Forecasts generated using protected time series change significantly from those that are generated using confidential data. While prior experiments have demonstrated severe degradations in forecast accuracy from a VAR model applied to differentially private time series, little is known about how data protection affects other forecasting models . We analyze the effects of several data protection methods (top and bottom coding, additive noise, differential privacy, and cluster-based swapping) on both simple and complex forecasting models.

Across shorter forecast horizons, we find that data protection generally degrades forecast accuracy.

At moderate levels of differential privacy, we find that simple models have better accuracy than complex models for all forecast horizons.

(Need a statement that makes more sense than the below for non-diffential privacy protection methods)

Under top coding, complex models have better accuracy than simple models for short horizons, and worse forecast accuracy than simple models for long horizons.

Under bottom coding, simple models outperform complex models for all horizons, even when the complex models have higher forecast accuracy using the original data.

First, we find that forecast accuracy under top and bottom coding improves by up to 13% relative to forecasts using the original data for long forecast horizons.

With data protection, we recommend using simpler forecasting models, such as exponential smoothing, which offer higher accuracy than complex forecasting models.

We investigate the drivers of these results and offer guidance for practitioners on selecting a forecast model under various data protection approaches.