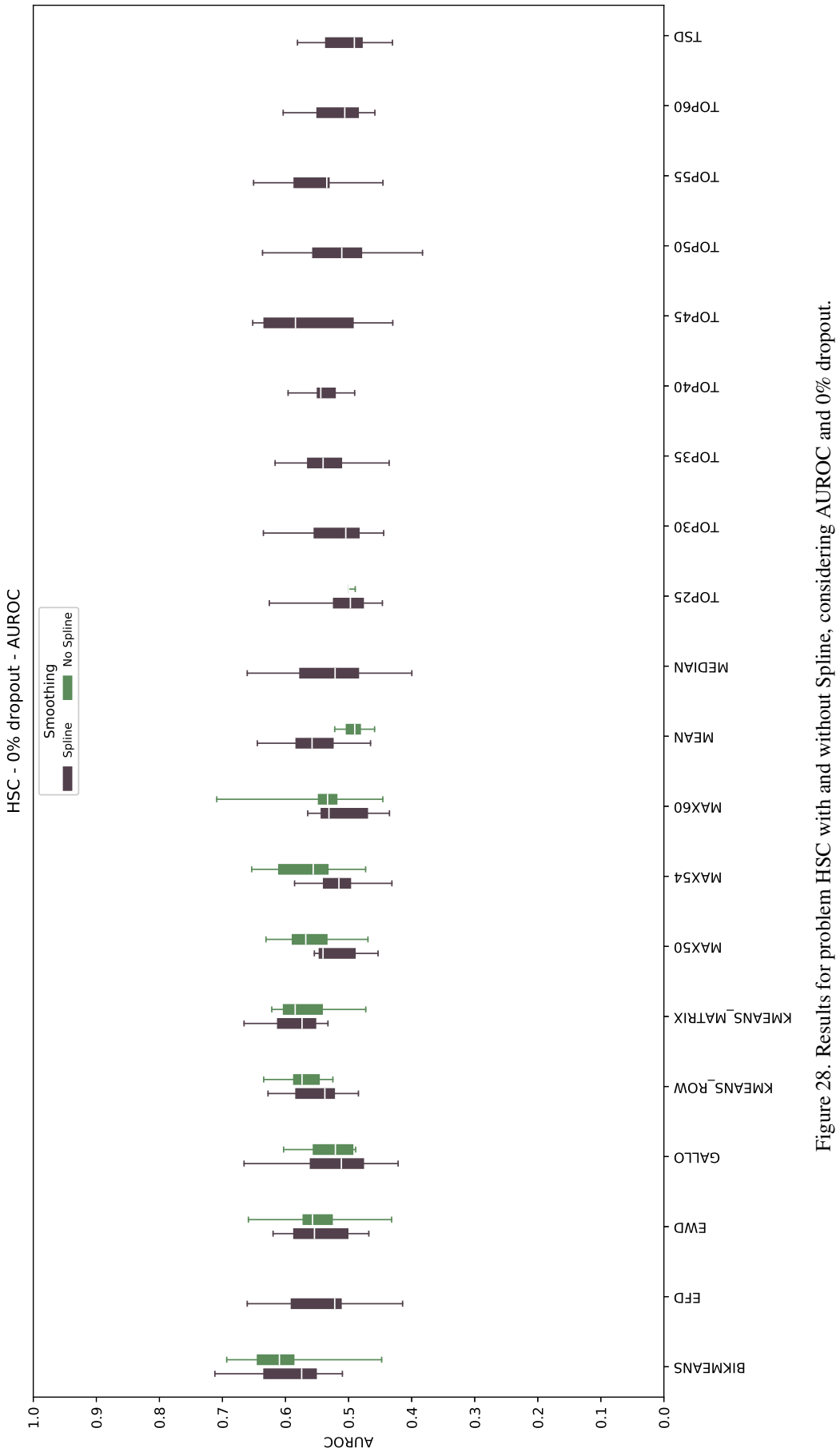

Figure 28. Results for problem HSC with and without Spline, considering AUROC and 0% dropout.

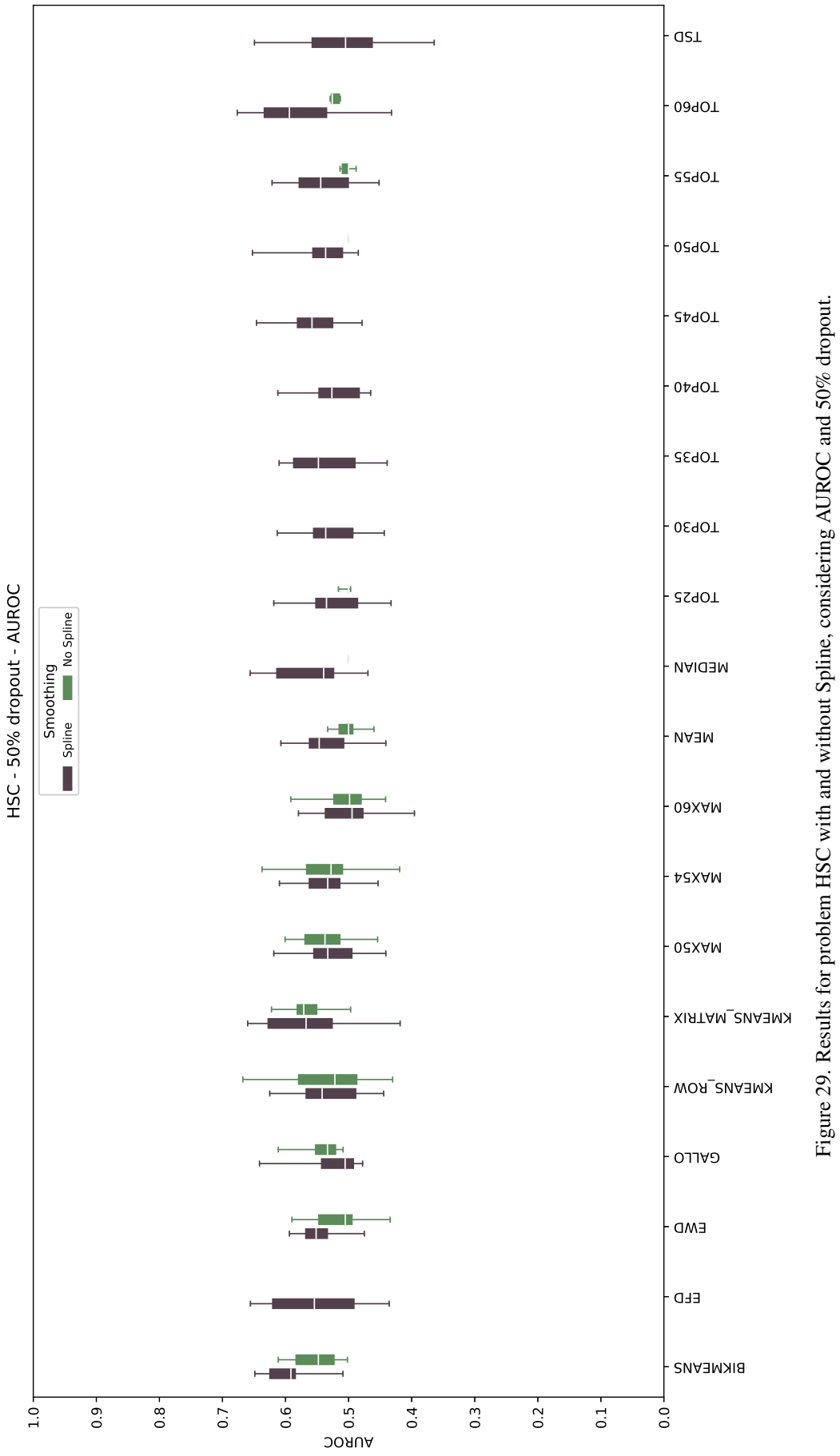

Figure 29. Results for problem HSC with and without Spline, considering AUROC and 50% dropout.

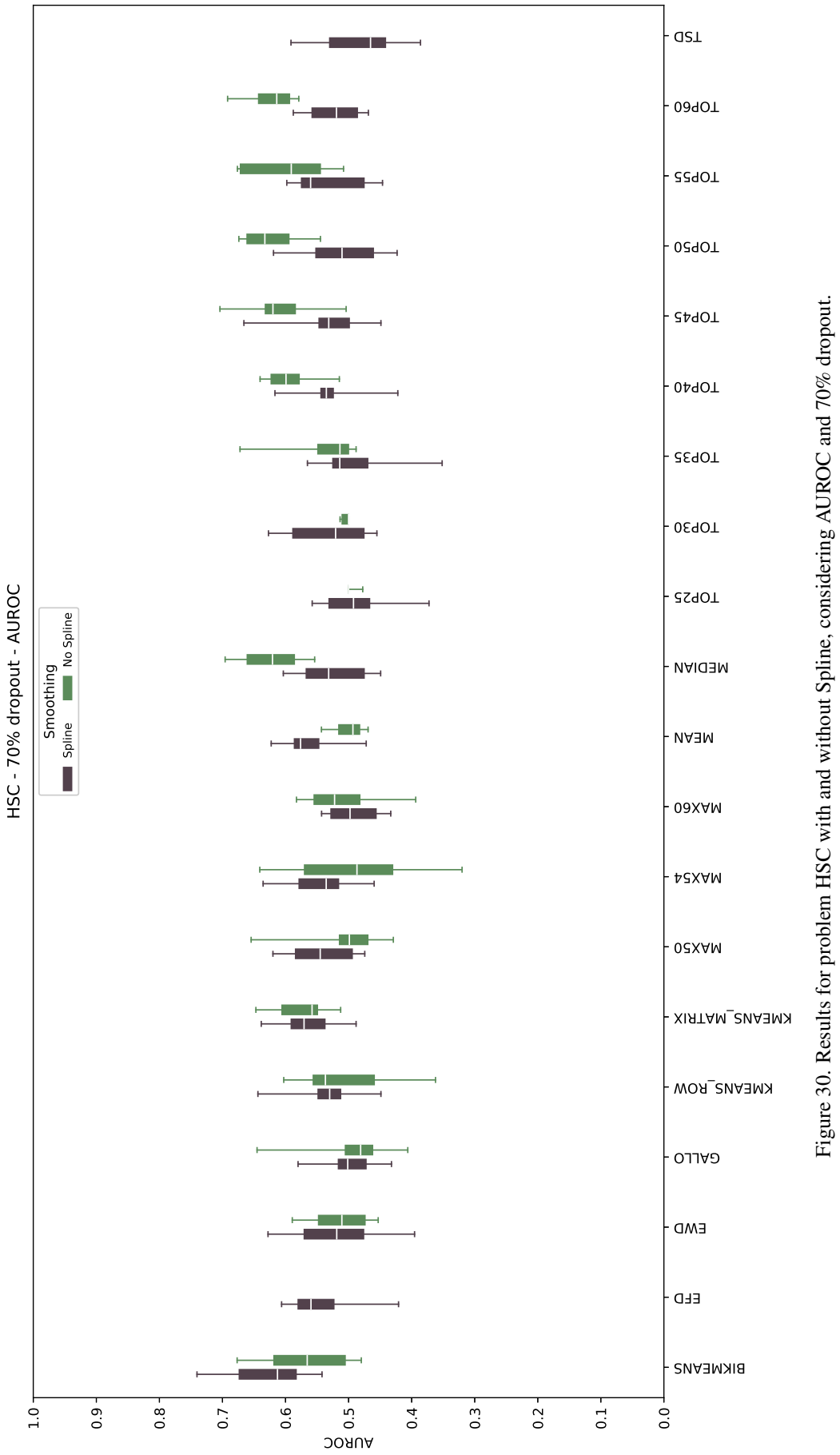

Figure 30. Results for problem HSC with and without Spline, considering AUROC and 70% dropout.

3 Parameter Analysis

In this seciton we perform the parameter analysis of Max -X%Max and Top%X approaches. The Max -X%Max parameter are analyzed in [50%, 54%, 60%], once 54% is the reference parameter [1]. When considering Top%X, we performed experiments considering [25%, ..., 60%].

3.1 Max -X%Max

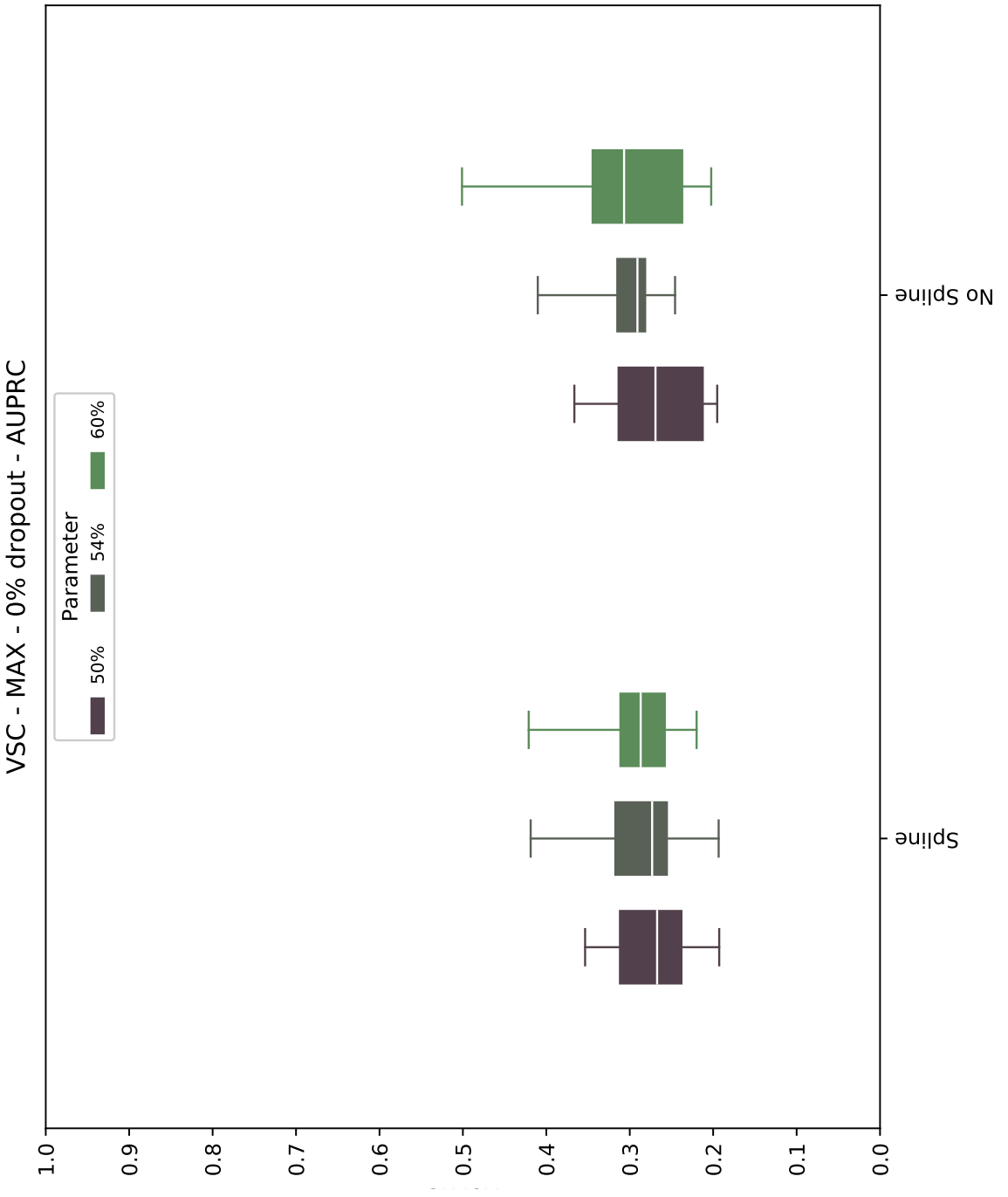

Figure 31. Results for problem VSC with and without Spline, considering AUPRC and 0% dropout with parameters in range [50%, 54%, 60%].

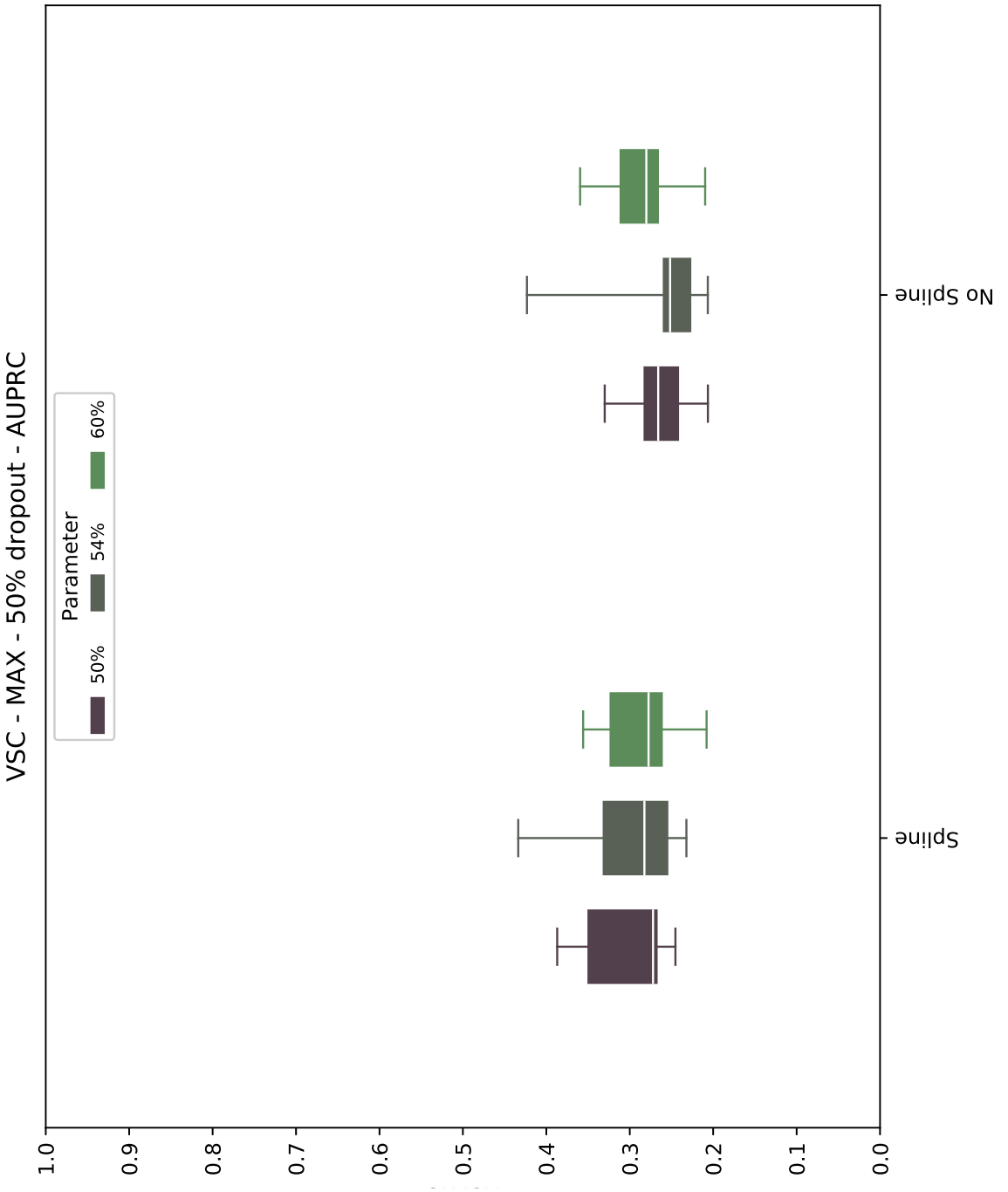

Figure 32. Results for problem VSC with and without Spline, considering AUPRC and 50% dropout with parameters in range [50%, 54%, 60%].

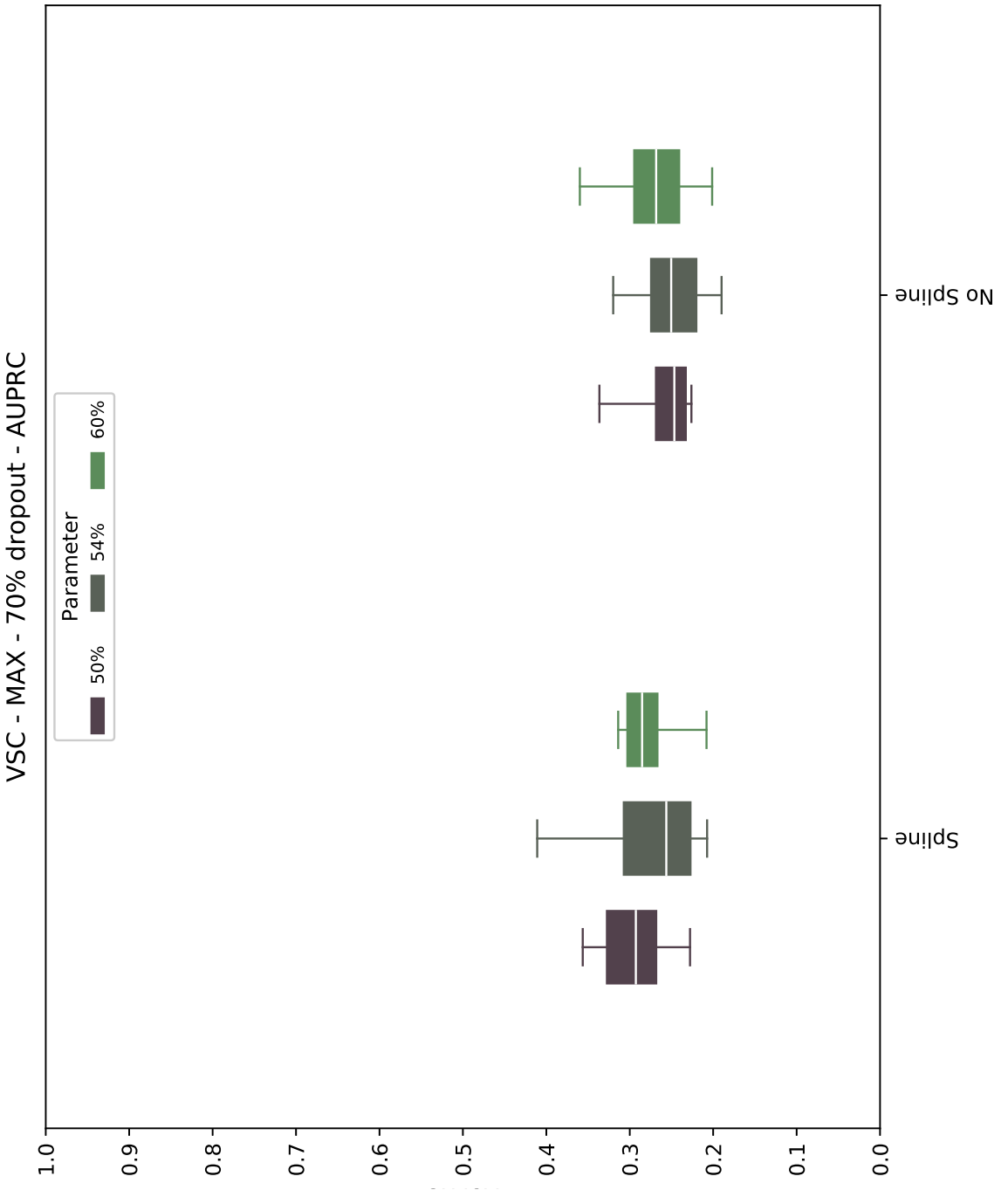

Figure 33. Results for problem VSC with and without Spline, considering AUPRC and 70% dropout with parameters in range [50%, 54%, 60%].

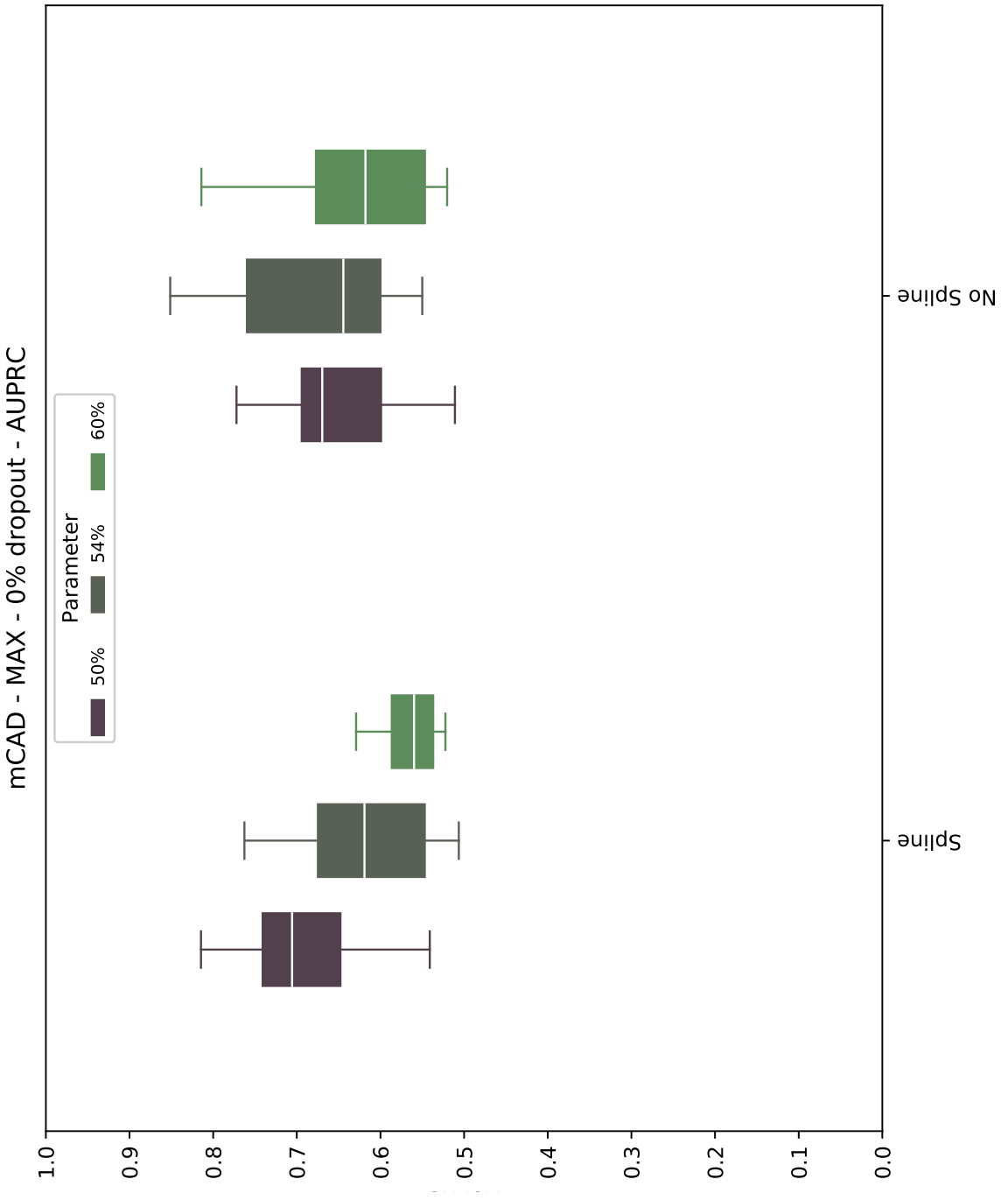

Figure 34. Results for problem mCAD with and without Spline, considering AUPRC and 0% dropout with parameters in range [50%, 54%, 60%].

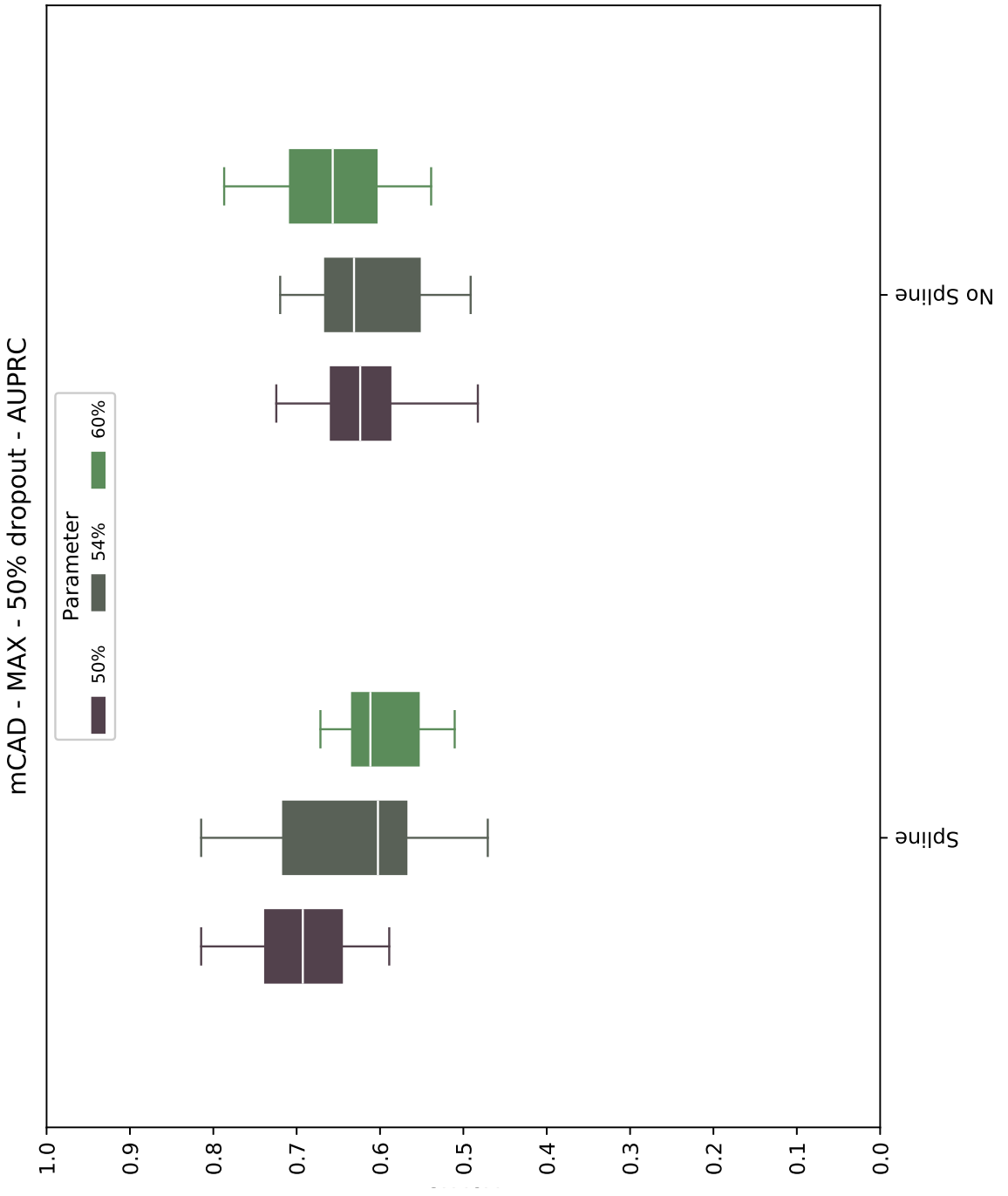

Figure 35. Results for problem mCAD with and without Spline, considering AUPRC and 50% dropout with parameters in range [50%, 54%, 60%].

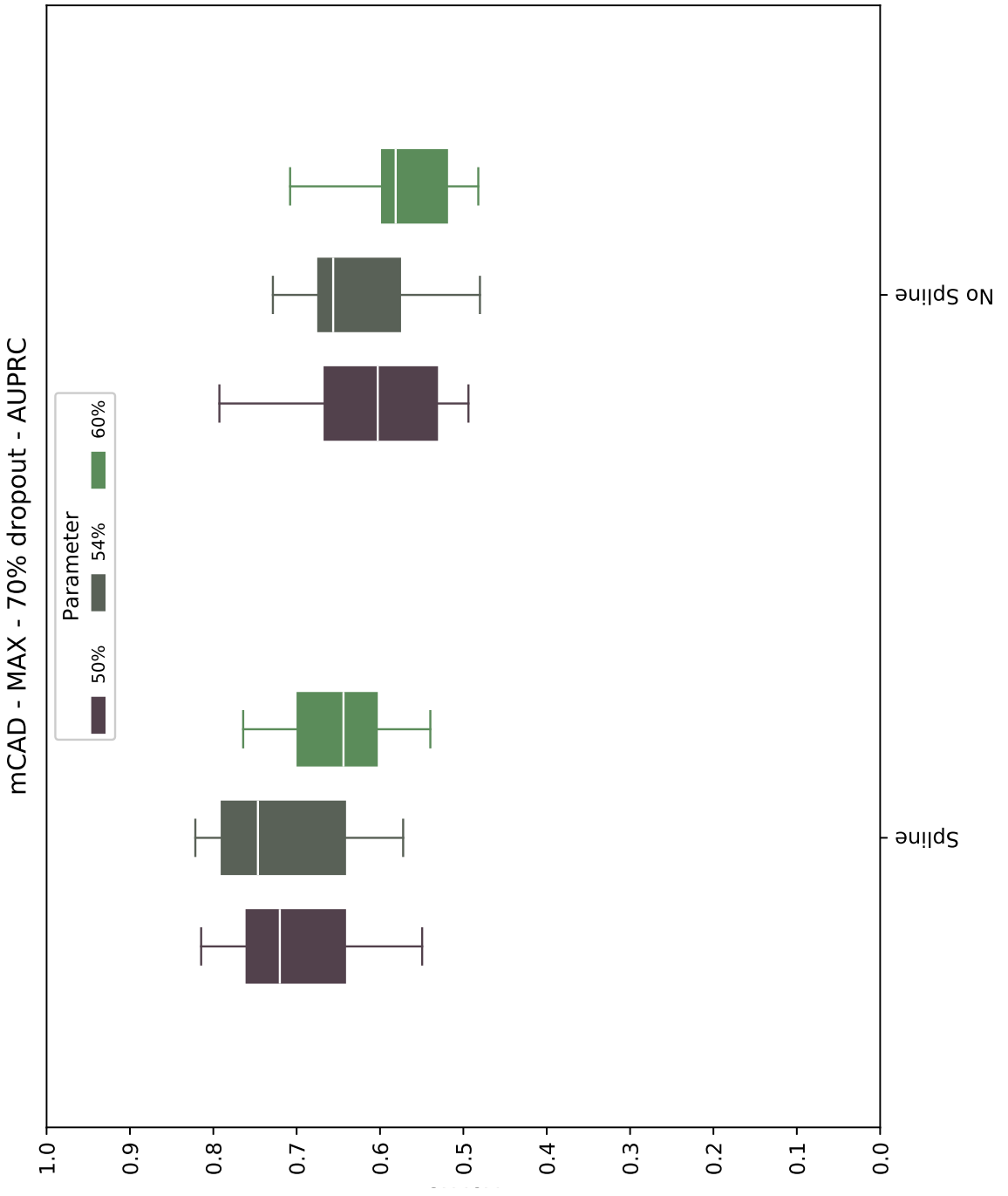

Figure 36. Results for problem mCAD with and without Spline, considering AUPRC and 70% dropout with parameters in range [50%, 54%, 60%].

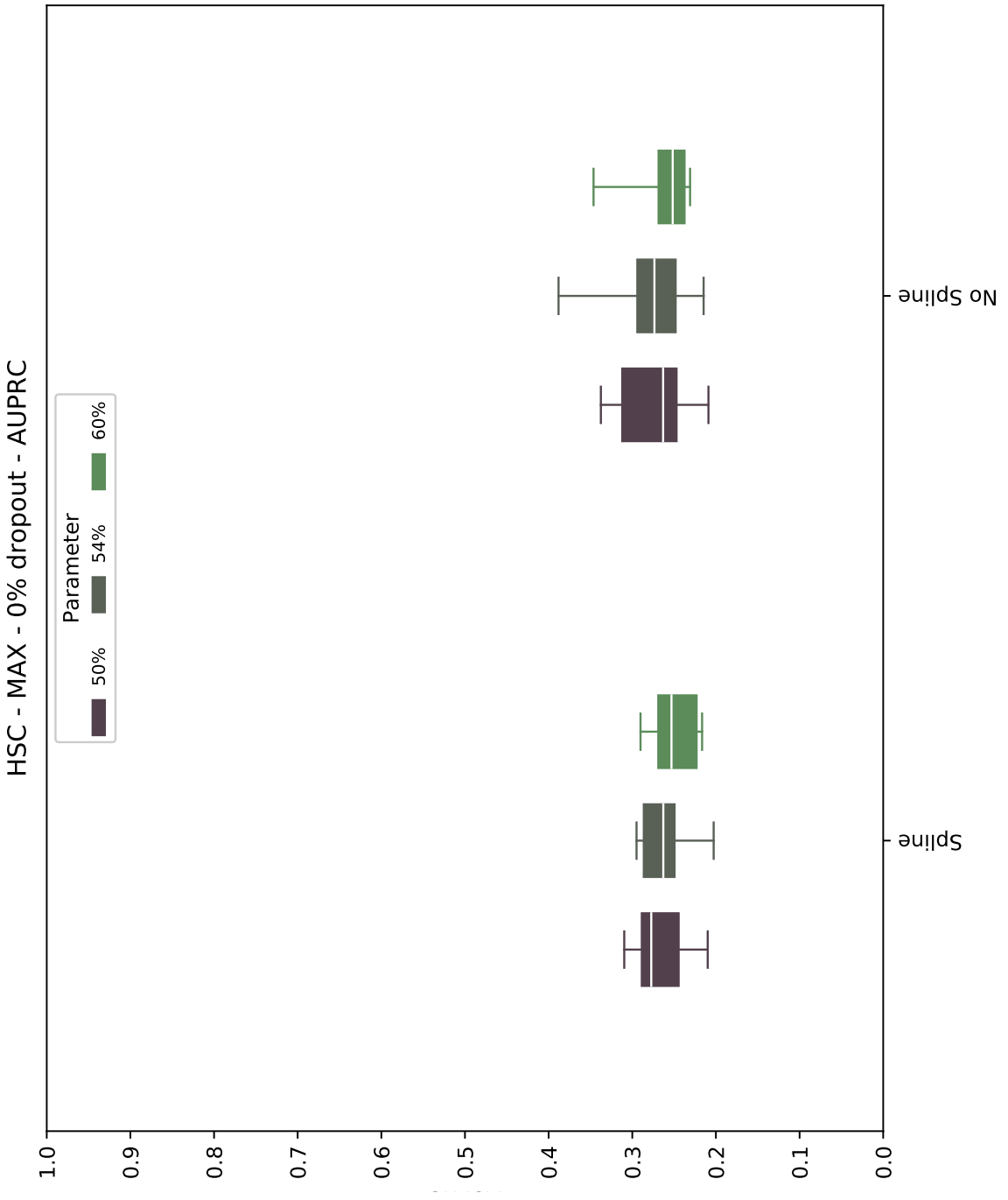

Figure 37. Results for problem HSC with and without Spline, considering AUPRC and 0% dropout with parameters in range [50%, 54%, 60%].

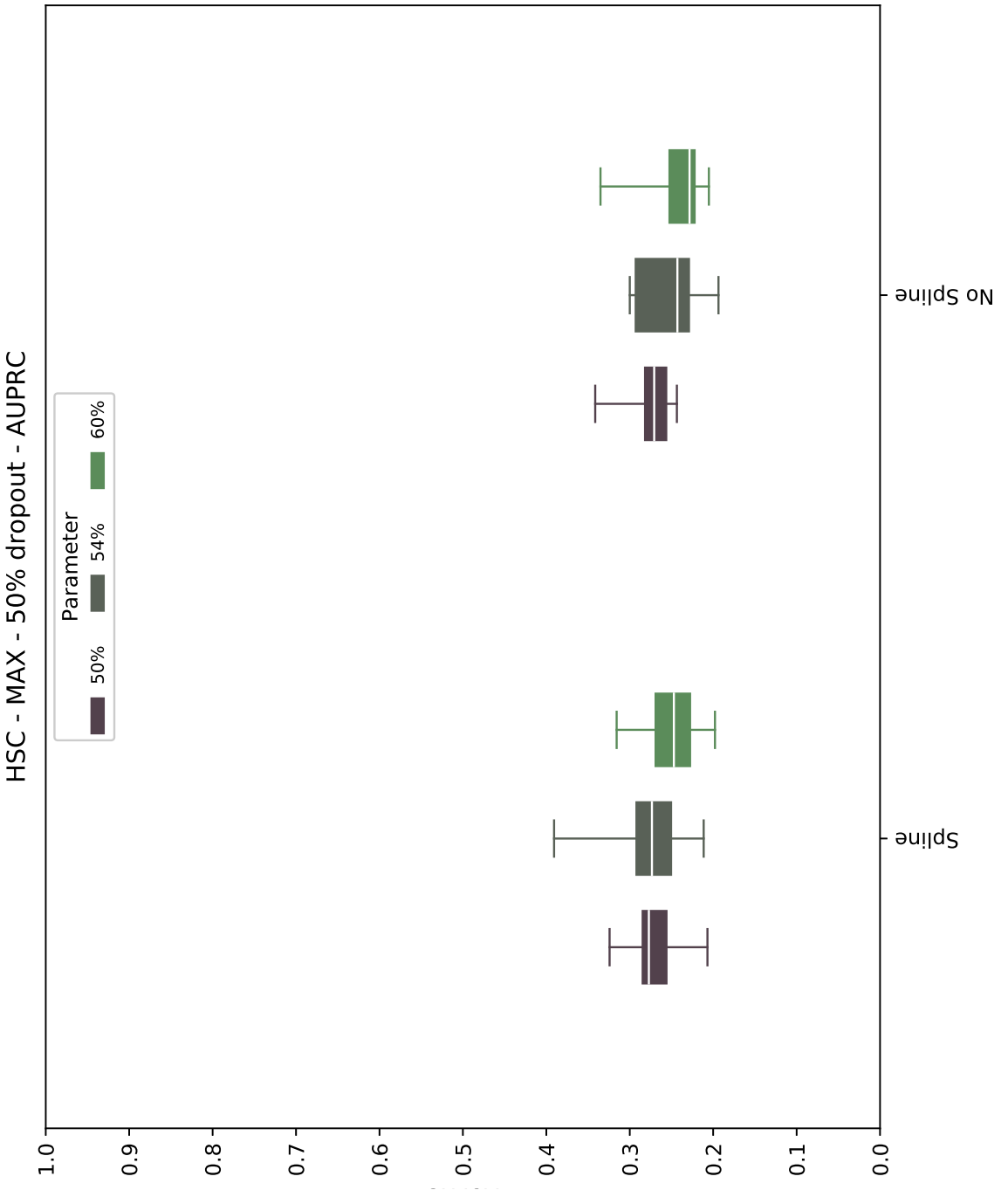

Figure 38. Results for problem HSC with and without Spline, considering AUPRC and 50% dropout with parameters in range [50%, 54%, 60%].

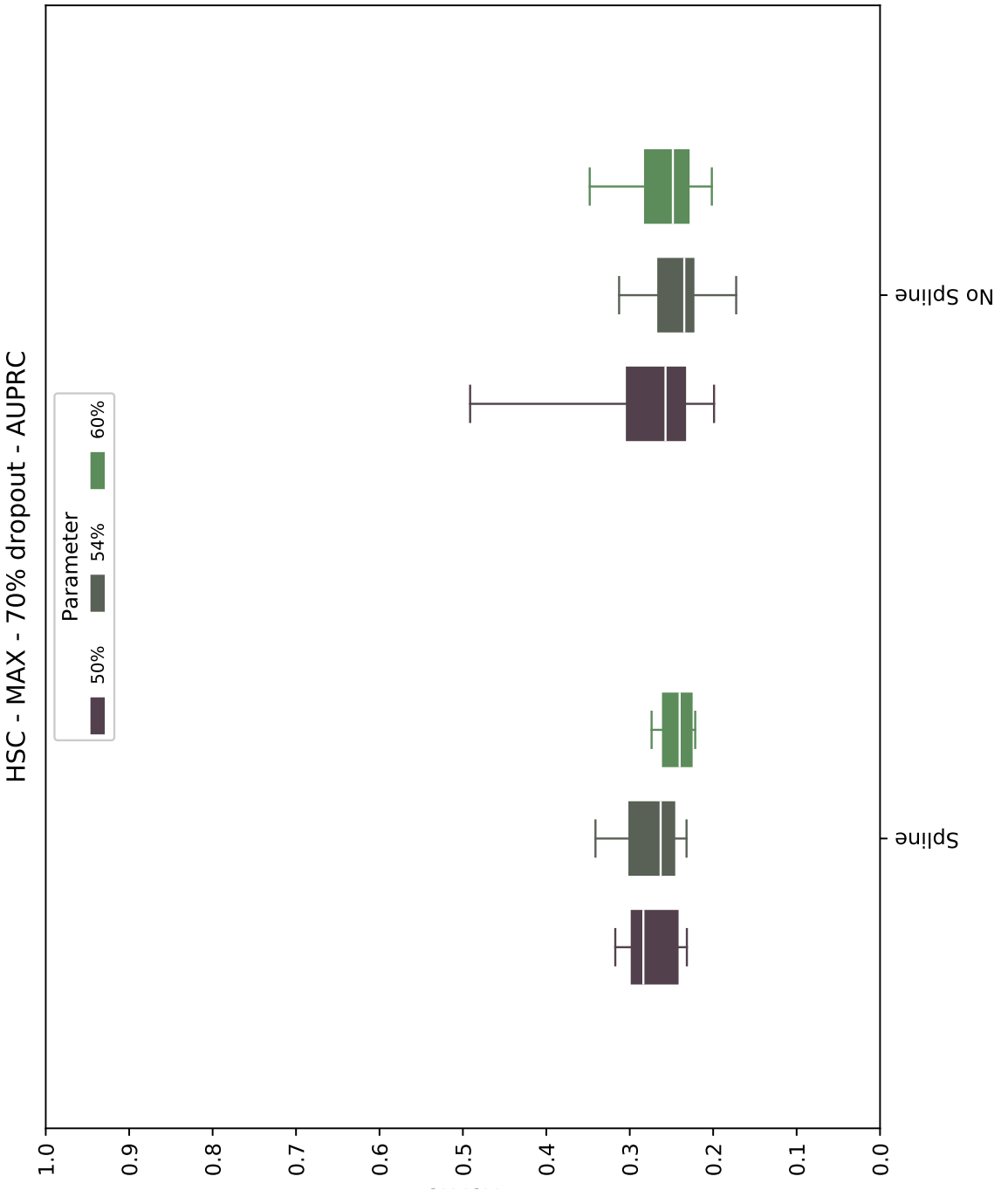

Figure 39. Results for problem HSC with and without Spline, considering AUPRC and 70% dropout with parameters in range [50%, 54%, 60%].

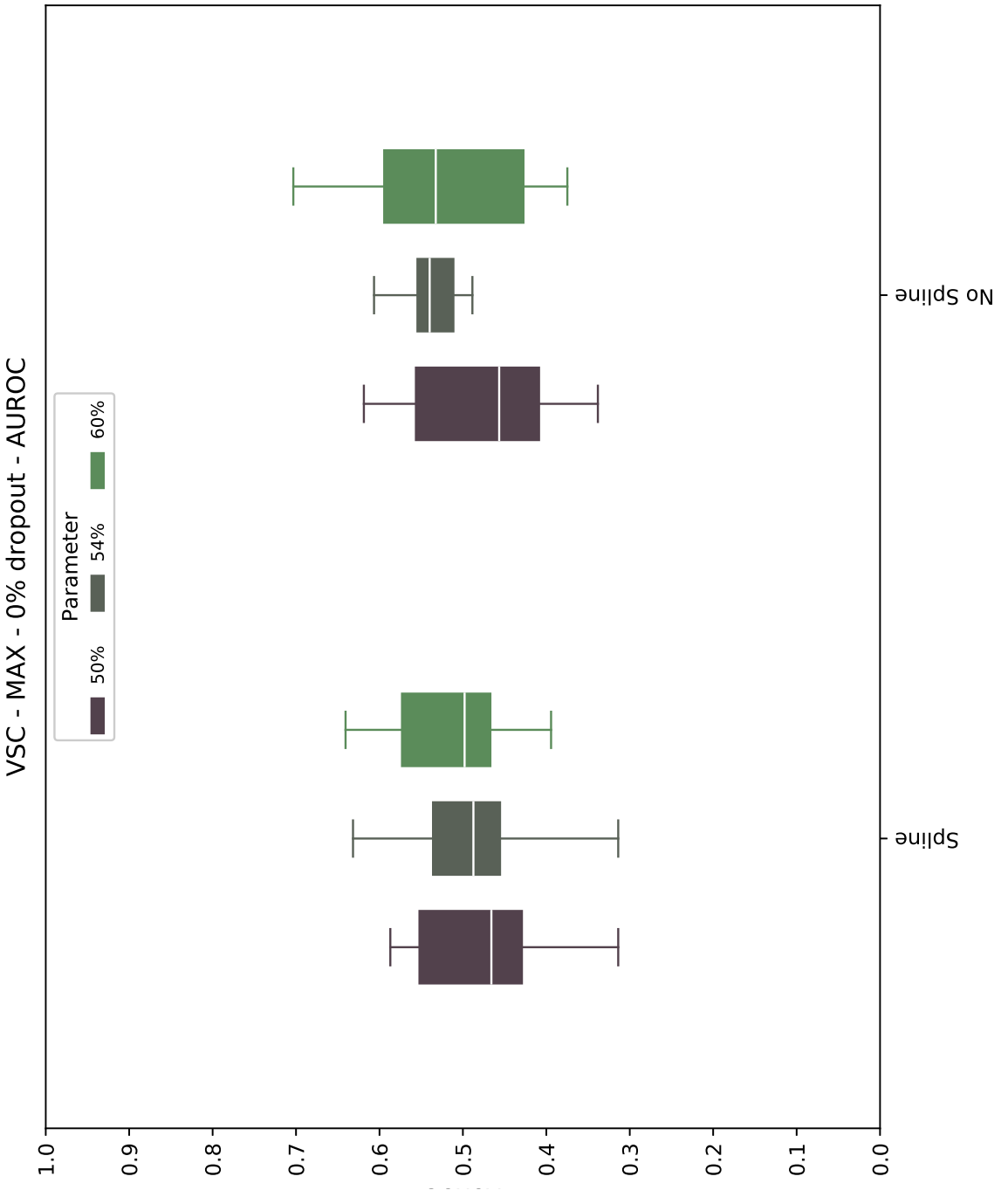

Figure 40. Results for problem VSC with and without Spline, considering AUROC and 0% dropout with parameters in range [50%, 54%, 60%].

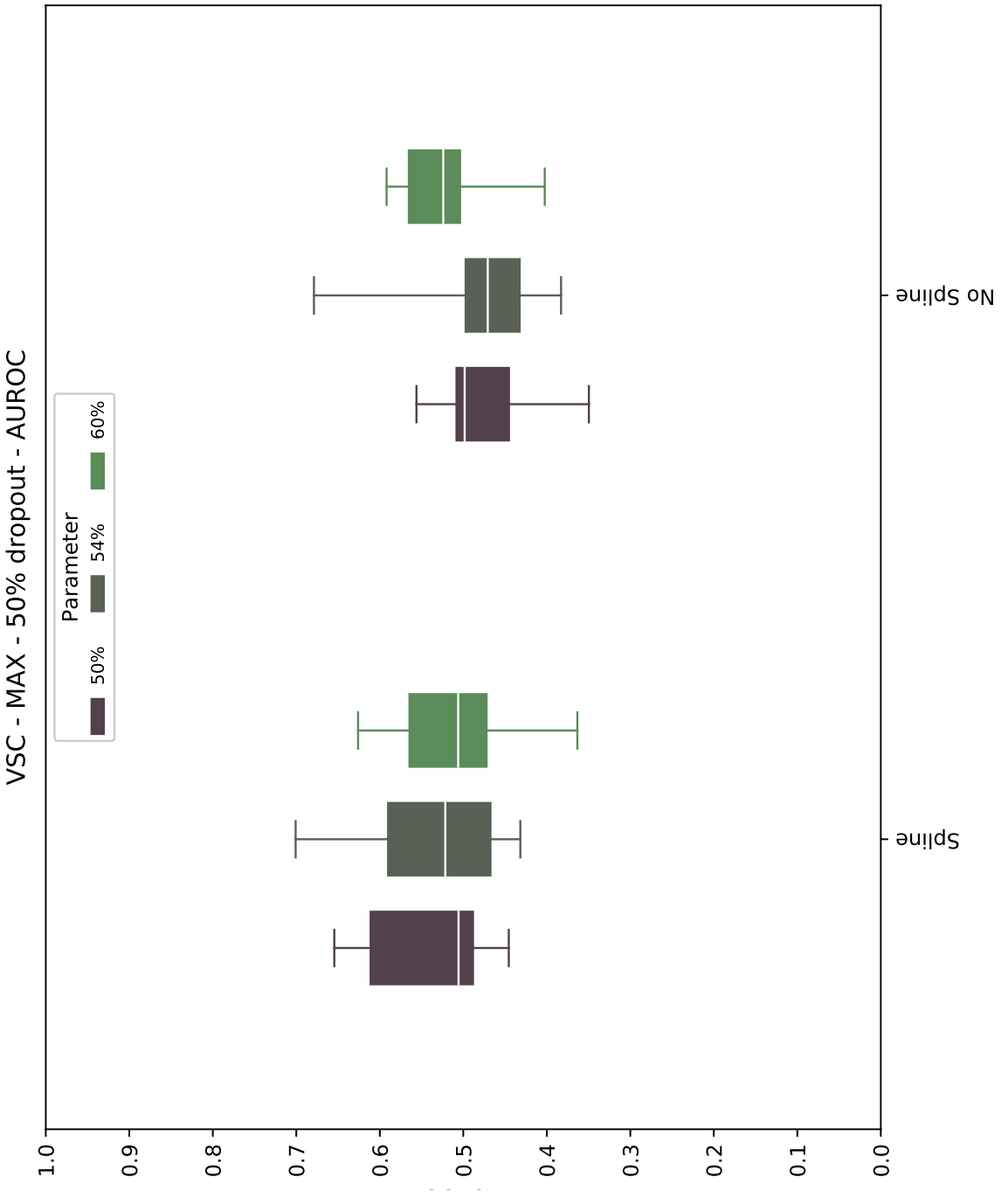

Figure 41. Results for problem VSC with and without Spline, considering AUROC and 50% dropout with parameters in range [50%, 54%, 60%].

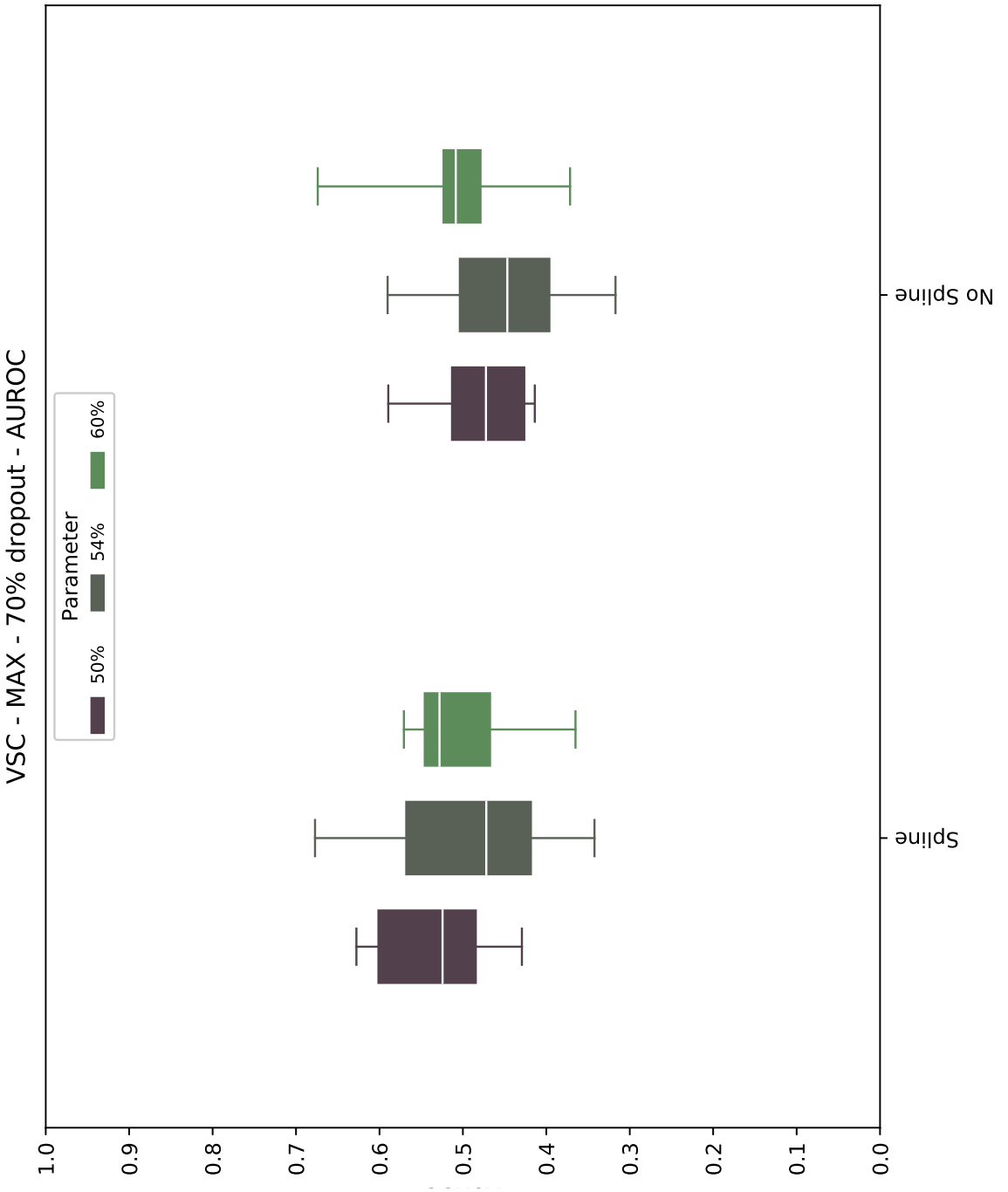

Figure 42. Results for problem VSC with and without Spline, considering AUROC and 70% dropout with parameters in range [50%, 54%, 60%].

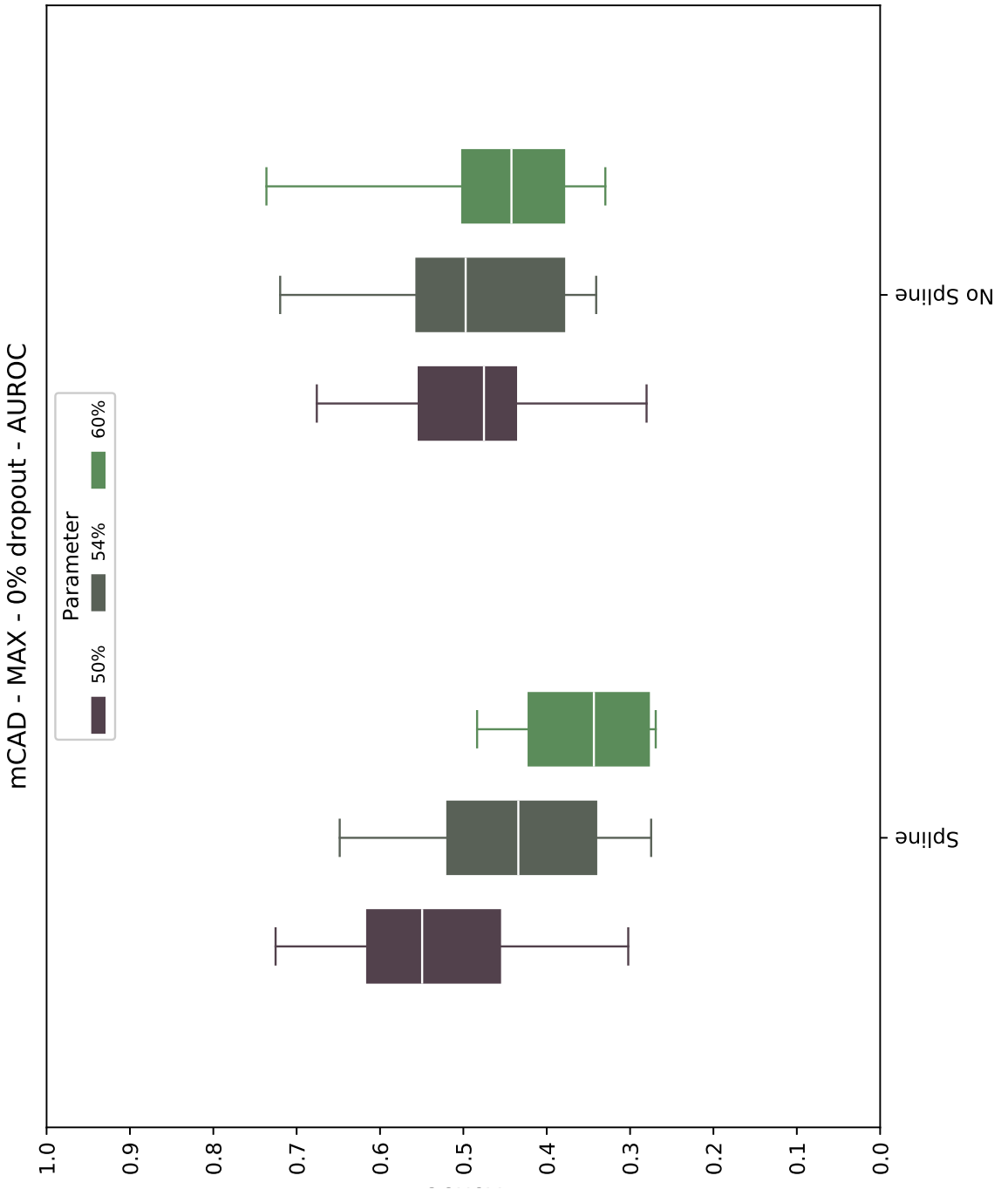

Figure 43. Results for problem mCAD with and without Spline, considering AUROC and 0% dropout with parameters in range [50%, 54%, 60%].

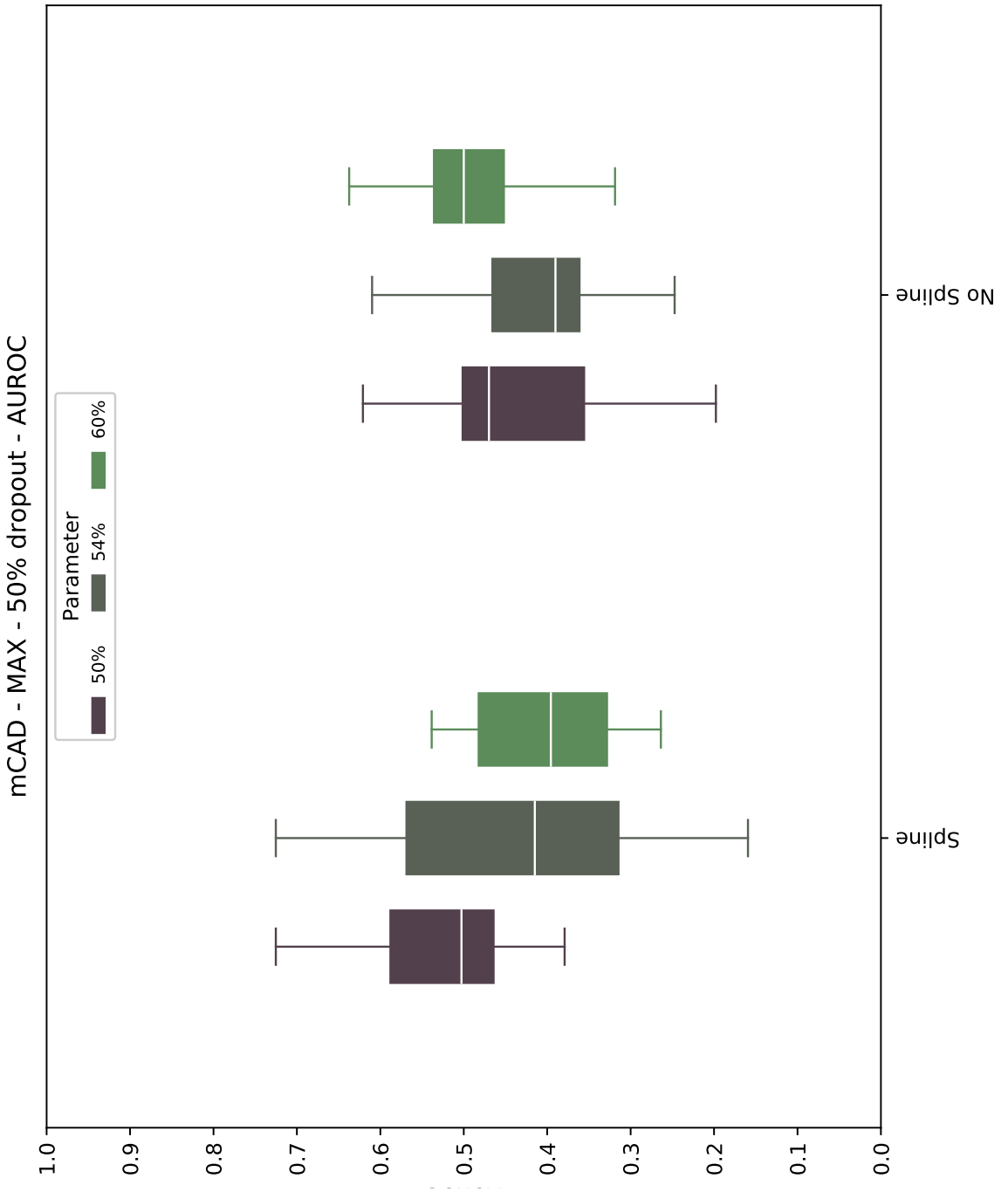

Figure 44. Results for problem mCAD with and without Spline, considering AUROC and 50% dropout with parameters in range [50%, 54%, 60%].

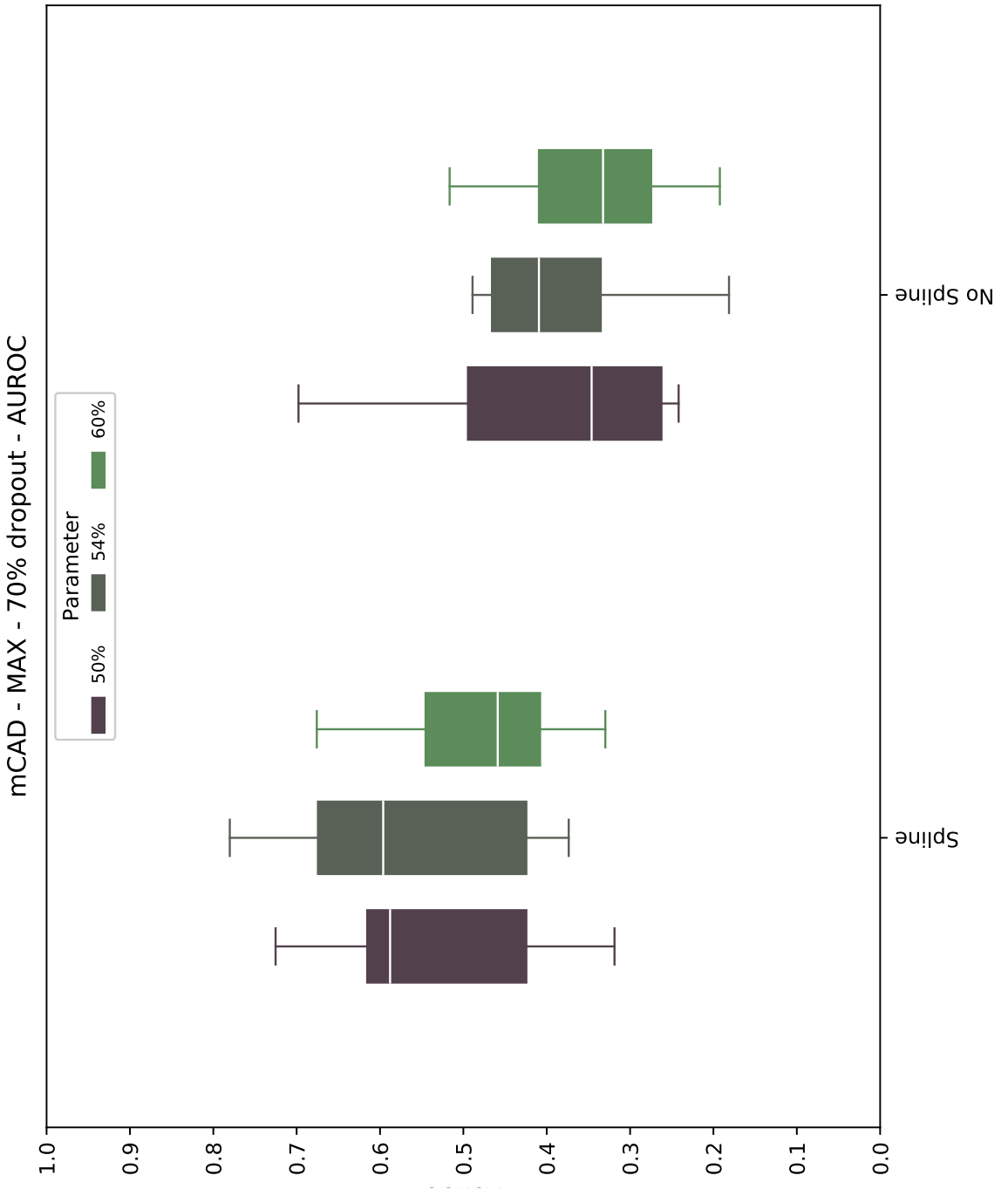

Figure 45. Results for problem mCAD with and without Spline, considering AUROC and 70% dropout with parameters in range [50%, 54%, 60%].

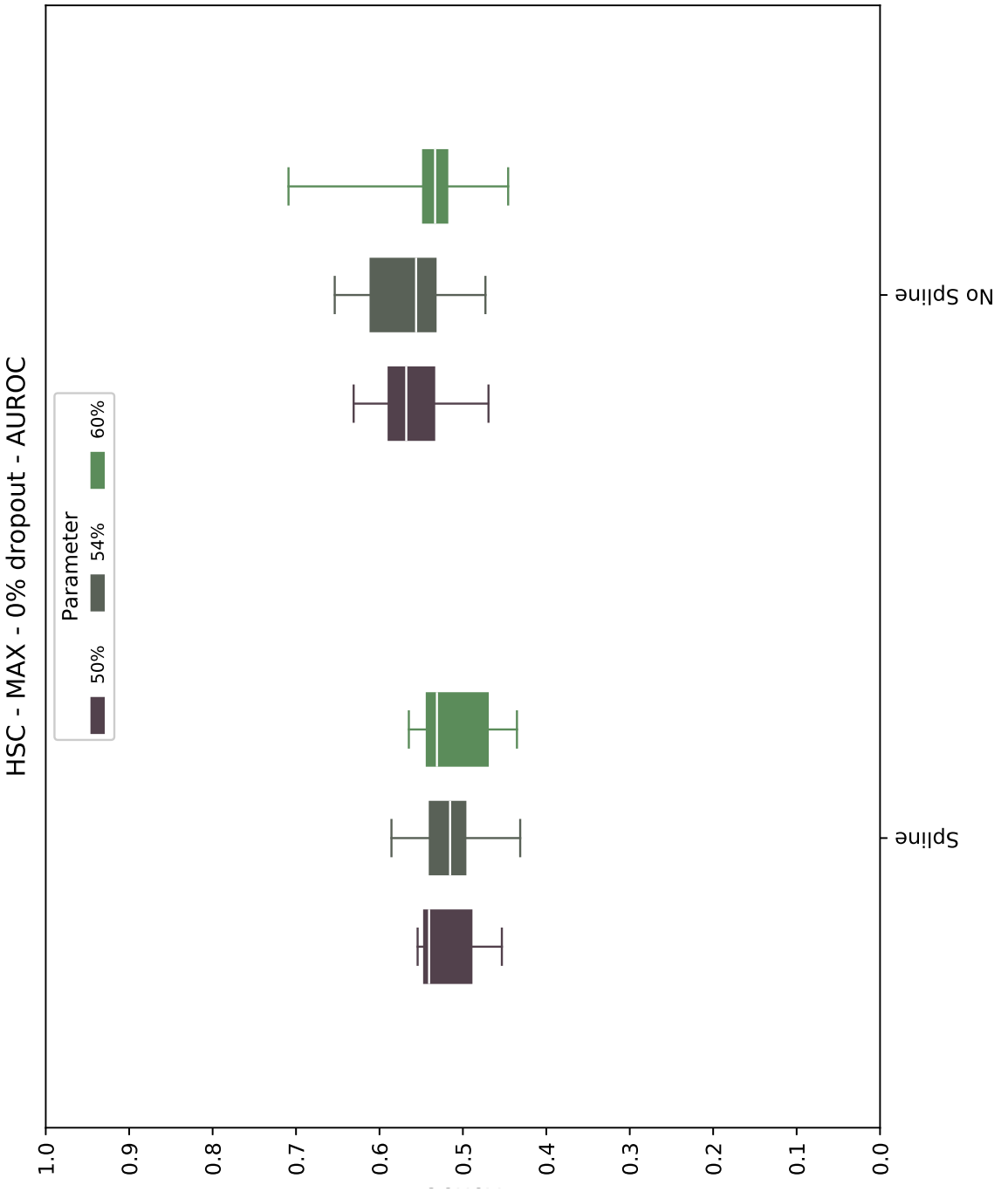

Figure 46. Results for problem HSC with and without Spline, considering AUROC and 0% dropout with parameters in range [50%, 54%, 60%].

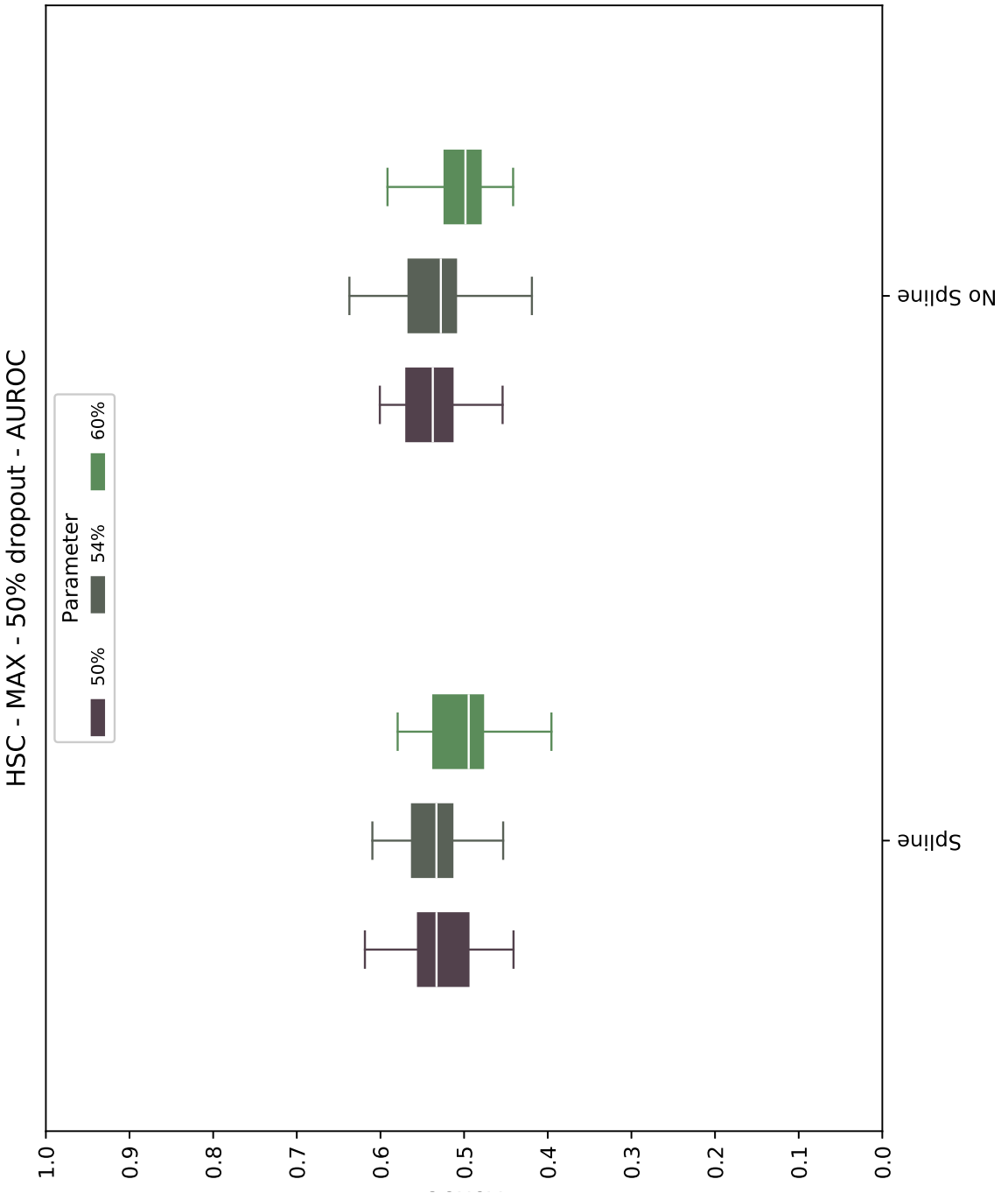

Figure 47. Results for problem HSC with and without Spline, considering AUROC and 50% dropout with parameters in range [50%, 54%, 60%].

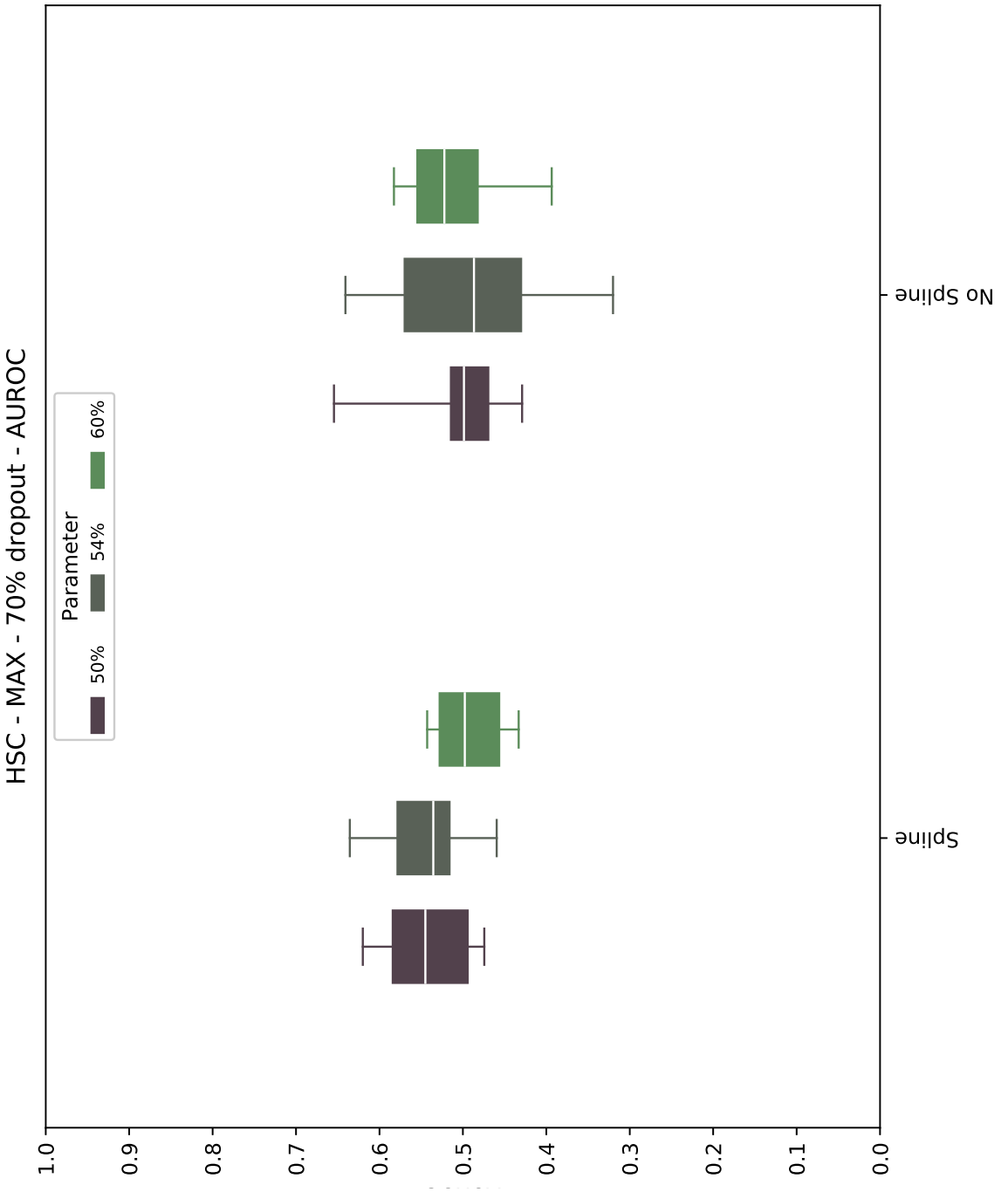

Figure 48. Results for problem HSC with and without Spline, considering AUROC and 70% dropout with parameters in range [50%, 54%, 60%].

3.2 Top%X

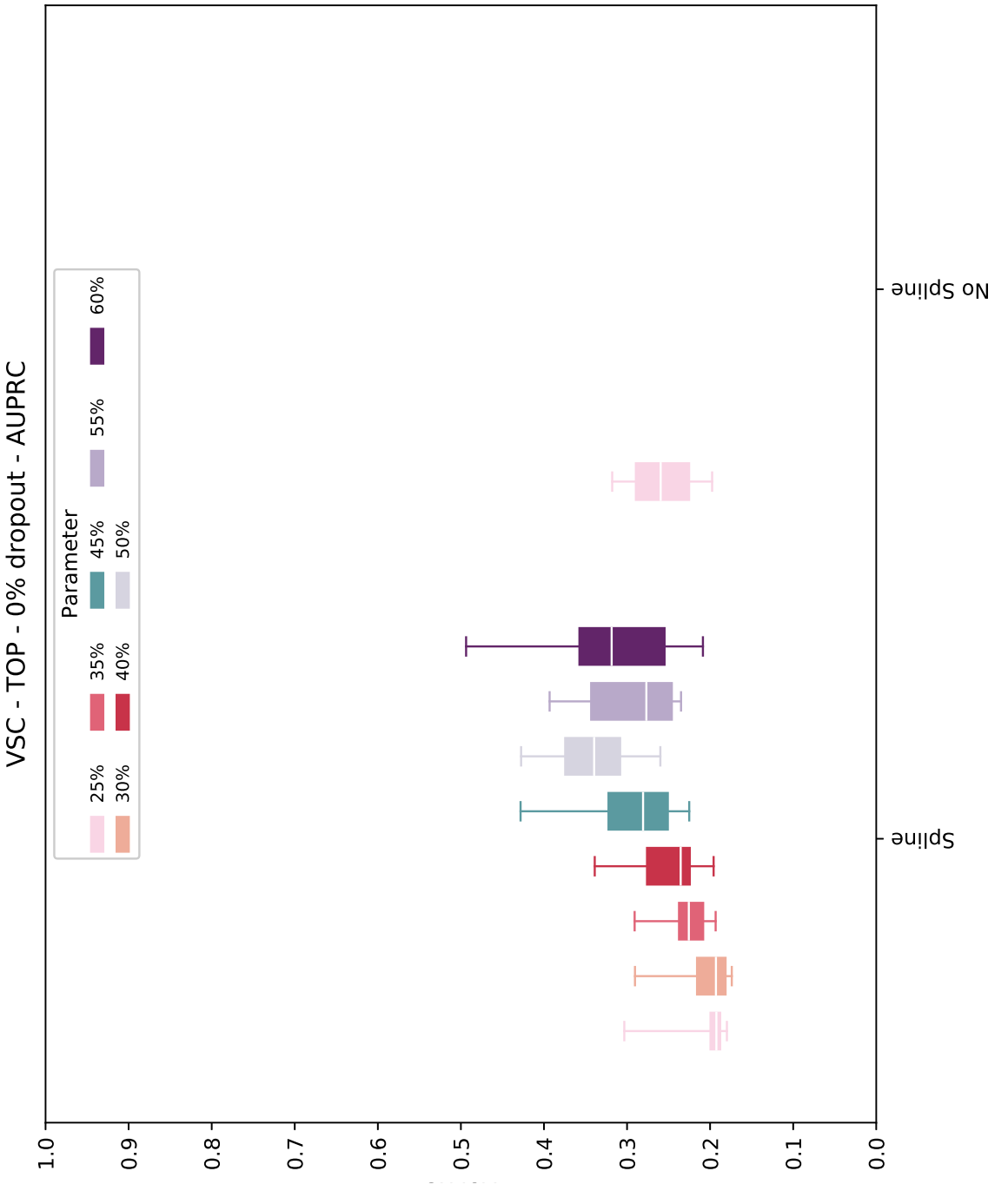

Figure 49. Results for problem VSC with and without Spline, considering AUPRC and 0% dropout with parameters in range [25%, ..., 60%].

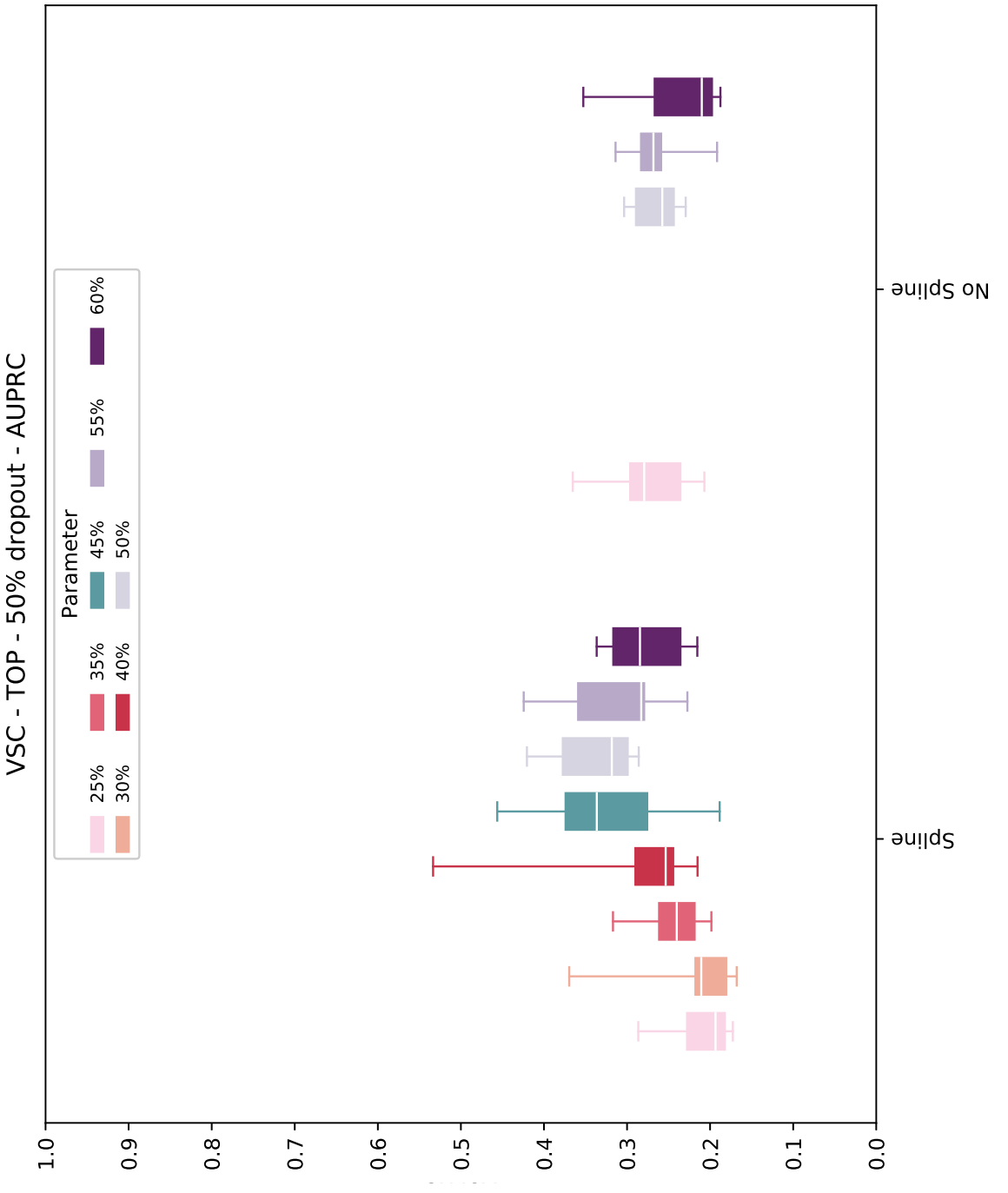

Figure 50. Results for problem VSC with and without Spline, considering AUPRC and 50% dropout with parameters in range [25%, ..., 60%].

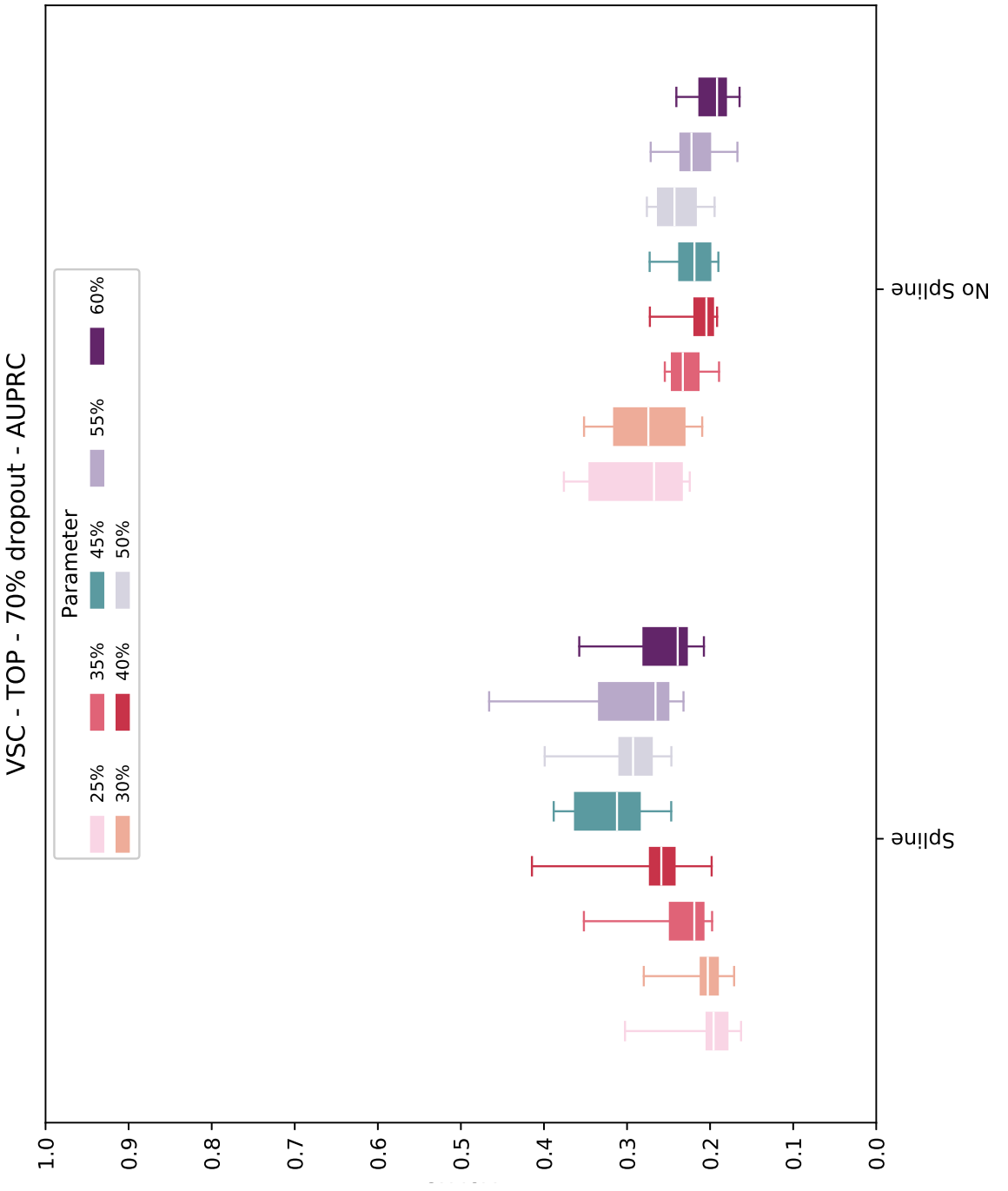

Figure 51. Results for problem VSC with and without Spline, considering AUPRC and 70% dropout with parameters in range [25%, ..., 60%].

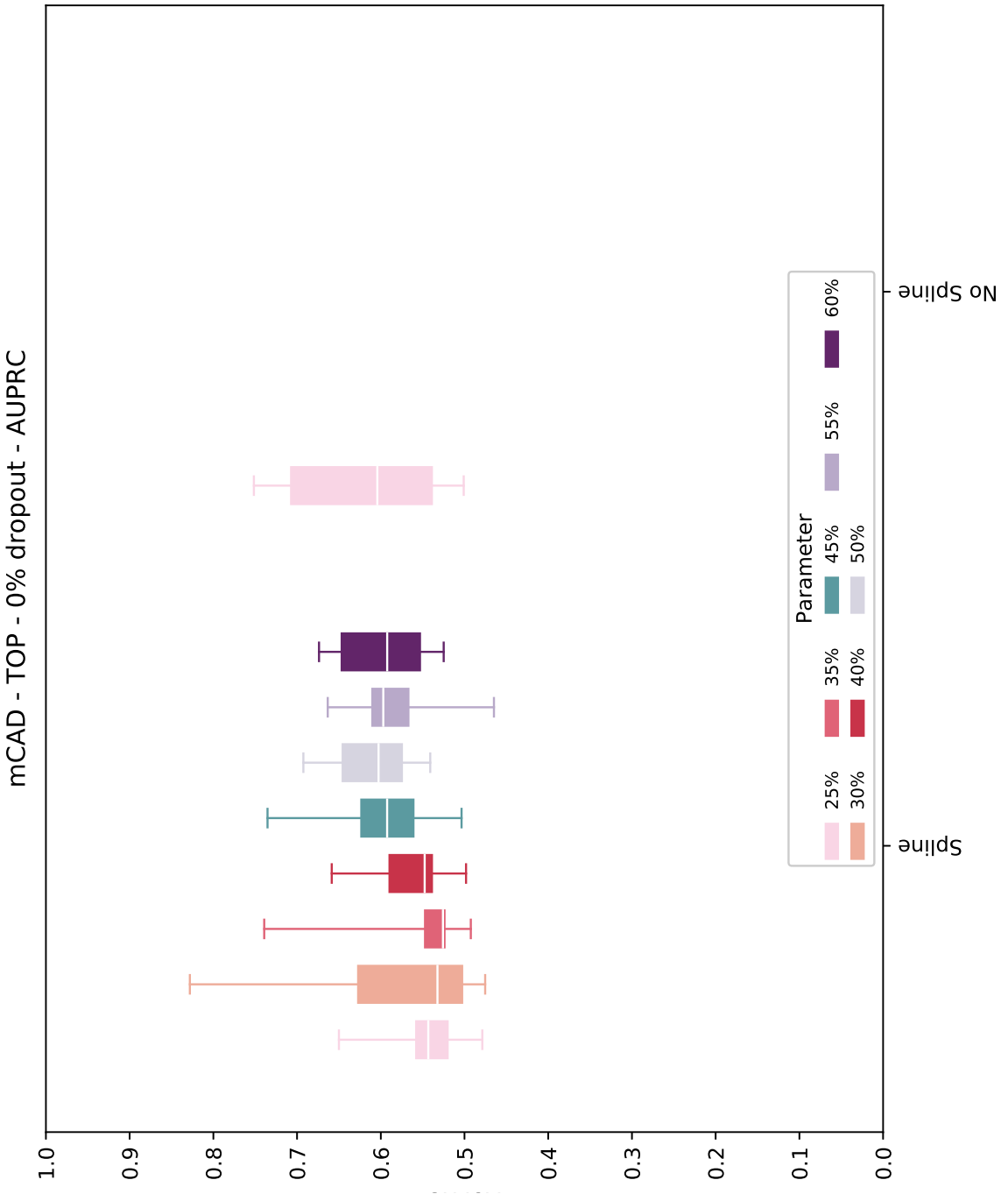

Figure 52. Results for problem mCAD with and without Spline, considering AUPRC and 0% dropout with parameters in range [25%, ..., 60%].

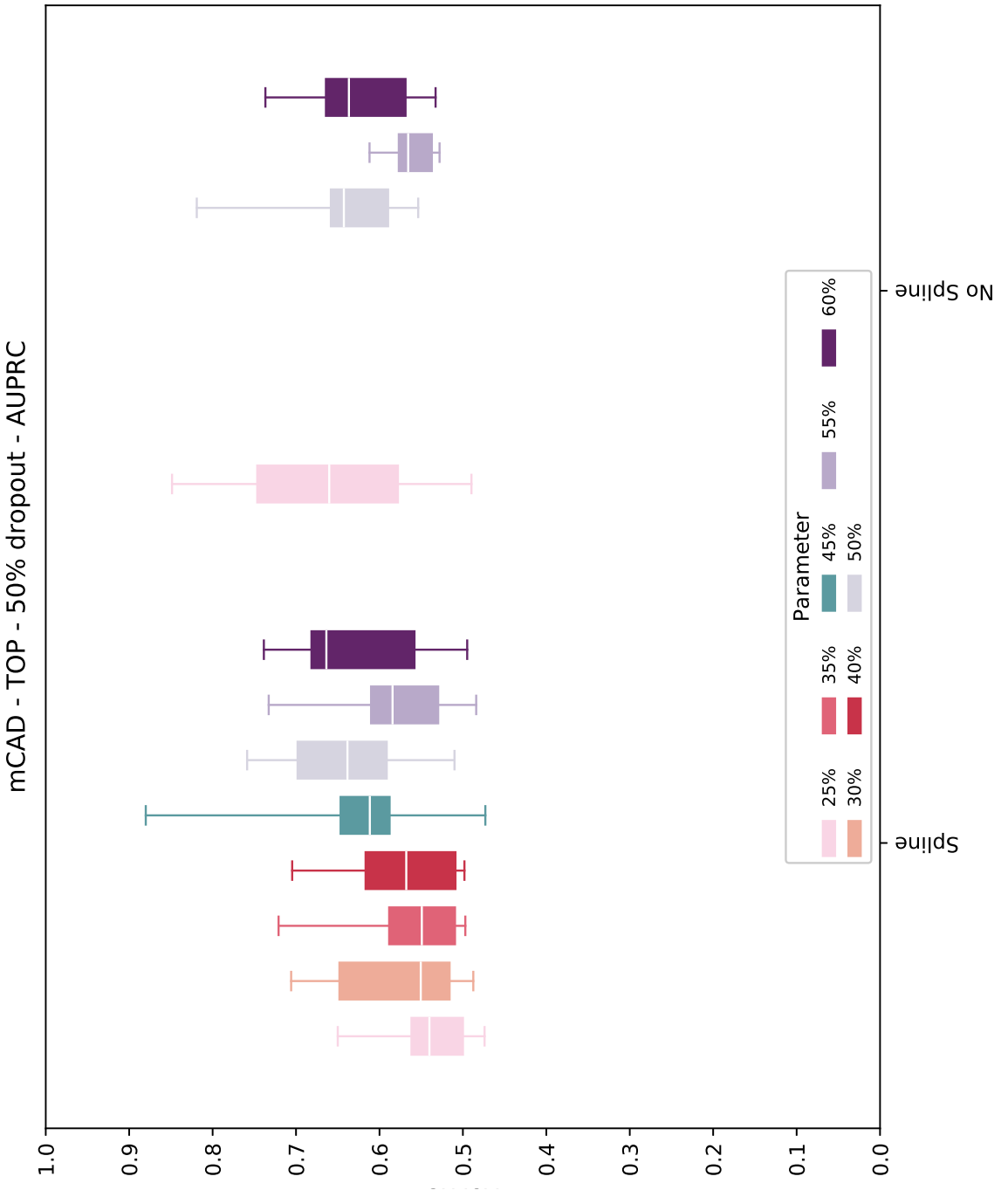

Figure 53. Results for problem mCAD with and without Spline, considering AUPRC and 50% dropout with parameters in range [25%, ..., 60%].

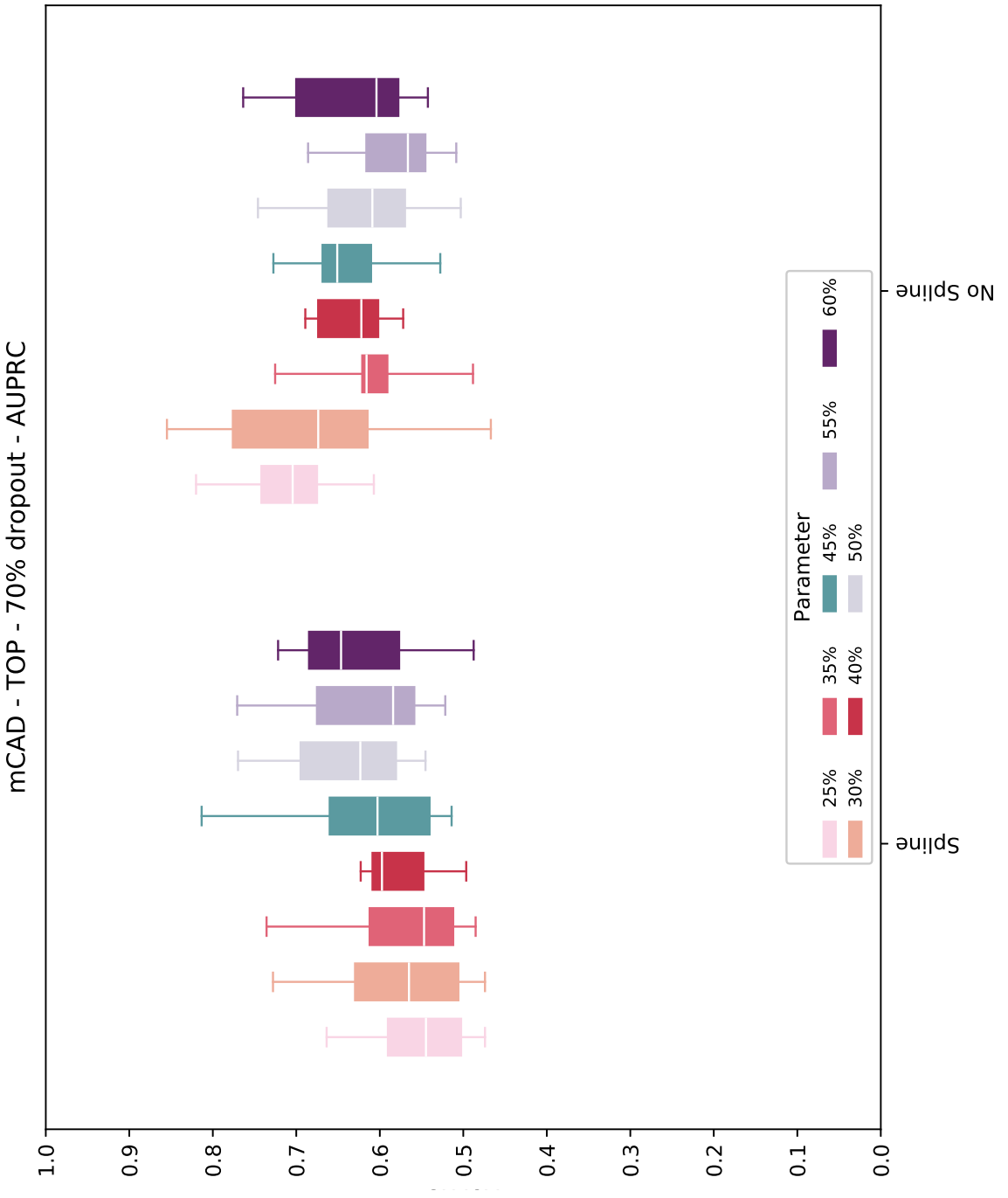

Figure 54. Results for problem mCAD with and without Spline, considering AUPRC and 70% dropout with parameters in range [25%, ..., 60%].

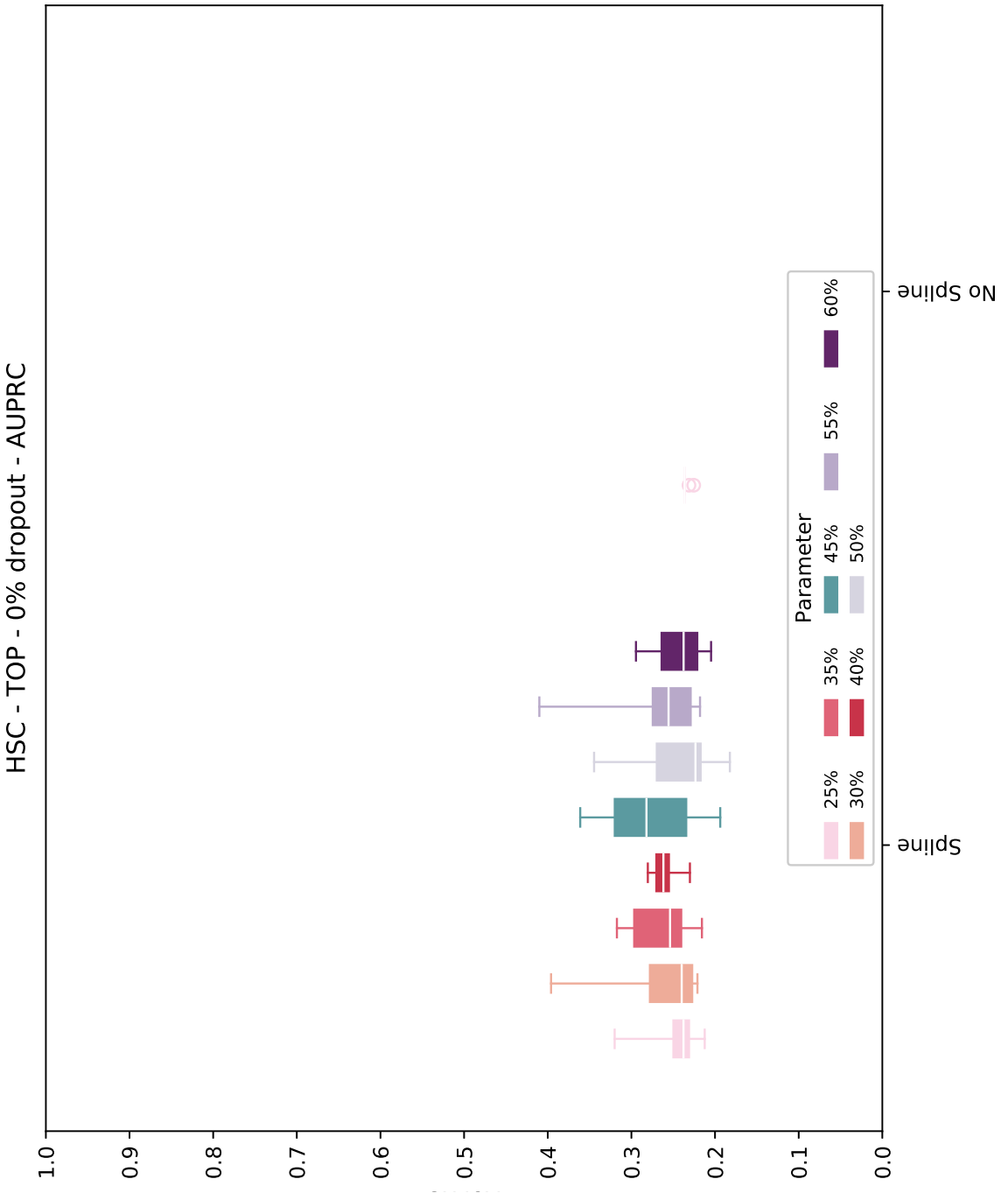

Figure 55. Results for problem HSC with and without Spline, considering AUPRC and 0% dropout with parameters in range [25%, ..., 60%].

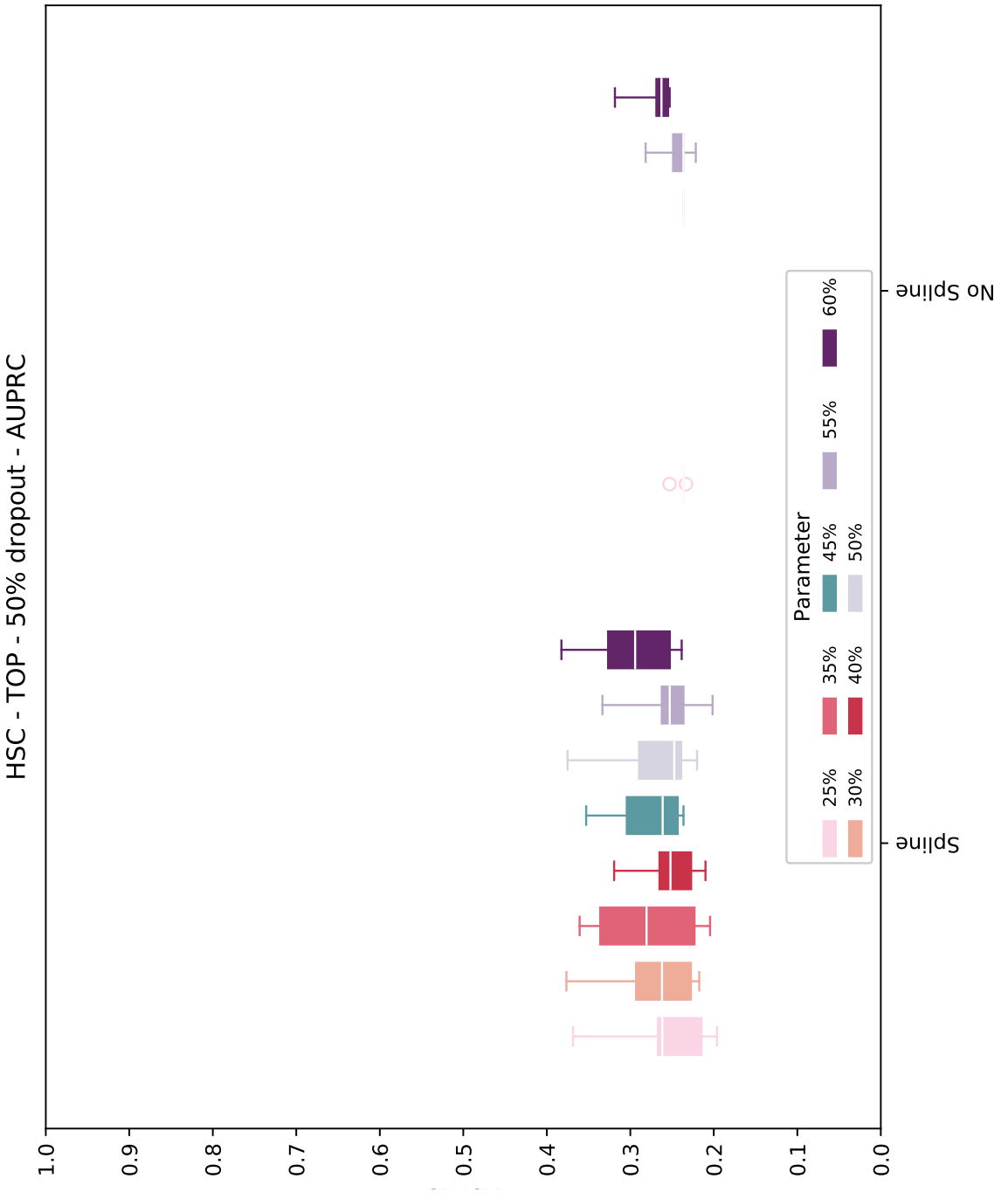

Figure 56. Results for problem HSC with and without Spline, considering AUPRC and 50% dropout with parameters in range [25%, ..., 60%].

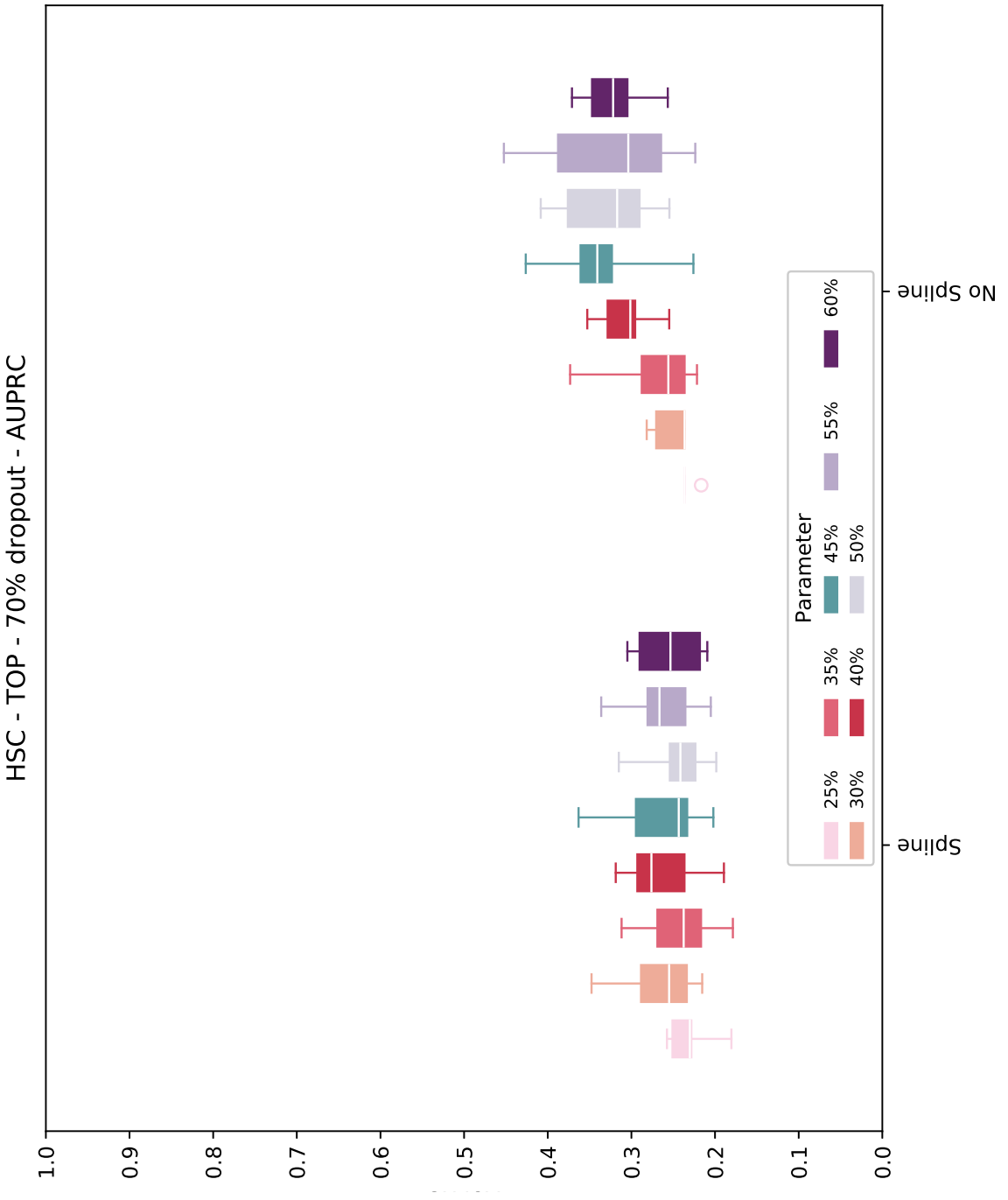

Figure 57. Results for problem HSC with and without Spline, considering AUPRC and 70% dropout with parameters in range [25%, ..., 60%].

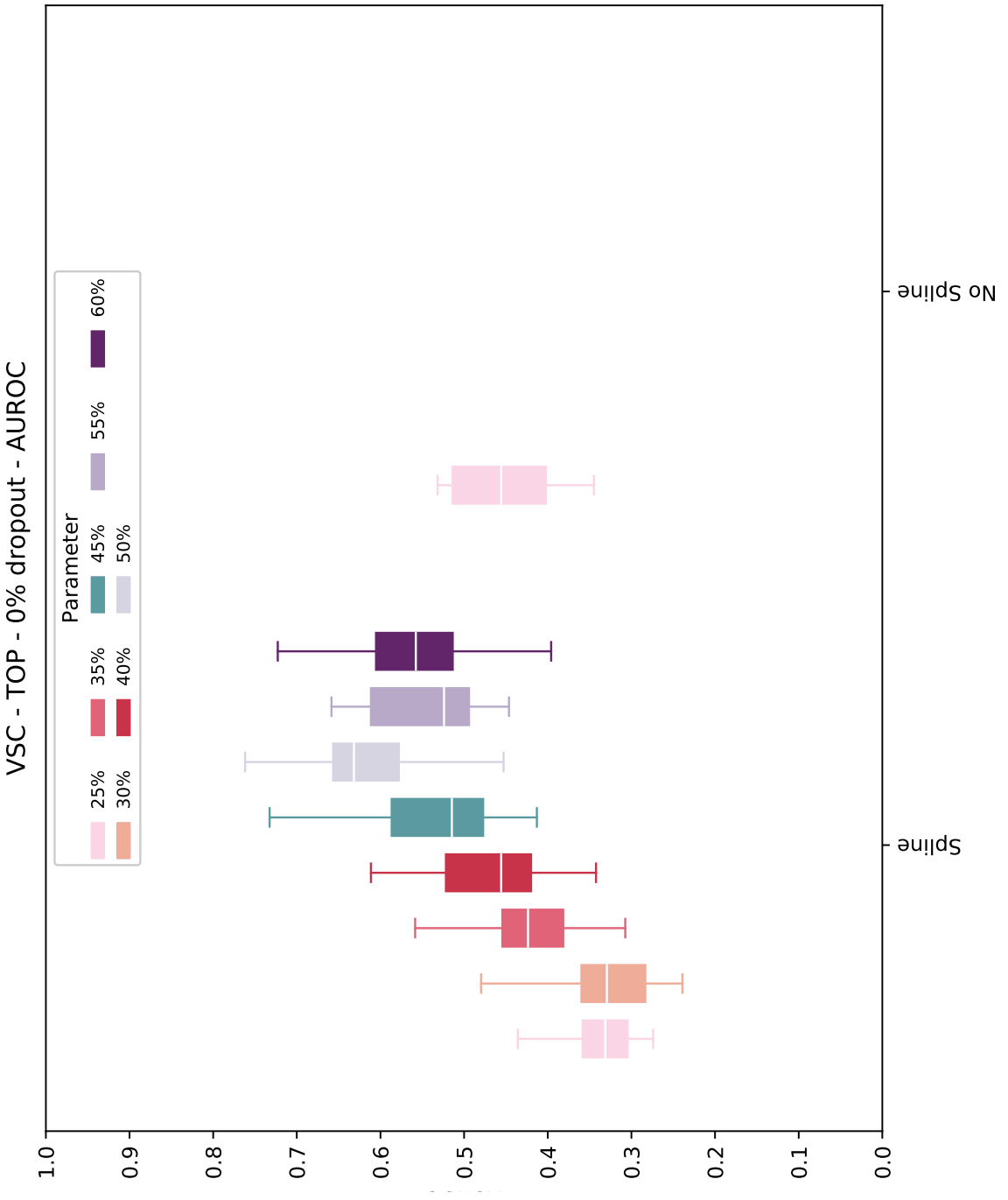

Figure 58. Results for problem VSC with and without Spline, considering AUROC and 0% dropout with parameters in range [25%, ..., 60%].

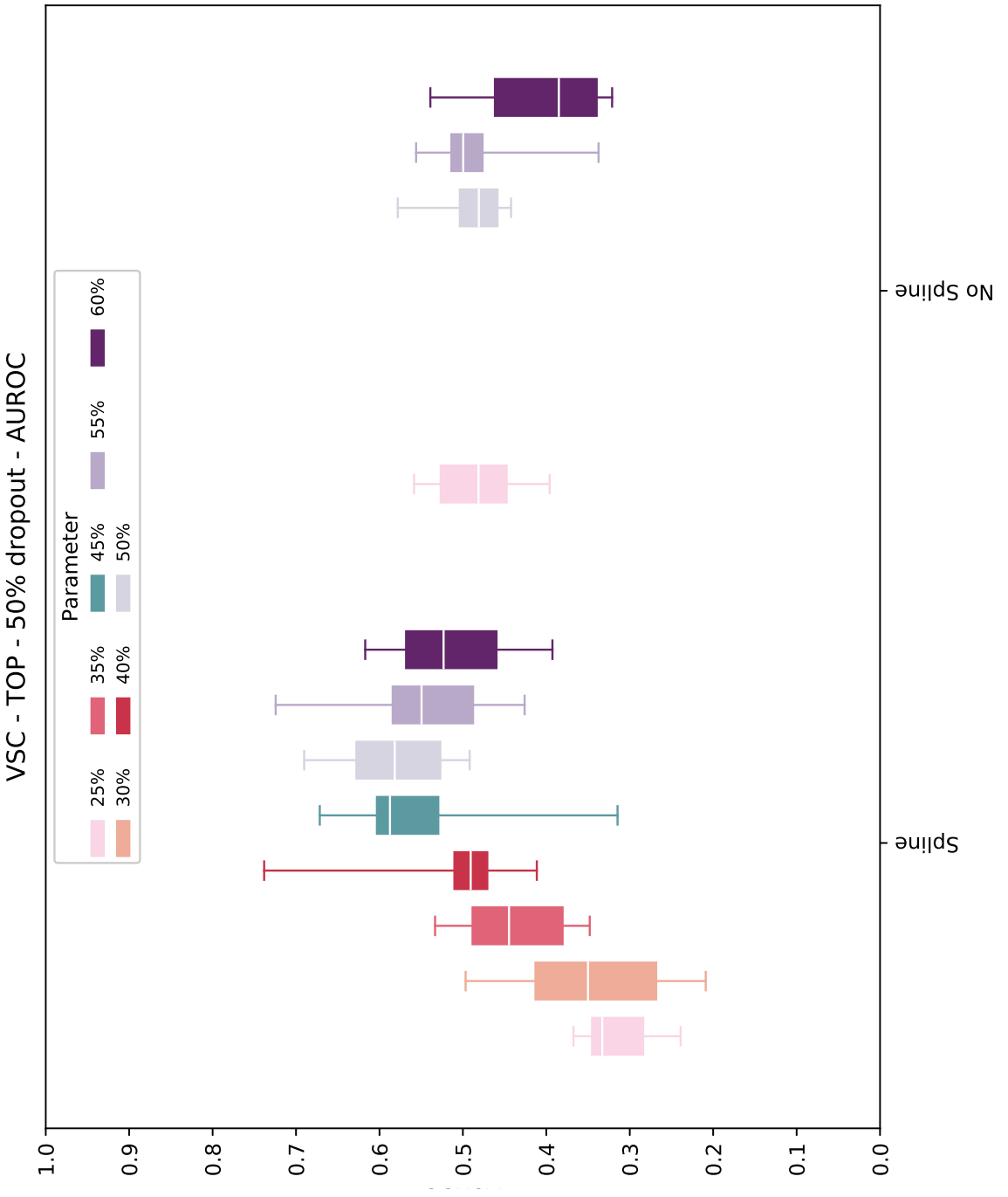

Figure 59. Results for problem VSC with and without Spline, considering AUROC and 50% dropout with parameters in range [25%, ..., 60%].

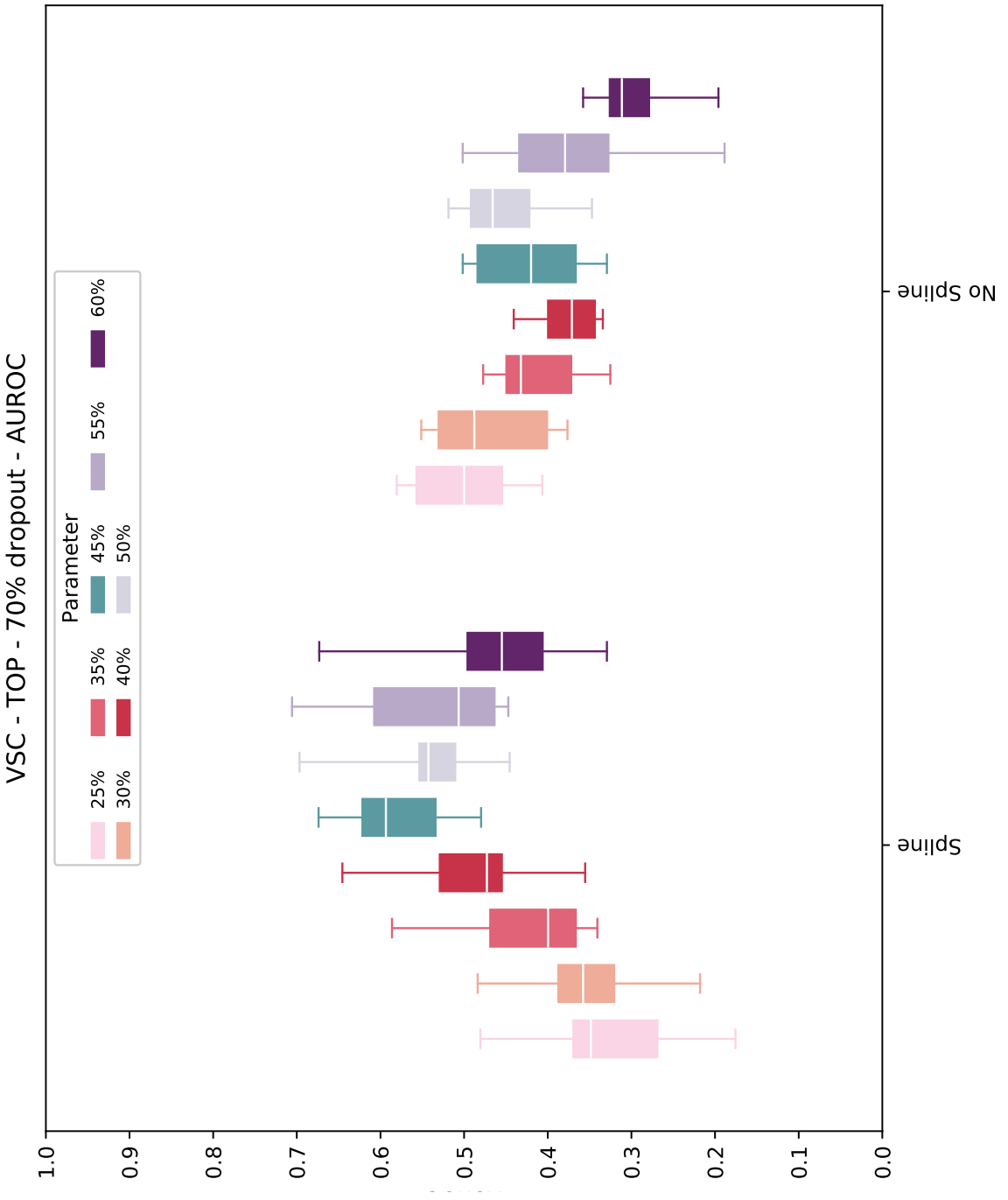

Figure 60. Results for problem VSC with and without Spline, considering AUROC and 70% dropout with parameters in range [25%, ..., 60%].

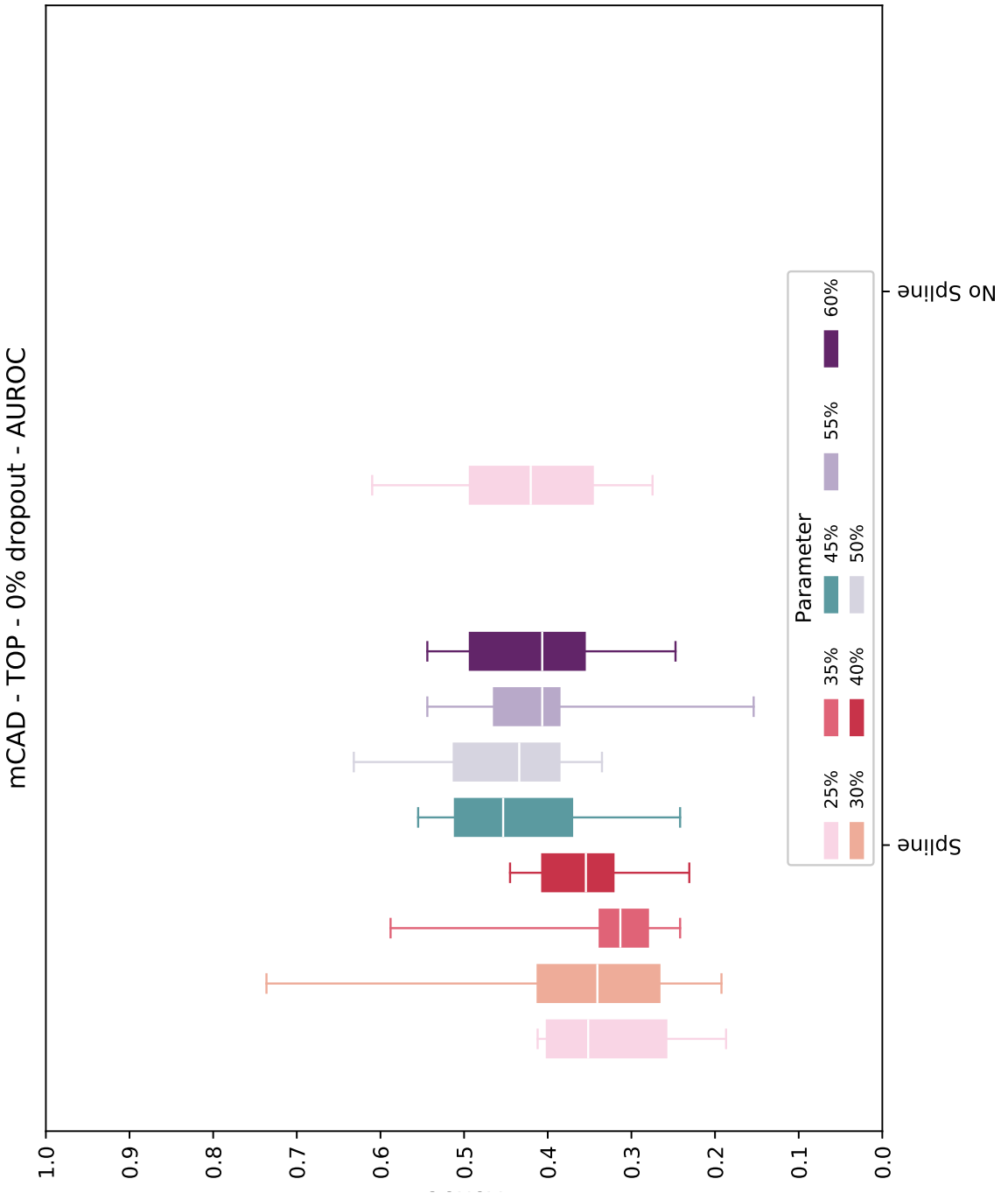

Figure 61. Results for problem mCAD with and without Spline, considering AUROC and 0% dropout with parameters in range [25%, ..., 60%].

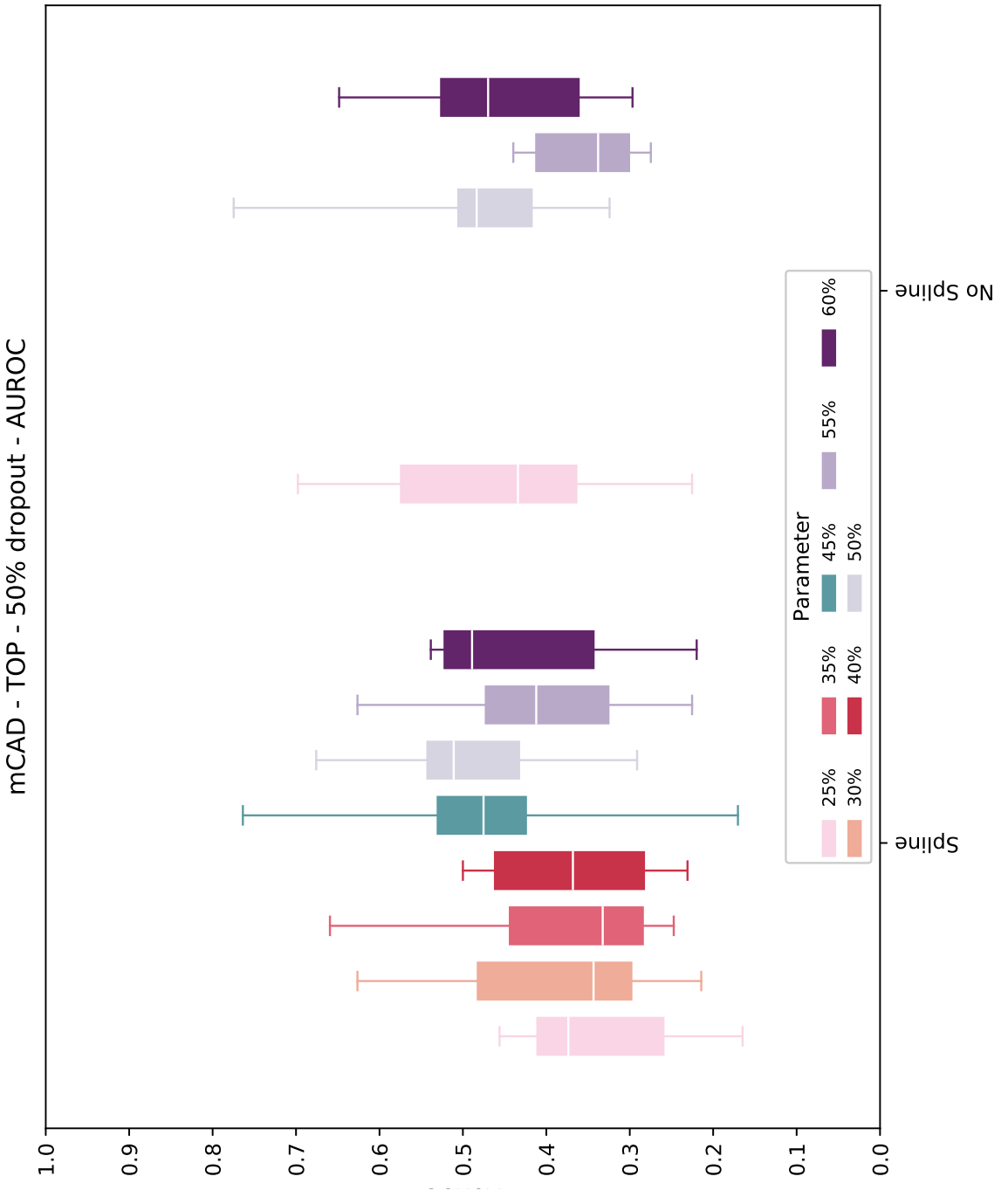

Figure 62. Results for problem mCAD with and without Spline, considering AUROC and 50% dropout with parameters in range [25%, ..., 60%].

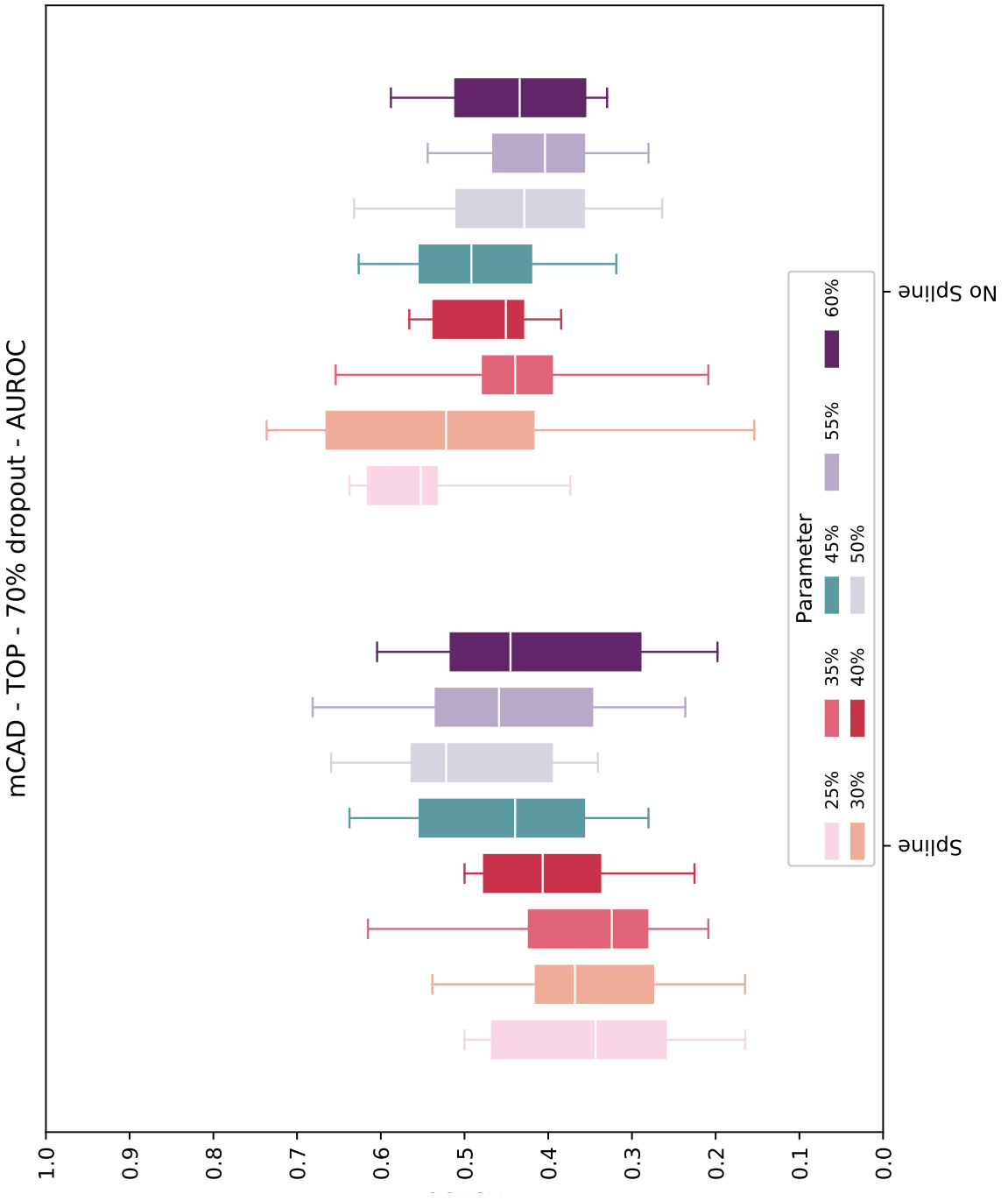

Figure 63. Results for problem mCAD with and without Spline, considering AUROC and 70% dropout with parameters in range [25%, ..., 60%].

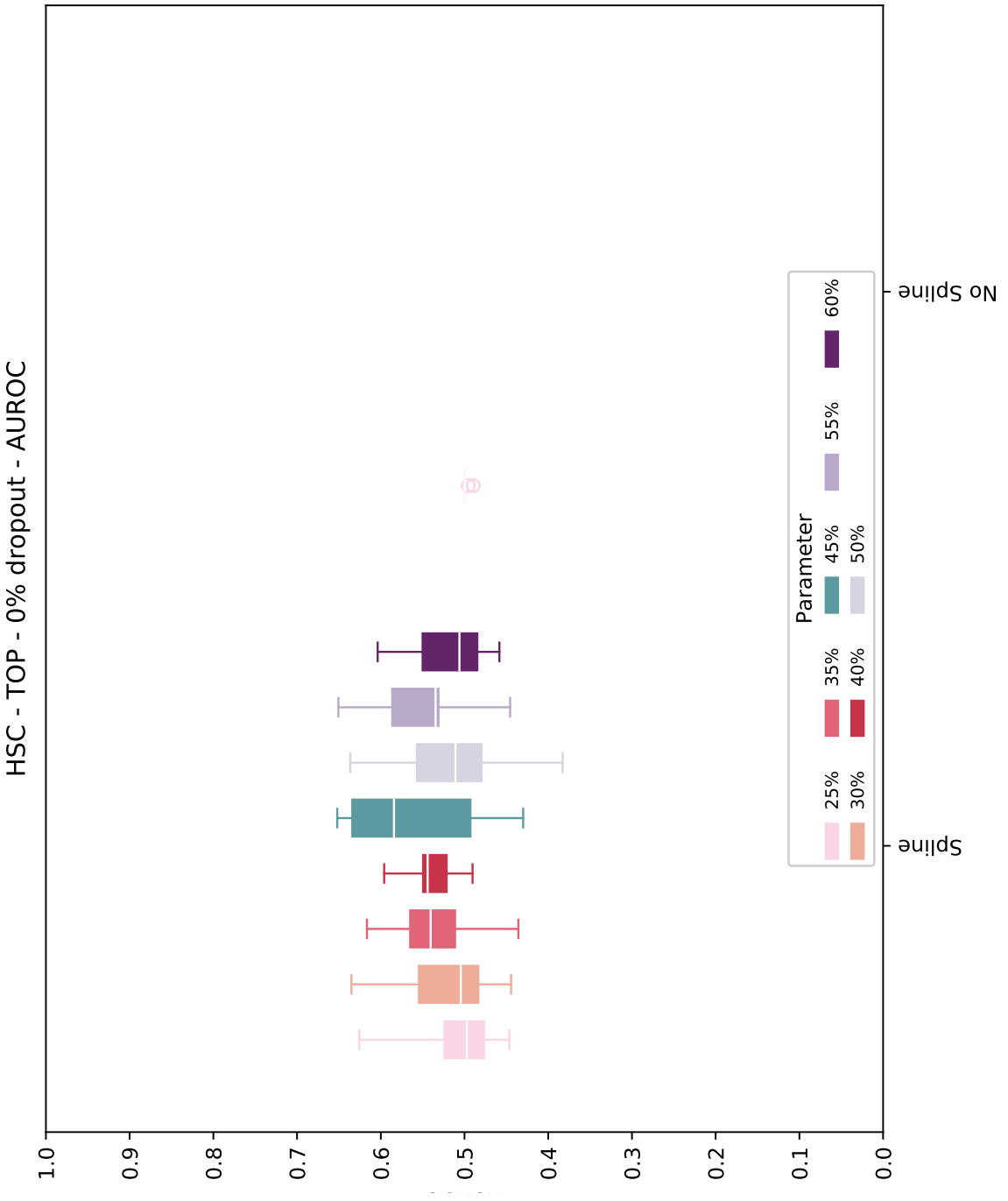

Figure 64. Results for problem HSC with and without Spline, considering AUROC and 0% dropout with parameters in range [25%, ..., 60%].

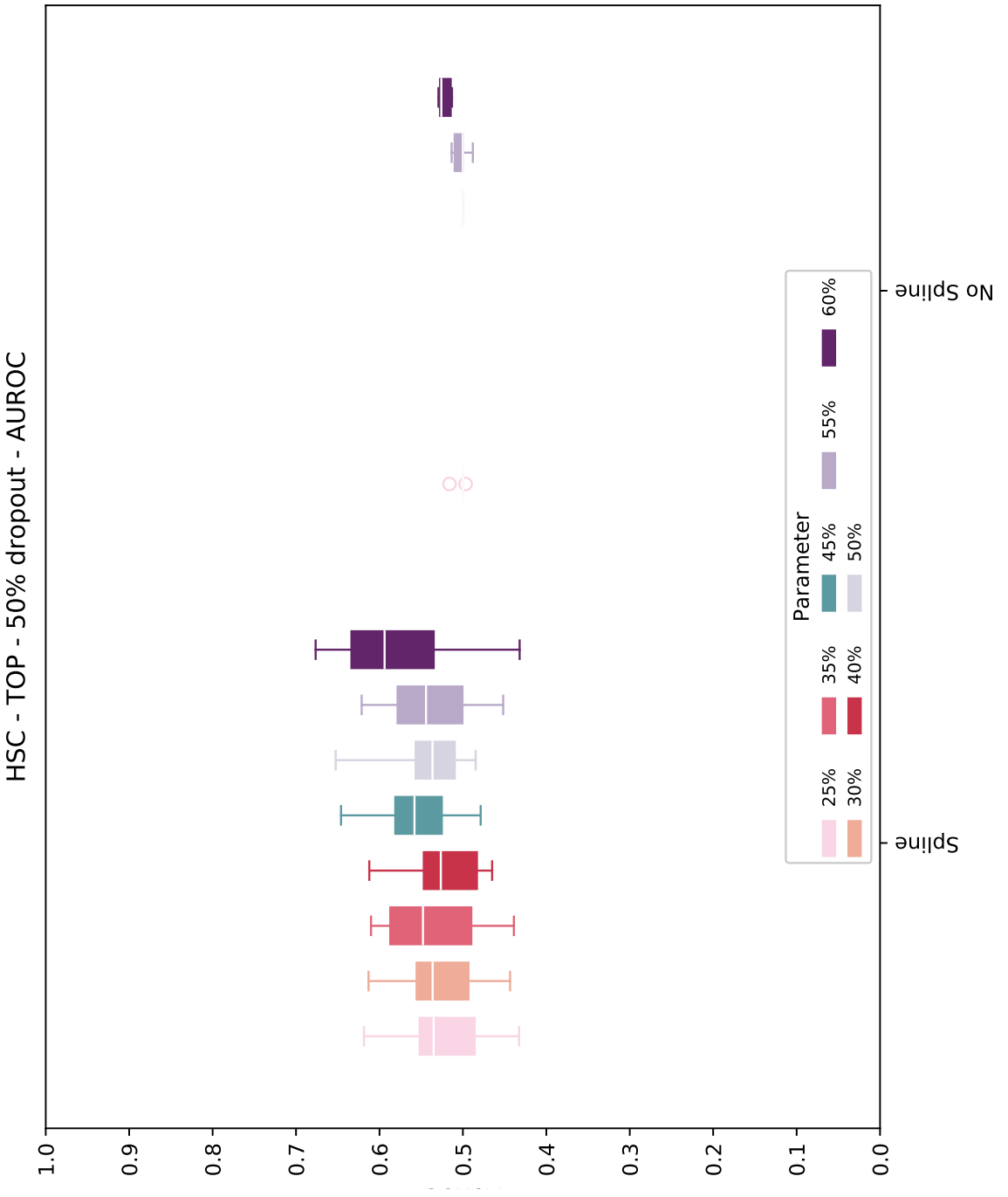

Figure 65. Results for problem HSC with and without Spline, considering AUROC and 50% dropout with parameters in range [25%, ..., 60%].

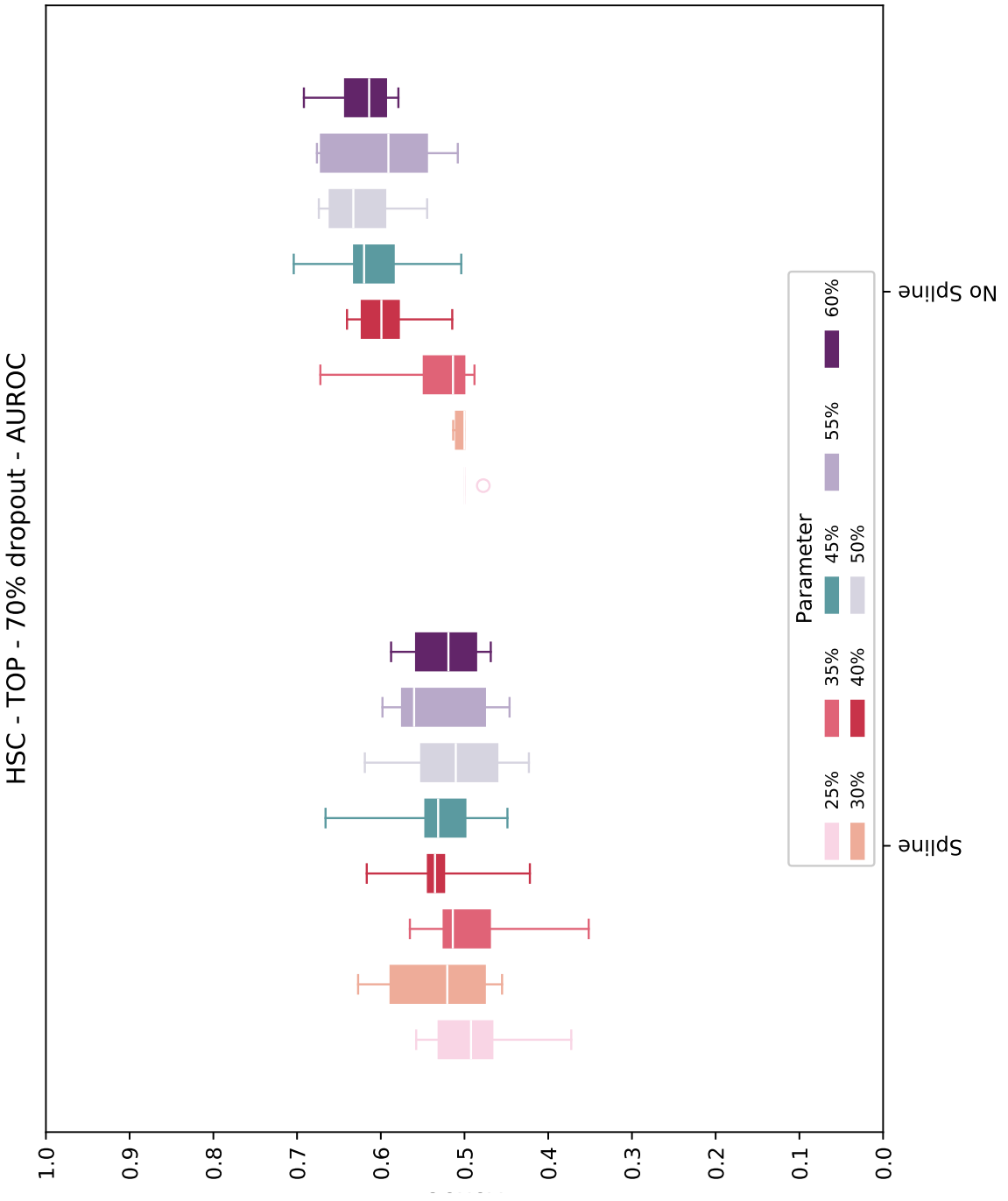

Figure 66. Results for problem HSC with and without Spline, considering AUROC and 70% dropout with parameters in range [25%, ..., 60%].

References

[1] S. C. Madeira and A. L. Oliveira. An evaluation of discretization methods for non-supervised analysis of time-series gene expression data. *INESC-ID Technical Report*, vol. 42, pp. 2005, 2005.