Training loss loss for different choices of wd