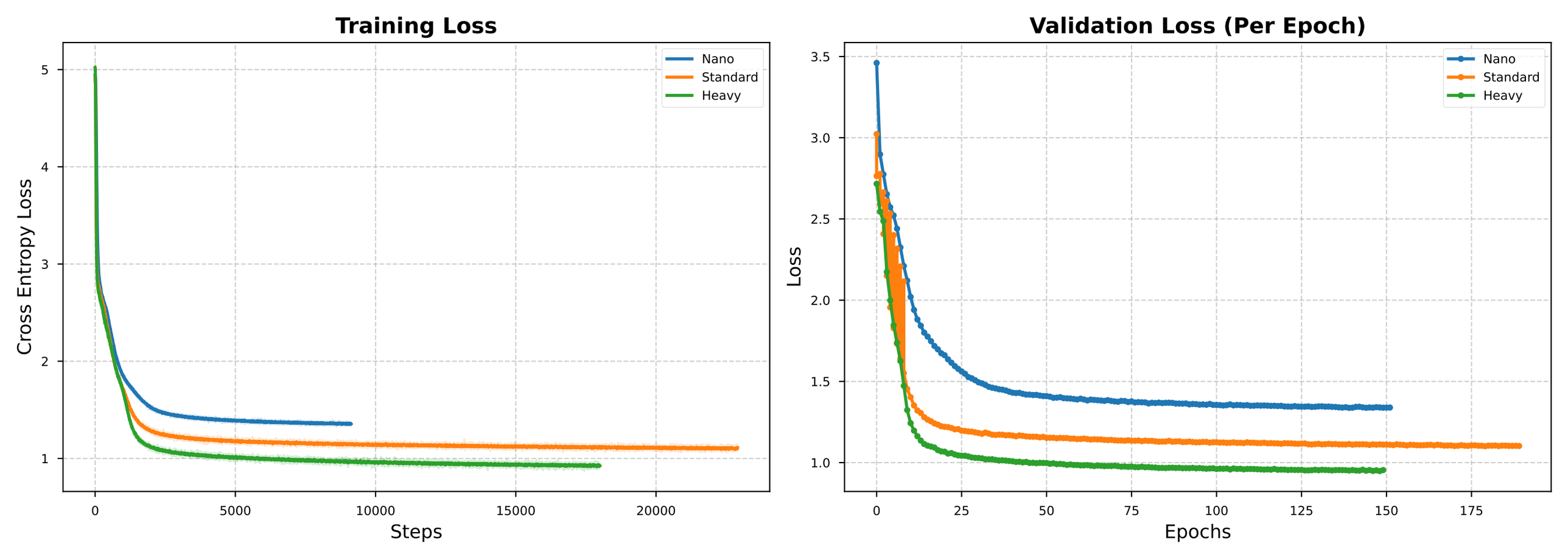
Training Loss
Cross Entropy Loss
Steps
Nano
Standard
Heavy
0
5000
10000
15000
20000
1
2
3
4
5
Validation Loss (Per Epoch)
Loss
Epochs
Nano
Standard
Heavy
1.0
1.5
2.0
2.5
3.0
3.5
0
25
50
75
100
125
150
175