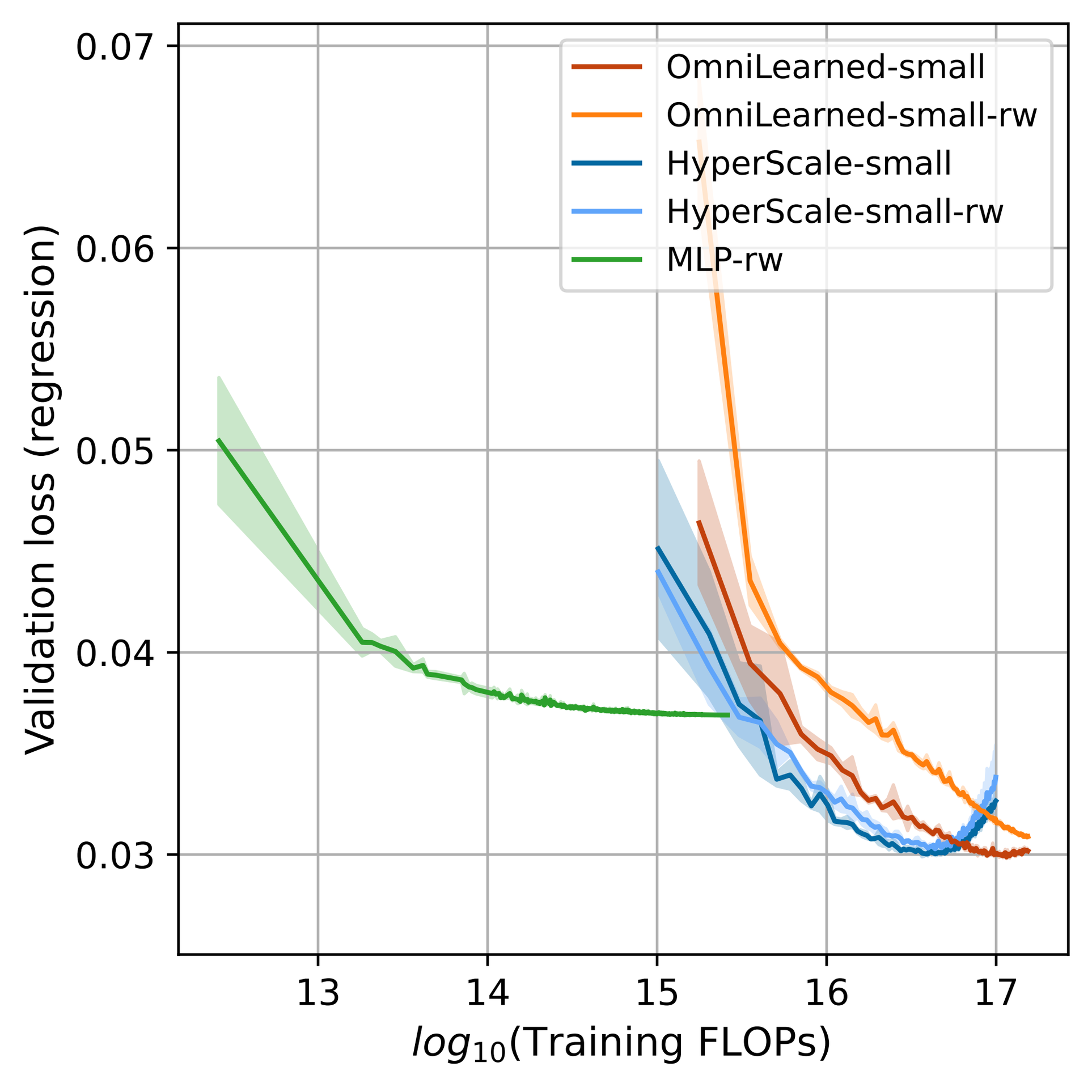

OmniLearned-small
OmniLearned-small-rw
HyperScale-small
HyperScale-small-rw
MLP-rw
Validation loss (regression)
0.07
0.06
0.05
0.04
0.03
13
14
15
16
17
log_{10}(Training FLOPs)