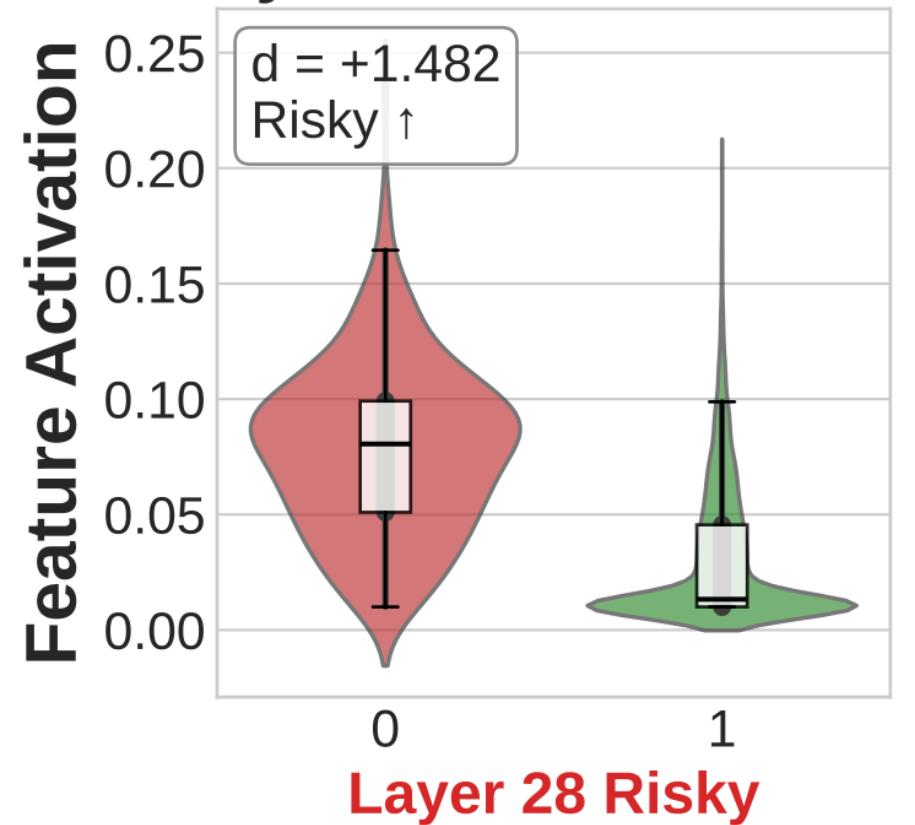
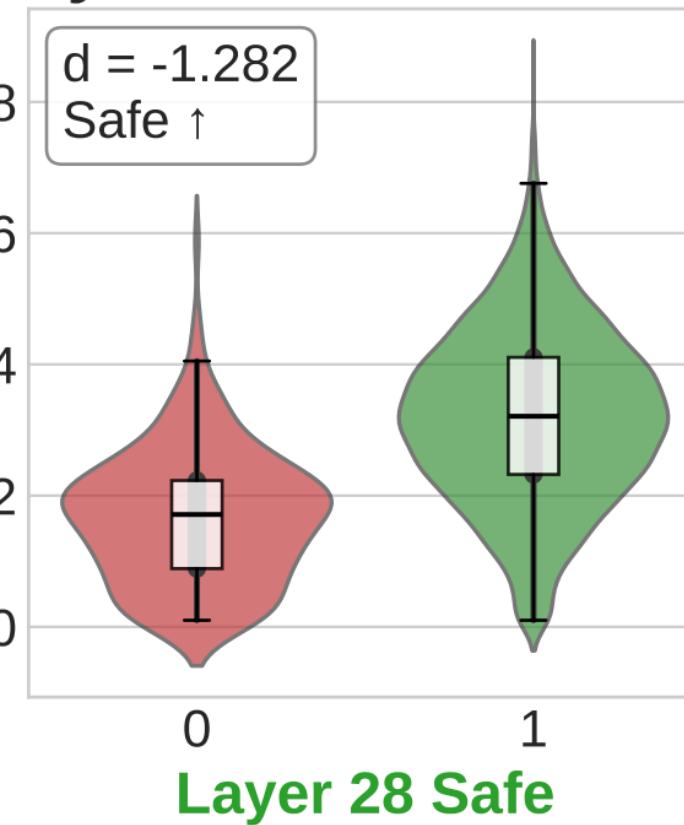


# SAE Feature Separation Analysis

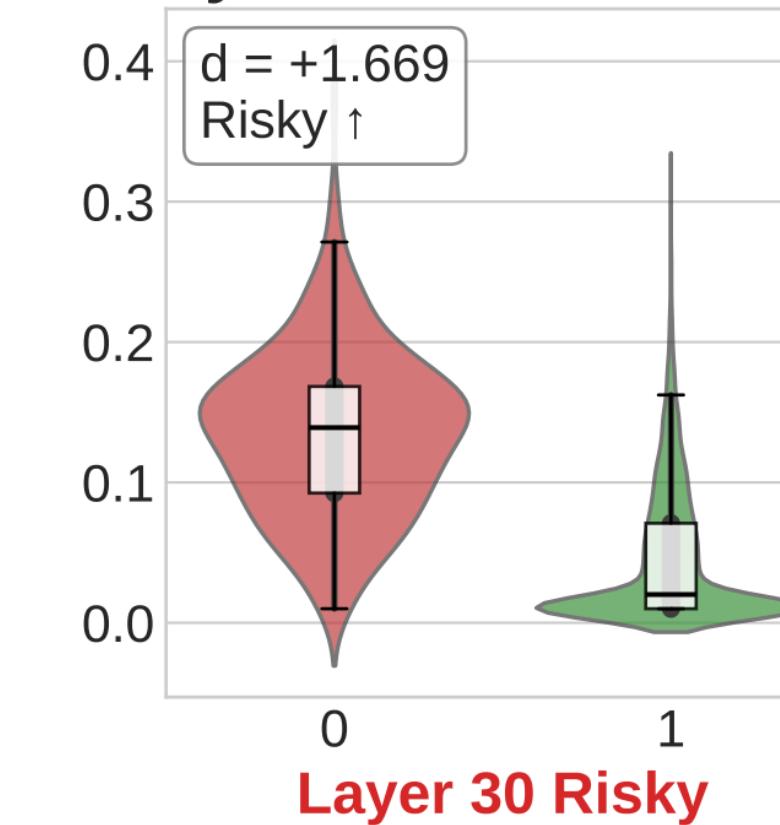
## Layer 28 Feature 25651



## Layer 28 Feature 18936



## Layer 30 Feature 16827



## Layer 30 Feature 18141

