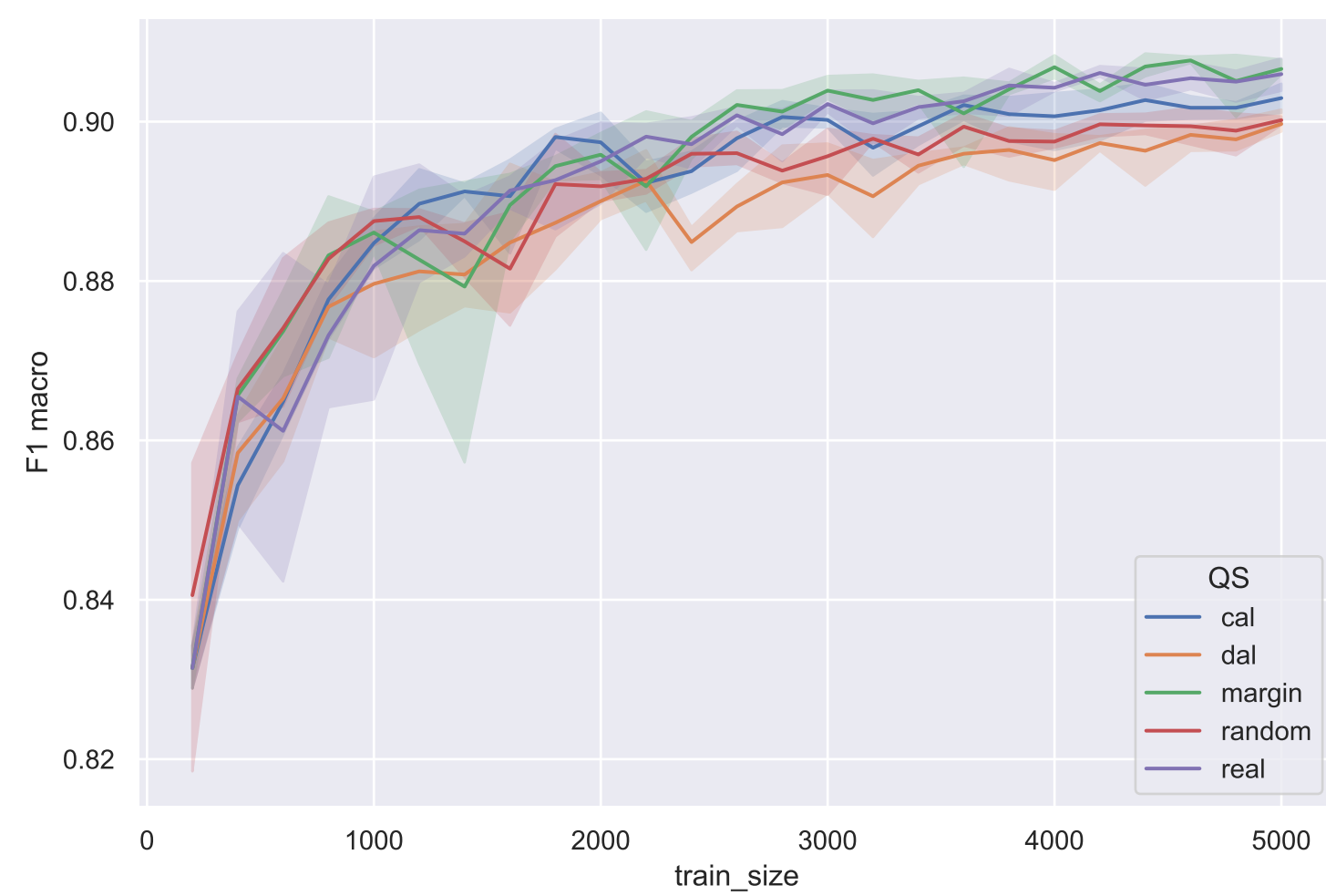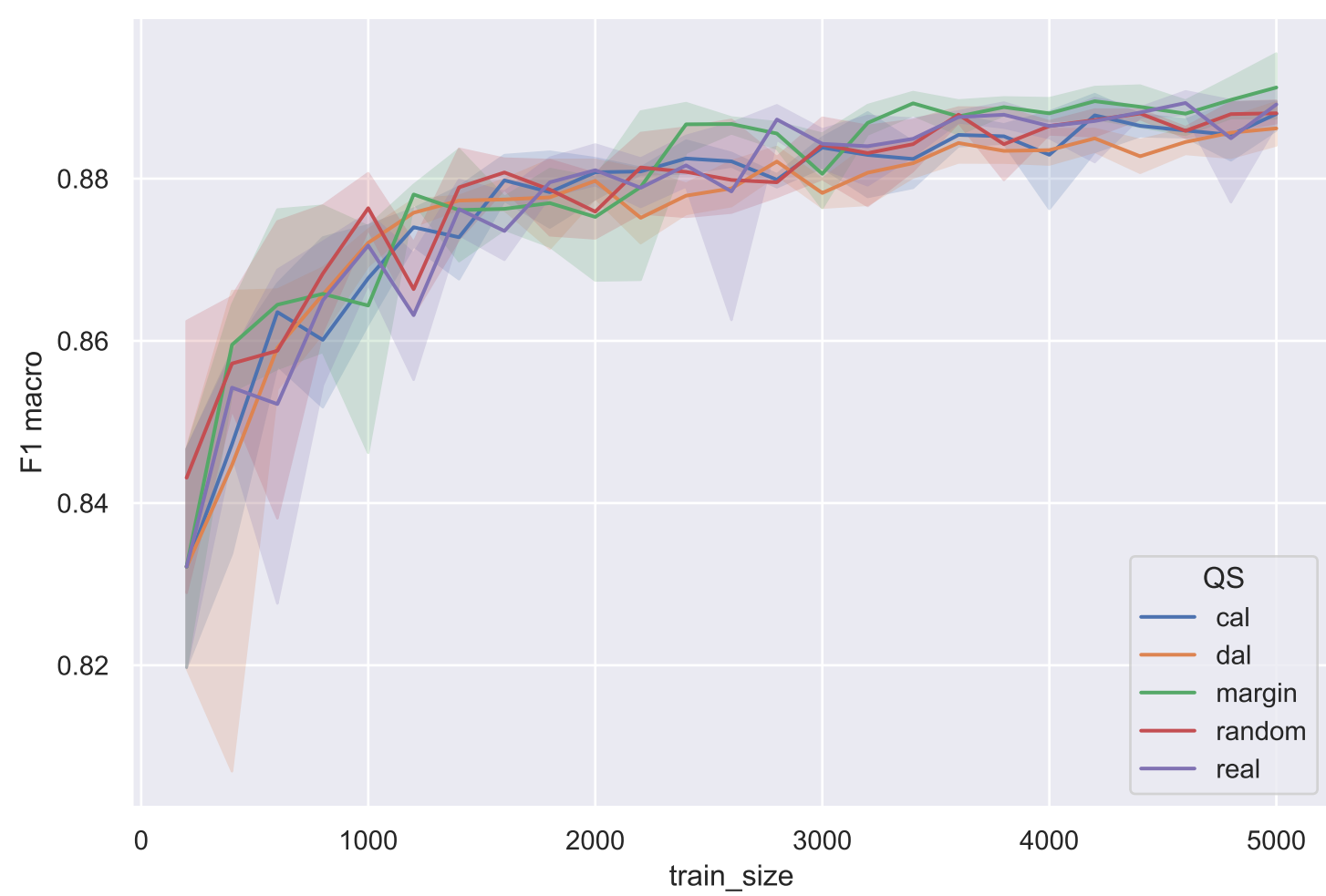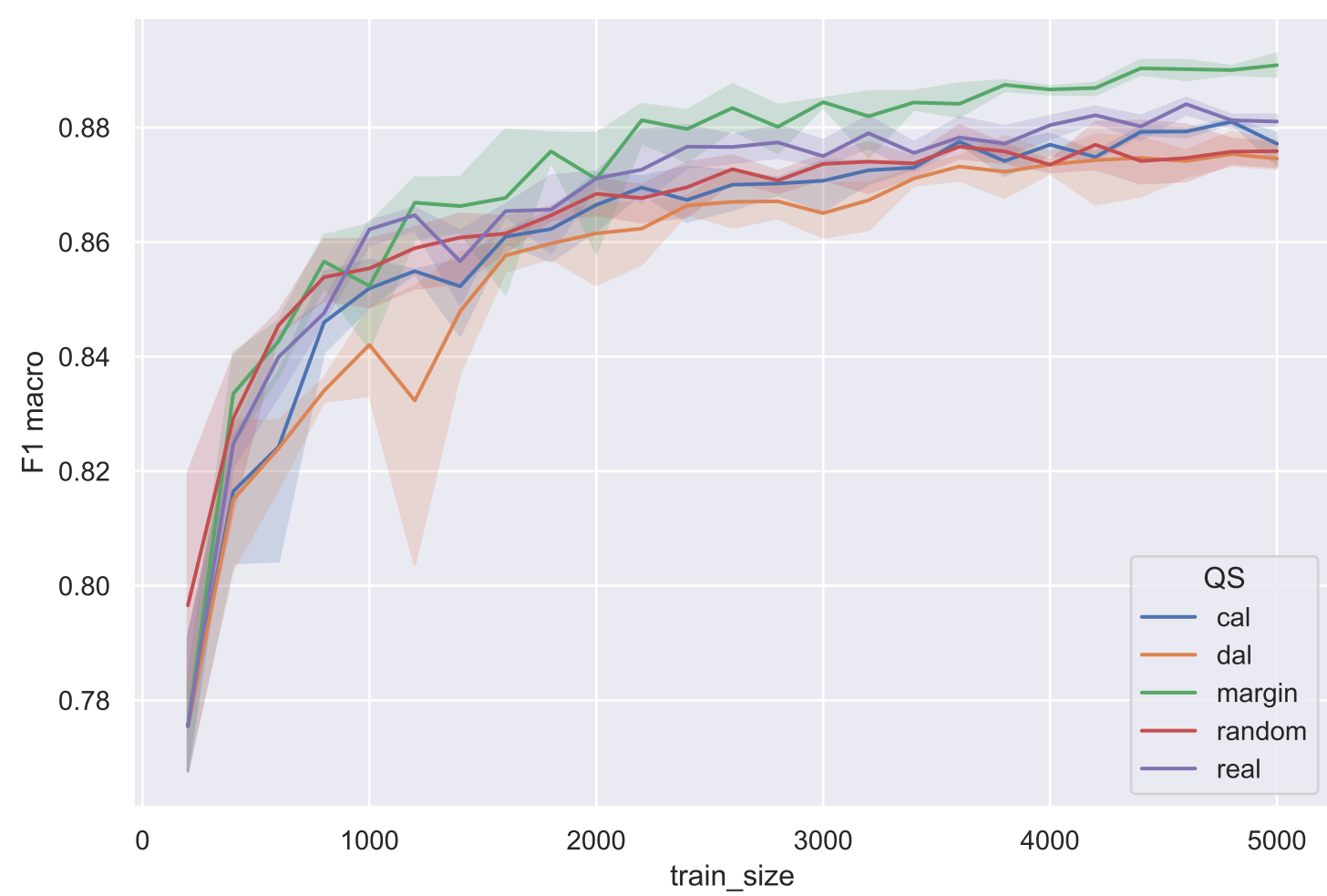dataset: agnews, batch size=200