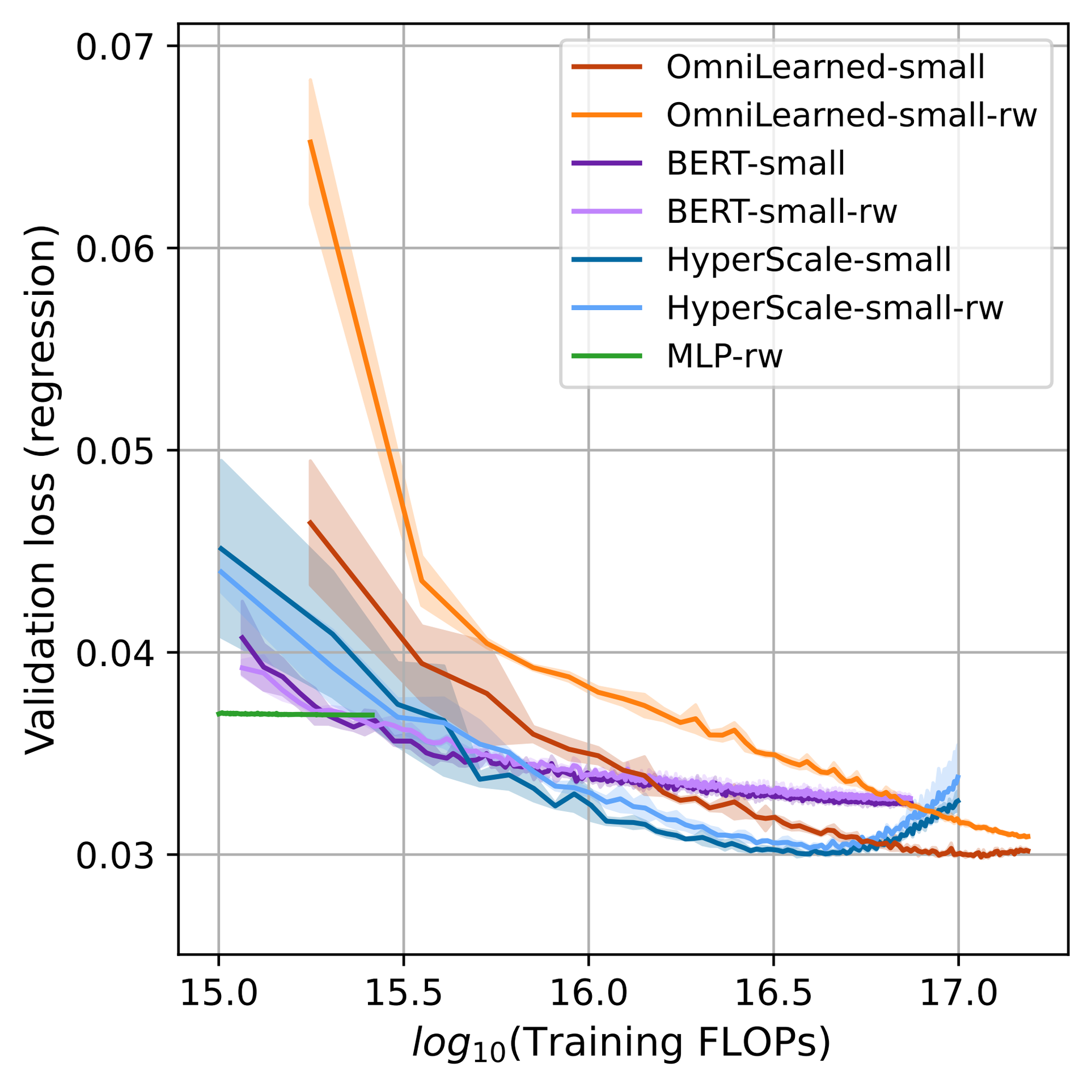

0.07
0.06
0.05
0.04
0.03
Validation loss (regression)
15.0
15.5
16.0
16.5
17.0
log_{10}(Training FLOPs)
OmniLearned-small
OmniLearned-small-rw
BERT-small
BERT-small-rw
HyperScale-small
HyperScale-small-rw
MLP-rw